Modelling Mountain Pine Beetle Population Dynamics and Management: A Case Study from the Cypress Hills

by

Mélodie Kunegel-Lion

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Ecology

Department of Biological Sciences

University of Alberta

Abstract

Outbreaks of insects are currently one of the main sources of disturbance in North American pine forests. The huge economic and ecological consequences of these outbreaks emphasize the need for effective pest management. For example, the mountain pine beetle (*Dendroctonus ponderosae*, Hopkins 1902) has killed more than half the commercial timber in British Columbia. Large-scale model predictions of mountain pine beetle have highlighted the impact of beetle pressure, weather, and stand characteristics on beetle location. Observational and experimental studies specify how these factors affect beetle presence during the course of its life-cycle. However, local-scale predictions could help inform pest management more accurately. Especially, there is a need to understand how the impacts of ecological and environmental factors on mountain pine beetle population change during the course of an outbreak. Population models typically incorporate management using 1) functional responses for biological control or 2) removal of a certain number of individuals at certain times for other types of control. This second method lacks a framework describing how management could be implemented in population models in a rigorous way. In this thesis, I use a combination of statistics, machine learning methods, simulations, and mathematical models to explore mountain pine beetle population dynamics and its management in the Cypress Hills interprovincial park. Boosted regression trees are able to predict accurately mountain pine beetle infested trees presence on a local scale using ecological and environmental variables. Logistic regressions using similar variables further demonstrate which factors affect beetle presence and how they vary depending on the outbreak phase. Virtual experiments of mountain pine beetle management show that managers are efficient in detecting infested trees using previous-year infestation information. Yet, efficiency could be further improved by the addition of weather, stand characteristics, and topography to inform detection. Using a mathematical model along with individual-based simulations, I demonstrate that we could apply the functional response framework to human-pest interactions in order to simulate population dynamics as well as compare management strategies. Simulations of a semi-empirical model describing the interaction between beetle and pine populations show that direct control affect beetle outbreak duration and pine mortality even at a low level and a moderate control can eradicate infestations in the long term. However, a significant control level is needed to achieve a quick suppression. Assessing management and considering ways of improvement are important as pest damages in North America are predicted to increase as a consequence of climate change.

Preface

This thesis is an original work by Mélodie Kunegel-Lion. Some of the research included in this thesis was conducted as part of collaborations detailed below. M.A. Lewis was the supervisory author in all chapters and was involved with concept formation and manuscript composition.

A version of Chapter 2 has been published as: Kunegel-Lion, M., McIntosh, R.L. & Lewis, M.A. 2019. Management assessment of mountain pine beetle infestation in Cypress Hills, SK. *Canadian Journal of Forest Research* 49: 154–163. R.L. McIntosh directed data collection and contributed to manuscript edits.

A version of Chapter 3 has been published as: Kunegel-Lion, M., Goodsman, D.W. & Lewis, M.A. 2019. When managers forage for pests: Implementing the functional response in pest management. *Ecological Modelling* 396: 59–73. D.W. Goodsman initiated the conceptual design and contributed to the mathematical structure section and manuscript edits.

A version of Chapter 4 has been submitted to *Scientific Reports* and is currently under review.

A version of Chapter 5 has been submitted to *Journal of Environmental Management* and is currently under review.

The data used in this thesis was provided by Rory L. McIntosh, Insect & Disease Expert at the Forest Service Branch of the Saskatchewan Ministry of Environment.

A la mémoire de mon père

Acknowledgements

Thank you to my supervisor, Dr. Mark Lewis, for giving me such a great opportunity to work in his lab and broaden my horizons in mathematical ecology. Mark, I have been extremely fortunate to have you as my academic mentor. Your guidance, support, and encouragements helped me challenge myself. I have learnt a lot about mathematical modelling, statistics, ethics, and good practice in science thanks to the stimulating discussions during lab retreats and weekly lab meetings that you organized.

Many thanks to my committee members, Drs. Nadir Erbilgin and Fangliang He, for taking the time to be on my supervisory committee and providing valuable feedback.

I thank the University of Alberta and the NSERC TRIA Network for the financial support I received. I further thank TRIA-Net for the enriching experiences during the yearly meetings.

I would like to thank the entire Lewis Research Group for all the feedback I received during lab talks, for the interesting discussions at lunch, and for the amazing support. Thanks to Devin Goodsman for starting a discussion about functional responses in pest management which eventually grew into chapter 3. I would like to express special thanks to Kim Wilke-Budinski for her incredible help with any administrative issue and more.

I would also like to thank the TRIA-Net research team for sharing and discussing their findings and knowledge on mountain pine beetle in such a cheerful and friendly way, helping me understand different perspectives and see the big picture.

I thank Brian Poniatowski and Rory McIntosh for providing the mountain pine beetle data set and for the useful discussions around technical and surveillance aspect in the Cypress Hills interprovincial park. Thanks to them, I had the wonderful opportunity of spending a couple of days in Cypress Hills watching mountain pine beetle specialists work, which help me develop my knowledge on mountain pine beetle and understand the challenges of its management.

Finally, I am grateful to my partner, Florian Barré-Minos. His enthusiasm about moving to and living in Canada helped me overcome all the challenges associated with it. And I am grateful to my family, especially my mother and my sister, for their incredible support and encouragements.

Contents

List of Tables

List of Figures

Chapter 1

General introduction

Forest insect outbreaks can have huge environmental and economic consequences. For instance, the European spruce beetle (*Ips typographus*) killed thousands of hectares of mature spruce trees in central Europe between 1992 and 2000 (Wermelinger, 2004). In North America, insect and pathogen outbreaks are currently one of the main forest disturbances. They cause, in some places, more damages than fires and intense climatic events such as storms (Dale *et al.*, 2001). For example, the eastern spruce budworm (*Choristoneura fumiferana*) is a defoliator of fir and spruce that has affected millions of square kilometres of forests (Venier & Holmes, 2010). The southern pine beetle (*Dendroctonus frontalis*) impact on commercial pine forests has cost timber producers more than one billion U.S. dollars in southern U.S. over the course of 28 years (Pye *et al.*, 2011). Similarly, the mountain pine beetle (MPB; *Dendroctonus ponderosae*, Hopkins 1902) outbreak occurring in western Canada pine forests since the early 2000s has killed more than half of the merchantable timber volume (Walton, 2013). Therefore, forest pest management is crucial and frequently implemented in order to decrease the repercussions on the environment and human communities (Fettig & Hilszczański, 2015).

Managers employ various control tactics in order to reduce the impact done to the forest (Fettig & Hilszczański, 2015). To prevent damages and reduce economic losses, managers can harvest stands of commercial trees with a high susceptibility for infestation before the pest can reach them. They can also decrease stand susceptibility, for example by thinning treatments or prescribed fires (Fettig *et al.*, 2007). On the other hand, to directly control existing pest, they can survey areas at risk and implement control methods that will ensure pest mortality such as biological control, felling and treatment of infested trees, or insecticide application (Fettig & Hilszczański, 2015). For example, the great spruce beetle (*Dendroctonus micans*) is managed using the predatory beetle *Rhizophagus grandis* in Europe and Asia (Fettig & Hilszczański, 2015) and MPB infested trees are often treated using a fall and burn tactic (Alberta Sustainable Resource Development, 2007; Saskatchewan Ministry of Environment, 2016). Determining areas at risk can be done using modelling tools such as statistics, machine learning methods, population models, or agent-based models.

By definition, an adequate pest management affects considerably pest population dynamics by reducing population growth and/or dispersal. Decreases in pest population size due to direct control is rarely constant over time and rather clustered to some time periods convenient for managers given the insect life-cycle. Similarly, the disruption of dispersal events may be limited to specific managed areas. The extent of changes in growth and dispersal may also differ depending on the details of the management strategy. Therefore, studies investigating pest population dynamics must take the specifics of management into account in order to have accurate results.

In this thesis, I will explore pest population dynamics and management in the context of a MPB outbreak using phenomenological and mechanistic models.

1.1 Ecology of the mountain pine beetle

The mountain pine beetle is an eruptive bark beetle infesting a large variety of pine species in western North America (Negrón & Fettig, 2014). In Canada, its main host expands from the lodgepole pine (*Pinus contorta*) in British Columbia to the jack pine (*Pinus banksiana*) east of the Rocky mountains (Little, 1971; Cullingham *et al.*, 2011). MPB is an aggressive species alternating low population size called an endemic phase and cyclic outbreaks called an epidemic phase (Safranyik, 2004). Endemic populations survive by attacking weak and stressed pines with the help of other bark beetle species and pathogens (Smith *et al.*, 2011). In favourable conditions, such as warm winter temperatures, MPB populations increase to epidemic levels where they are able to successfully mass attack large and vigorous pine trees. At the peak of an outbreak, epidemic populations have enough individuals to infest large patches of pine forests. Populations decline because of a lack of susceptible hosts or unfavourable weather conditions such as cold snaps in the fall and spring leading to high larval mortality.

MPB typically has a one-year life cycle although a two-year life cycle has been reported in colder regions (Logan & Powell, 2001) and some evidence of bivoltinism, two generations per year, has been observed in the United States (Mitton & Ferrenberg, 2012; but see Bentz & Powell, 2014). In the early summer, beetles disperse, and attack susceptible pine trees. Females lay eggs in galleries under the bark which develop into larvae over the fall and winter. Beetles overwinter as larvae and resume their development into pupae and adult in the spring (Safranyik & Carroll, 2006). Fungal and bacterial associates facilitate beetle nutrition in the host by providing sustenance and concentrating host tree nutrients (Safranyik *et al.*, 1975; Bleiker & Six, 2007; Therrien *et al.*, 2015). Beetle attack and development lead to the host death. Therefore, MPB-infested pines are recognizable by the red colour of their crown one year after the attack.

Although new MPB infestations often appear in the neighbourhood of previous infestations, beetles have been observed hundreds of kilometres away from the closest pine forests (Safranyik & Carroll, 2006). Indeed, MPB generally has a short-distance dispersal behaviour where individuals actively fly a few meters above ground within a stand in order to find and colonize new hosts (Carroll & Safranyik, 2004). However, when climatic conditions are favourable, some beetles can engage in a long-distance dispersal behaviour. This behaviour happens when ascending air currents bring beetles above the tree canopy where they get caught in the wind and thus, get transported hundreds or thousands of kilometres (Chen & Jackson, 2017).

Adult beetles usually have a synchronized 10-days emergence window between mid-July and mid-August during which females initiate host selection and colonization (Safranyik

& Carroll, 2006). The mechanism of MPB host selection is disputed. While some studies report that pioneer beetles use visual cues in this process, other support random landing (Billings *et al.*, 1976; Hynum & Berryman, 1980; Safranyik & Carroll, 2006). More recently, some evidence was found that MPB could use host volatiles to differentiate between host and non-host species (Gray *et al.*, 2015). After landing, the tree suitability is assessed using gustatory cues, *i.e.* cues related to taste, before starting colonization (Raffa & Berryman, 1982).

During the colonization process, beetles release aggregation pheromones in order to attract enough conspecifics to produce a mass attack and overwhelm the tree defences (Safranyik & Carroll, 2006). Indeed, MPB presents a strong Allee effect: a certain number of beetles is needed to overcome the tree defences, and below it, the attack fails (Allee, 1931; Goodsman *et al.*, 2016). Therefore, MPB success depends on host mortality, thereby inducing selective pressure for strong host defences (Raffa & Berryman, 1983). These tree defences consist of the exudation of toxic resin containing phytochemicals that prevent MPB from attracting conspecifics and inhibit the formation of galleries and oviposition (Raffa & Berryman, 1983; Erbilgin *et al.*, 2017a). These defences depend on the tree species and the co-evolution history with MPB (Rosenberger *et al.*, 2017; Erbilgin, 2019).

1.2 Mountain pine beetle management

The MPB outbreak that started in 1999 in western Canada has been the most destructive insect outbreak Canada has ever known. An increase in the number of susceptible hosts in addition to favourable climate, such as warmer temperatures, have permitted this outbreak to gain such an extent that more than half the commercial timber volume in British Columbia was killed by MPB (Walton, 2013) and MPB populations were able to cross the Rocky Mountains (de la Giroday *et al.*, 2012).

MPB, and bark beetles in general, have a complex response to climate change (Bentz *et al.*, 2010). For example, temperatures control developmental thresholds used to synchronize the emergence of adults at optimal times of the year. Similarly, temperatures play a role in the cold hardening process allowing larvae to survive the coldest months of the year. Therefore, changes in temperatures throughout the year can either disrupt or favour seasonality as well as larval survival. Overall, modelling results show that MPB risk will increase over the century although it will stay relatively low in jack pine forests in Canada (Bentz *et al.*, 2010).

MPB outbreaks can have a long-lasting effect on biodiversity and communities (Chan-McLeod, 2006; Saab *et al.*, 2014). Outbreaks can kill entire pine stands in one generation. The eventual defoliation of the tree canopy, loss of tree cover, loss of living bark, and decrease in cone production following pine mortality negatively impact some of the invertebrates, mammals, and birds using pine forests as a food source, a foraging habitat, a resting or nesting habitat, or a shelter from predators. Those changes are reflected at the landscape level as the loss of trees gives opportunity for understorey species to develop. Local increases in species preferring open habitats will critically change communities and potentially lead to eradication of some mature pine forests in favour of open habitats (Chan-McLeod, 2006). However, dead trees provide fuel for fires which can help pine recolonization by favouring cone opening and prevent the loss of forest systems (Lotan *et al.*, 1985).

Rural communities are sensitive to these landscape-level changes. These communities are mainly affected by MPB through the loss of timber (Dhar *et al.*, 2016). However, the destruction of pine forests disrupt other provisioning services, such as water supplies and food, as well as water quality. As a result, forest-dependent communities, relying on these ecosystem services, might be heavily affected. In addition, potential negative effects on landscape aesthetics could alter tourism and recreation activities although it requires further research (Dhar *et al.*, 2016).

Cypress Hills, an interprovincial park located at the limit between the provinces of Alberta and Saskatchewan, is the most eastern MPB infestation in Canada (Safranyik & Carroll, 2006). In the park, covering 400 km^2, the main pine species is the lodgepole pine. This pine population is isolated from the main lodgepole and jack pine ranges (Cullingham *et al.*, 2012). The closest MPB outbreaks are south of the park in the

United States and west in British Columbia and Alberta. Due to the long-distance dispersal ability of the MPB and the dominant wind coming from the southwest, Cypress Hills could become a stepping stone for MPB infestations further north and east. Controlling MPB in Cypress Hills is therefore crucial.

The first discovery of MPB in Cypress Hills was in 1979 (Ono, 2004), followed by the first reported outbreak in the 1980s (Taylor *et al.*, 2006). Between 1983 and 1987, managers implemented a semiochemical baiting strategy (Cerezke, 1989) involving MPB pheromones and pine monoterpenes. These baits were found to be a good monitoring tool. Furthermore, they helped gather beetles in easy-access locations for sanitation cutting and burns. The extreme winter temperatures in 1984-1985 helped control the outbreak to lower population levels.

In 2006, a new outbreak started in Cypress Hills. It peaked in 2012-2013 and continues today in what seems to be the outbreak decline. In this thesis, I study this outbreak in the Saskatchewan portion of the park. There, and since 2006, the Saskatchewan Forest Service implemented a "zero-tolerance" policy designed to find and control as many infested pines as possible (Saskatchewan Ministry of Environment, 2016). This strategy utilizes aerial and ground surveys to detect currently infested trees. During aerial surveys of the entire Saskatchewan portion of the park, red-top trees (*i.e.* pines killed by MPB the previous year and displaying a red crown) are geolocalized. The cause of their mortality is later confirmed on the ground and circular surveys with a radius of 50 meters are implemented around them to search for newly infested pines. In locations with high densities of red-top trees, the entire area is checked using line surveys. All the infested trees detected are controlled for MPB, by either peeling the bark or using a fall and burn tactic. Both methods ensure that any live brood is killed. Undetected infested trees become red the following year. Therefore the location of all infested trees can be either detected during surveys or back-estimated from the red-top trees. This management policy provides an extremely detailed data set in time and space of a managed outbreak from start to decline.

1.3 Pest management in population models

Pest management can be divided into preventive measures, such as habitat modification to prevent growth or spread, and direct control lowering the number or density of pest individuals (Carroll *et al.*, 2006; Epanchin-Niell & Hastings, 2010). In this thesis, I focus on the direct control of MPB individuals.

They are not an established frameworks to model direct pest management in population dynamic models. Pest populations are, by necessity, controlled. Researchers have focused on questions such as *when*, *where* and *how much* in the context of optimal control (Epanchin-Niell & Hastings, 2010). However, the question of how to include pest management in a population model is less studied. A common way of including control in a model is to remove a certain number of individuals at certain times and locations (*e.g.* Liu *et al.*, 2006; He *et al.*, 2012). It can also be incorporated as a parameter changing the growth or spread of the pest population (Saphores & Shogren, 2005). Moreover, when management is implemented via biological control, the impact of a predator on a pest can be described by functional responses (*e.g.* Janssen & Sabelis, 1992; Mills & Getz, 1996; Van Den Berg *et al.*, 1997). As a novel way of incorporating pest management in population models, I introduce in this thesis the concept of human-pest functional responses based on the predator-prey functional response framework and where managers take the role of predators foraging on pest populations.

Functional responses describe the interaction between prey and predators by indicating how many prey can be consumed by a predator in a certain time given the prey density in the environment (Holling, 1959a). There are three main types of functional responses: a linear increase (type I), a curvilinear increase (type II), and a sigmoid curve (type III). An infinite linear increase of the number of prey consumed by a predator given the prey density is considered unrealistic. Instead, a type I functional response is interpreted as follows. As prey density increases, the number of prey consumed linearly increases until it reaches a sudden and complete saturation at high prey density, preventing the predator for consuming more prey than is possible in a given time (Jeschke *et al.*, 2004).

This functional response is characteristic of filter feeders. A type II functional response characterizes predators whose handling or digesting time progressively prevents them from getting more prey at high prey density (Jeschke *et al.*, 2002). In a type III functional response, some mechanisms make it harder for the predator to find prey at low prey density inducing a sigmoid shape (Hassell, 1978). In the context of resource management, and although they do not refer to it this way, fisheries models often include a linear type I functional response to describe the effect of harvest on fish populations whether it is a constant rate of harvest or depending on the fish density (Sutherland, 2001).

1.4 Modelling tools

Models in ecology are usually divided into phenomenological and mechanistic models. Phenomenological methods test for relationships between a response variable and predictors whereas the goal of mechanistic models is to describe the processes behind such relationships. In my thesis, I use both approaches in order to explore MPB population dynamics and its management from different perspectives.

Many studies have developed population models describing the mechanisms of MPB population dynamics. In particular, partial differential equations (*e.g.* Logan *et al.*, 1998; Strohm *et al.*, 2016) and integro-difference equations (*e.g.* Lewis *et al.*, 2010; Rodrigues *et al.*, 2015; Goodsman *et al.*, 2016) have been used to describe the spatial and temporal dynamics of MPB populations, continuously and discretely in time respectively. In stage-structured models, each species stage is modelled with a different equation in order to model, for example, beetle dispersal and attack dynamics in detail (*e.g.* Heavilin & Powell, 2008). in contrast, agent-based models are a way to simulate individual or subgroup behaviour using a set of rules to obtain the overall population behaviour instead of mathematical equations describing directly the population dynamics (*e.g.* Riel *et al.*, 2004; Nelson *et al.*, 2018). When management is included in population models, it is typically done as a separate module impacting population sizes at a specific point during a time step (*e.g.* Strohm *et al.*, 2016).

On the other hand, statistics, such as logistic regressions, are typically used to make spatial predictions of MPB infestations from empirical data (*e.g.* Aukema *et al.*, 2008; Preisler *et al.*, 2012; Sambaraju *et al.*, 2012). Lately, machine learning algorithms, such as classification and regression trees, are more and more used for this purpose (*e.g.* Sidder *et al.*, 2016). Spatial predictions are especially useful to inform MPB management by indicating future areas at high risk of infestation, thereby allowing improved detection or the implementation of effective preventive measures.

Machine learning methods are particularly useful in ecology. Decision tree ensembles, for example gradient boosting and random forest, are particularly common. They typically perform better than generalized linear models (Marmion *et al.*, 2009; Youssef *et al.*, 2016). Although inference is inherently weakened with machine learning, such methods are able to produce highly accurate predictions. Indeed, machine learning methods are not within the framework of hypothesis testing. The hypothesis testing approach involves carefully writing a statistical model to quantitatively describe a scientific hypothesis and testing this model against a null model or simultaneously against alternative hypotheses. Instead, machine learning methods flexibly categorizes outcomes in the parameter space. Therefore, the choice of the type of analysis depends on the study goal.

1.5 Dissertation outline

In this thesis, I use the modelling tools described above to answer questions related to MPB population dynamics and its management. In chapter 2, I use a boosted regression tree to make predictions of MPB locations a year ahead using ecological and environmental variables. Then, I use virtual experiments to assess the detection strategy implemented by the Forest Service in Saskatchewan by comparing it to a random strategy and a strategy informed by the model predictions of the first part. I also explore the option of increasing the survey area and study the changes in management costs associated with it. In chapter 3, I show that pest management activities can be modelled using the predator-prey framework. In particular, we can draw functional responses describing the relationship between number of pests controlled and pest density and obtain a correspondence with the three types of predator-prey functional responses described in Holling (1959a). Using an application to MPB and simulations, I show that this can help assess control efficacy and the feasibility of management goals. In chapter 4, I use logistic regressions to explore how the environmental and ecological factors driving MPB infestations differ depending on the outbreak phase. In chapter 5, I build a semi-empirical model where I include mechanisms of MPB dispersal and growth. Then, I simulate various control effort to explore the effect of MPB management on pine mortality and outbreak suppression. Finally, in the general conclusion, I discuss the lessons learnt for MPB population dynamics and management.

Chapter 2

Management assessment of mountain pine beetle infestation in Cypress Hills, Saskatchewan

2.1 Introduction

The mountain pine beetle (MPB; *Dendroctonus ponderosae*, Hopkins 1902) epidemic has caused extensive mortality in North American pine forests, which is in conflict with human objectives in many places. At a large scale, the epidemic is linked to climate change as well as population dynamics that shift intermittently between endemic and epidemic states (Carroll *et al.*, 2004; Shore *et al.*, 2006; Raffa *et al.*, 2008; Preisler *et al.*, 2012). MPB's spread is unaffected by most environmental barriers such as low mountain ranges and fragmented forests due to its ability to disperse long distances (de la Giroday *et al.*, 2012; Bentz *et al.*, 2016). To better control MPB populations, we need to determine areas at risk and assess the efficiency of current detection strategies.

The MPB is a bark beetle that infests and kills various species of pines. In western North America, lodgepole pine (*Pinus contorta*, Dougl. ex Loud. var. *contorta* Engelm) is the primary MPB host although MPB is a threat to almost all pine species (Safranyik & Carroll, 2006). During an epidemic, beetles coordinate their attacks, using aggregation pheromones, to form a "mass attack" and overwhelm the defences of large and healthy trees (Bordon, 1982). Therefore, an epidemic population of MPB presents a threat to healthy pine stands.

The MPB is primarly univoltine, meaning that a new generation is produced over a year (see Mitton & Ferrenberg, 2012; Bentz & Powell, 2014; Mitton & Ferrenberg, 2014). In summer, the beetles disperse and reproduce, and the females lay eggs in galleries they excavate under the bark. Beetles usually overwinter as larvae. In spring, they resume their development and finally emerge as adults later in the summer (Safranyik & Carroll, 2006). Trees are seriously injured by the gallery excavation process and the development of MPB larvae and their associated blue stain fungi, and generally die and turn red by the end of the MPB life-cycle. During the following years, attacked trees become grey. As a result, red-top trees, infested during the summer of the previous year are easily spotted during aerial surveys of stands, becoming a good proxy for the status of the previous year's MPB infestation levels.

At a landscape level, two types of dispersal strategies have been observed for MPB (Safranyik & Carroll, 2006; Robertson *et al.*, 2007): long-distance dispersal, passive downwind flight over the canopy, and short-distance dispersal, active flight a few meters above ground. Researchers estimate the short-distance dispersal range to be within a stand (Safranyik & Carroll, 2006) at the order of 20 to 50 meters, although some beetles can go as far as 100 meters (Robertson *et al.*, 2007). In contrast, long-distance dispersal range is tens to hundreds of kilometres (Safranyik & Carroll, 2006; Jackson *et al.*, 2008). While short-distance dispersal is much more common than long-distance dispersal (Safranyik *et al.*, 1989; Chen & Walton, 2011), the MPB's epidemic behaviour associated with outbreaks arising from long-distance dispersal can pose a threat to entire regions of pine forests.

In Canada, since 2006, a local MPB epidemic has emerged in the Cypress Hills area, located in the southwest of Saskatchewan and southeast of Alberta. The Cypress Hills inter-provincial park comprises the West Block, divided between Alberta (219 km^2) and Saskatchewan (126 km^2), and the Center Block, in Saskatchewan (58 km^2). For the purpose of this paper, our study focuses on the Saskatchewan portion of the park. Therefore the use of "the park" and "Cypress Hills" in the text refers to the Saskatchewan portion. The local MPB population is endemic to the park and probably came from southern populations in Montana, USA (R. L. McIntosh, pers. comm.). It could have been partly sustained by beetle flights from the south and west. Indeed, during spring and summer, during MPB dispersal, the dominant wind comes from the southwest.

Studying and controlling MPB in the Cypress Hills area is essential for two reasons. First, as an inter-provincial park and national heritage, Cypress Hills has significant natural, economic, and cultural values. Second, even though this park is somewhat isolated compared to lodgepole and jack pine ranges (Little, 1971; Cullingham *et al.*, 2012), the presence of a MPB epidemic, in association with the long-distance dispersal ability of the insect and the wind direction, makes the Cypress Hills area a possible stepping-stone facilitating the infestation of the remainder of Saskatchewan and regions further east. Therefore, there is an urgent need for analysis of management and for infestation prediction in Cypress Hills.

Aware of the need for management, the Forest Service Branch of the Saskatchewan Ministry of Environment has implemented a "zero-tolerance" policy designed to catch and control as many short-distance infestations as possible. This requires intensive surveillance to implement early detection and rapid aggressive response actions. The policy operates according to the following procedure. In early fall, after MPB have colonized new trees, an aerial survey of the park extent is conducted to collect georeferenced data on potential red-top trees, which are dead or dying trees infested by MPB the previous year. These are later ground-truthed for MPB attacks. Then, 50 meter-radius circular survey plots are drawn around each of the red-top trees confirmed to have been killed by MPB. The survey plots are searched for green attacked trees, which are trees recently infested by MPB during the summer. These are later controlled in late fall/winter which usually consists of felling and burning massively infested trees, ensuring that beetles are killed. The survey plot can be spatially extended if green infestations are spotted close to the plot's limits (Saskatchewan Ministry of Environment, 2016). In addition to these measures, areas presenting high densities of red-top trees are entirely surveyed and controlled. No detected infestations are left untreated. Such intensive control is expensive. Therefore, there is a need to determine how effectively this strategy is working.

Given this management strategy and the MPB context in Canada, our study aims to answer the following question: Are there ways to improve detection strategies without increasing management costs? If managers completely removed infested trees coming from MPB short-distance dispersal inside the park, the remaining source of infestation would be due to long-distance dispersal events from outside the park which are often considered spatially random when observed at a small scale (Long *et al.*, 2012; Powell *et al.*, 2018). Therefore, we hypothesize that a random search would be as efficient as a local search around red-top trees. Moreover, we hypothesize that, if factors others than distance to previous infestations influence the location of new infestations, then a search based on predictions from such factors would be more efficient than a local search around red-top trees. However, the management survey might not be big enough to include all infestations from short-distance dispersal events. Therefore, we make the third hypothesis that, as the search area increases, the detection efficiency will increase too.

2.2 Material and methods

2.2.1 MPB predictions

To predict MPB infestation a year ahead in Cypress Hills, we used the generalized boosted classification model which is a machine learning algorithm. Boosted classification trees generate results with an excellent fit for a binary response by successively fitting a tree to the previous tree's residuals to reduce the final error variance significantly (StatSoft Inc., 2013).

Data

The covariates and response variable values were distributed discretely in space and time. We applied a grid of 18 317 cells of size 100m×100m to the Cypress Hills park extent. For each cell for each year, the observation consisted of a set of environmental and ecological covariates plus the response variable. The response variable was the presence/absence of MPB derived from the presence/absence of green attacked trees in each cell of the grid based on data from the Forest Service ground survey. From the Forest Service surveys, we got the locations of green infestations controlled by managers and we deduced which trees had been attacked in the previous year using the red-top trees.

We used 14 covariates related to topography, weather, vegetation, and beetle pressure (Table 2.1). The weather variables were: the highest maximum daily temperature over the year, the overwinter survival probability of the larvae (Régnière & Bentz, 2007), and the average daily relative humidity in spring. Indeed, MPB dispersal is reduced with high temperatures (Safranyik & Carroll, 2006). The minimum temperatures in fall and winter impact MPB survival if the vulnerable stages–developing in the fall and at the end of the winter–are exposed to extreme temperatures (Cole, 1981; Safranyik & Carroll, 2006; Régnière & Bentz, 2007). Drought in the spring reduces pines' ability to defend themselves and increase MPB attacks' success rate (Safranyik, 1978; Lusebrink *et al.*, 2016). Additionally, MPB individuals need at least 833 degree-days above 5.5°C over a year to complete their growth (Safranyik *et al.*, 1975; Carroll *et al.*, 2006; Safranyik *et al.*, 2010). In the park, over the time period studied, the minimum number of degree-days above 5.5°C was 923, which is above the threshold and so degree-days was not included in our model. Furthermore, high numbers of degree-days are not an issue as MPB rarely present multivoltinism (Bentz & Powell, 2014). We included the MPB presence at the same location and in the neighbourhood the year before in order to take into account the spatio-temporal autocorrelation of the data (Fig. 2.1). The beetle pressure from outside the park was represented by the distance to the park southern border (illustrated on Fig. 2.2) which was close to external infestations not managed by

Table 2.1 – Description and range of the covariates used in the generalized boosted classification model.

Name	Description	Range	Unit
PineCover	Coverage of *Pinus albicaulis* (Whitebark Pine), *Pinus banksiana* (jack pine) and *Pinus contorta* (includes subspecies lodgepole pine and shore pine)	0 – 76.1	%
TMax	The highest maximum daily temperature from September of the previous year to August	27.3 – 36.7	°C
OWS	The overwinter survival probabilities of larvae (Régnière & Bentz, 2007) using a 5-year lookback	0.23 – 0.50	–
RH	Average daily relative humidity in spring	56.9 – 73.8	%
BP0	Presence of previous year mountain pine beetle infestation in the focus cell	0/1	–
BPn	Previous year mountain pine beetle pressure in the neighbouring cells: BPn = $\sum$ BP0 in adjacent cells of radius 1 + 0.5 × $\sum$ BP0 in adjacent cells of radius 2 + 0.25 × $\sum$ BP0 in adjacent cells of radius 3 (Fig. 2.1)	0 – 9.25	–
DistSouth	Distance from the grid cell centroid to the South infested border of the park	5 – 36660	m
Latitude	Latitude of the grid cell centroid	49.55 – 49.61	dec. °
Longitude	Longitude of the grid cell centroid	-110.01 – -109.43	dec. °
Year	Year of the survey	2007 – 2015	–
Elevation	Elevation at the grid cell centroid	1055 – 1386	m
Slope	Slope at the grid cell centroid	0 – 20.31	°
Northerness	Tendency of the slope to face North	+1 – -1	–
Easterness	Tendency of the slope to face East	+1 – -1	–

the Forest Service and potential sources of MPB. The rest of the variables included in the model were: pine cover, latitude, longitude, year, elevation, slope, and northerness and easterness derived from the aspect.

Topography data came from the Canadian Digital Elevation Map downloaded from the Geogratis website (geogratis.cgdi.gc.ca). We generated weather variables with the BioSIM software (Régnière *et al.*, 2014) at the location of each grid cell centroid. BioSIM uses data from surrounding weather stations and interpolates the weather variable values at each location of interest using a digital elevation map. The vegetation data came from Beaudoin *et al.* (2014). The authors computed these data from 2001 MODIS imagery, and the vegetation parameters were assumed constant over our time period.

We used data from the years 2007 to 2015. Randomly, we chose 75% of these data, years combined, *i.e.* 149 278 observations, to train the model. The remaining 25%, 49 502 observations, were used to validate the model.

Generalized Boosted Model

We trained the generalized boosted classification model using the `gbm` function of the `R` package `gbm` (Ridgeway, 2015) on the 14 covariates in the training set. The process analyzed the performance of 50 000 classification trees and performed a 10-fold cross-validation in order to find the best classifier. The algorithm implemented in the `gbm` function consisted of reducing a loss function between the observed and the predicted response values using Friedman's Gradient Boosting Machine (Ridgeway, 2015). The loss function was represented by a Bernoulli error distribution, which is adapted to a binary response. The `gbm` function output provides the probability of MPB presence at each location. We tested the accuracy of the model's prediction using the area under the receiver operating characteristic curve (AUC; Metz, 1978; Bradley, 1997), the false positive and false negative rates, and the misclassification rate which is the percentage of misclassified instances by the model. A receiver operating characteristic (ROC) curve (Metz, 1978) depicts, for a range of probability thresholds, the true positive rate (or 1 - false negative rate, also referred to as sensitivity) against the false positive rate (also

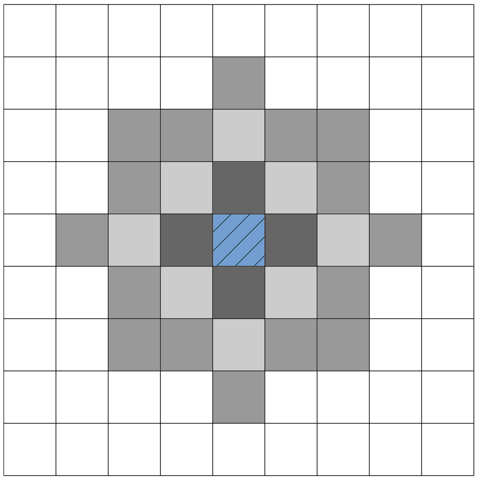

Figure 2.1 – Representation of the adjacent cells taken into account in the covariates (*cf.* Table 2.1). Striped blue: focus cell, dark grey: 4 adjacent cells (radius 1), light grey: next 8 adjacent cells (radius 2), medium grey: next 16 adjacent cells (radius 3).

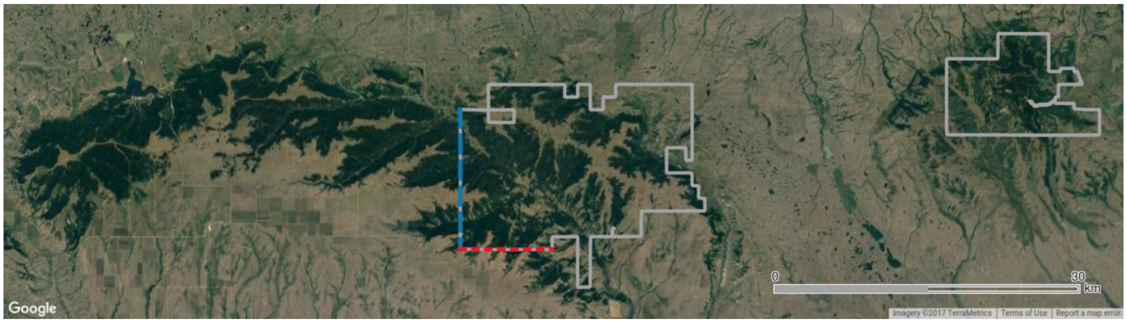

Figure 2.2 – Cypress Hills park boundaries in Saskatchewan (grey). The dotted red line represents the park border close to outside infestations in the South. The dashed blue line represents the park border with Alberta.

referred to as 1 - specificity). We used Youden's method (Youden, 1950) to determine the probabilities threshold which selects the farthest point from the diagonal on the ROC curve. A high AUC ($0 \leq \text{AUC} \leq 1$) represents a good performance of a binary classifier in terms of correspondence between observed and predicted values.

2.2.2 Assessing management

Data

To assess the detection strategies, we needed the exact locations of red-top trees for a focus year and the following year. In 2011 to 2013, the data from the Forest Service included an exhaustive survey of red-top trees' locations and the number of green infestations controlled around each red-top tree. The other years included infested areas in which red-top trees' locations were not specified. For this reason, we only used data from 2011 and 2012 for this analysis. Furthermore, the years 2011 and 2012 happened to have a similar number of red-top trees/survey plots: 292 for 2011 and 284 for 2012, which made the two years comparable.

For controlled green infestations, we used the location of the circular plot centres (±50 meters compared to the real locations of green infestations). For uncontrolled green infestations outside of survey plots, we used the location of red-top trees the year after. The total number of green infestations was 644 for 2011 and 936 for 2012.

Simulated detection strategies

To calculate the efficiency of the detection strategies, we simulated virtual experiments. For each year, we counted the number of green infestations in increasing virtual survey areas for three different strategies: 1) local search in circular plots of varying radius around red-top trees (similar to the current Forest Service strategy), 2) search in circular plots of varying radius randomly located in space, and 3) search in a varying number of 100×100m square plots placed at locations predicted by the boosted classification tree. In the predictions strategy, we used 100×100m square plots and not circular plots to match as much as possible the predicted locations from the classification tree. For the local and random searches, we used circular plots of increasing radius: from 50 to 100 meters by increment of 5, from 110 to 150 meters by increment of 10, 200, and 300 meters.

To be able to compare similar survey areas among detection strategies, we needed to be able to fix the number of search locations, and therefore the search area, from the classification tree output. We could simply select a certain number of locations with the highest probabilities. However, if the number of selected locations is small like is the case here, some locations with relatively high probabilities might not be chosen whereas locations with slightly higher probabilities due to random noise will. To bypass this issue, we introduced some noise by randomly sampling the locations using the model probabilities to the power of 3 as weight. We investigated the impact of variation in this exponent value in Appendix A. For the random and prediction strategies, we performed 500 simulations for each year.

Control efficiency

We calculated control efficiency for each year for each survey area with the equation

$$\text{control efficiency} = \frac{\#\text{ green infestations controlled}}{\text{total }\#\text{ green infestations in the park}}. \tag{2.1}$$

From the area controlled (*i.e.* the sum of every survey plot area), we obtained the net survey area by removing the overlapping areas. For each year,

$$\text{net survey area} = \begin{cases} \#\text{ plots} \times \pi r^2 - \text{overlaps} & \text{for local/random} \\ \#\text{ square plots} \times 100^2 & \text{for predictions} \end{cases}. \tag{2.2}$$

We then determined the relationship between net survey area and control efficiency. This was achieved by fitting a non-linear function, using the `nls` function of the `R` package `stats`, to control efficiency versus net survey area in the two cases: local search around red-top trees, local control efficiency $= f_{\text{local}}$(net survey area), and model predictions strategy, prediction control efficiency $= f_{\text{prediction}}$(net survey area). For the random search case, we fitted a linear function using the `lm` function of the `R` package `stats`: random control efficiency $= f_{\text{random}}(\text{net survey area})$.

Management cost

To determine cost-effective recommendations for managers, we also examined the relationship between net survey area and management cost. The management cost variable included the cost of aerial survey, the cost of control, and the cost of surveying all non-overlapping 50 meter-radius circular plots. It was available for the years 2010 to 2015. Within each year, the cost per unit (control cost per tree and survey cost per plot) did not vary depending on the location. However, since the cost per unit varied among years due to economic fluctuations, we took the median cost per unit over the years 2010 to 2015 and multiplied it for each year by the number of units in each category (number of controlled trees and circular plots per year). Thus, for each year:

$$\begin{aligned}\text{management cost} = \; & \text{median aerial survey cost} \\ & + \text{median control cost per tree} \times \#\ \text{trees controlled} \\ & + \text{median circular plot survey's cost} \times \#\ \text{plots.}\end{aligned} \tag{2.3}$$

The number of units in each category was available for the years 2006 to 2015. Therefore, we determined management cost values for 2006 to 2015. As a result, although total cost did vary year to year, the cost per plot and per tree did not. We fitted a linear regression line to the relationship between management cost and total area surveyed with circular plots (management cost $= g$(total area surveyed with circular plots) where $g(.)$ is a straight line function) using the `lm` function of the `R` package `stats`. The total area surveyed with circular plots does not contain overlaps (Saskatchewan Ministry of Environment, 2016) so this is equal to the net survey area with radius $= 50$ (Eq. 2.2). To get to the next step, we assumed that the management cost increases proportionally with the plot area. Thus, the cost of the total area from several survey plots is equal to the cost of the area of a single much larger survey plot. Hence, management cost $= g$(total area surveyed with circular plots) became management cost $= g$(net survey area). We then defined the “management cost per controlled tree” which is the management cost divided by the control efficiency for one year. Note that this cost per controlled tree is scaled by the total number of infestations in the park for each year. We explored the relationship between management cost per controlled tree and net survey area using the two regression equations: control efficiency $= f(\text{net survey area})$ and management cost $= g(\text{net survey area})$:

$$\begin{aligned}\text{management cost per controlled tree} &= \frac{\text{management cost}}{\text{control efficiency}}\\ &= \frac{g(\text{net survey area})}{f(\text{net survey area})}.\end{aligned}\tag{2.4}$$

The net survey area value corresponding to the minimum management cost per controlled tree would be the optimal area to survey.

However, one could also assign a cost θ to a missed green infestation as it would lead to several green infestations the following year. The cost of a missed green infestation θ times the number of missed green infestations is the avoided cost as it is the amount that would be saved in the future if these trees were actually controlled instead of being missed. In other words, θ is the marginal cost added to the following year cost if one green infestation is left and produces new infestations. Therefore, the total cost was defined as

$$\begin{aligned}\text{total cost} &= \text{management cost} + \text{avoided cost}\\ &= \text{management cost} + \theta \times \#\text{ missed infestations}.\end{aligned}\tag{2.5}$$

Thus, the total cost per controlled tree is the management cost plus the avoided cost divided by the control efficiency. Again, note that this cost per controlled tree is scaled by the total number of green infestations for each year. We then compared the optimal survey area for the management cost and for the total cost depending on the strategy used. We also investigated the dependence of the optimal survey area on θ in Appendix B.

2.3 Results

2.3.1 MPB predictions

The generalized boosted classification model has a good predictive ability (Fig. 2.3): the AUC value is 0.927. The probability threshold chosen from Youden's index is 0.003, which means that it is optimal in terms of misclassified instances to consider any probability value above this threshold as an infestation. Using this threshold, we calculated the confusion matrix (Table 2.2). The false negative and false positive rates calculated from it are, respectively, 0.187 and 0.118, which means that 18.7% of the infested locations are wrongly classified as non-infested and 11.8% of the non-infested locations are wrongly classified as infested. Additionally, the misclassification rate was 0.119 which means that 11.9% of the model results were misclassified compared to the observations.

We calculated the variables' impact on the classification tree output (*i.e.* relative importance). The MPB presence in the same location the year before is the most important variable (relative importance = 0.60), followed by the MPB pressure from neighbouring cells (0.26), the distance to the southern infested border of the park (0.10), and the overwinter survival (0.02). The remaining variables have each a relative importance below 0.01.

2.3.2 Assessing management

When increasing the radius of the circular plots or the number of square plots, and thus the area surveyed, the control efficiency increases and saturates for the local and predictions strategies (Fig. 2.4). The control efficiency of the search around random locations increases linearly with the net survey area. The local and predictions strategies are more efficient than the random search. For example, the local search reaches between 55.9% and 71.2% control efficiency at a 50-meters radius (current strategy), the predictions strategy between 54.3% and 63.3%, whereas it reaches only 0.01% con-

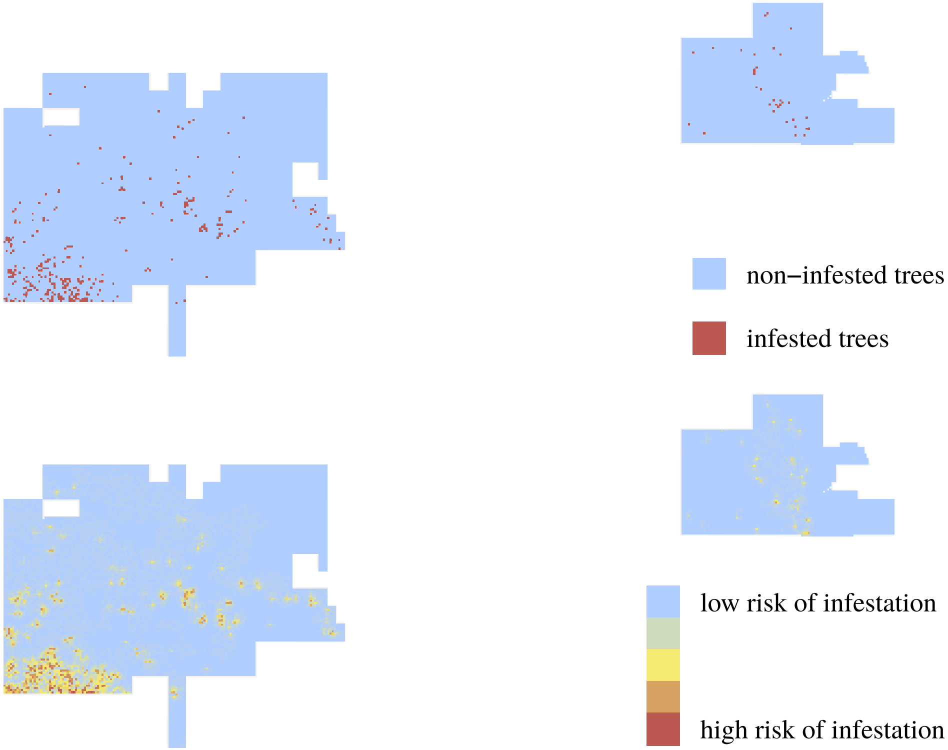

Figure 2.3 – Observations (top) versus predictions (bottom) of the mountain pine beetle infestation in Cypress Hills, Saskatchewan, for 2011. On a), a dark red color represents cells with infested trees whereas a blue color represents cells without infested trees. For b), the risk of infestation per cell ranges from blue (low risk) to dark red (high risk).

Table 2.2 – Confusion matrix showing the results of the model classification on the validation data set ($n = 49502$) using the threshold 0.003 chosen using the Youden's index.

		Observed	
		absence	presence
Predicted	absence	43 059	129
	presence	5 752	562

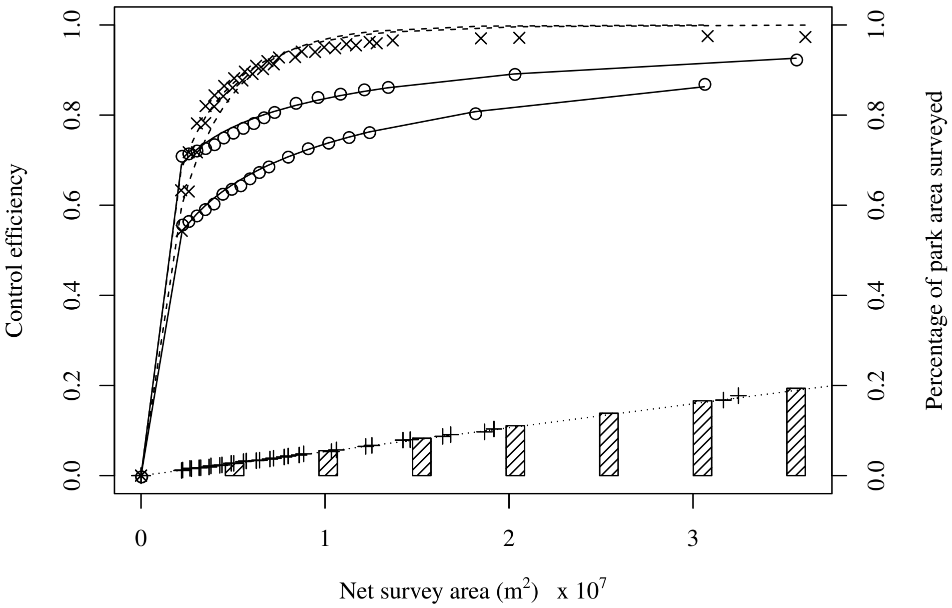

Figure 2.4 – Management control efficiency (= number of infested trees controlled in the park divided by the total number of infested trees) in relation to the net survey area (= total area controlled without overlaps). Solid lines and circles represent the local search around red-top trees for 2011 and 2012. Dashed lines and crosses represent the search at locations chosen from predictions for 2011 and 2012. Dotted lines and pluses represent the search around random locations for 2011 and 2012 combined. Each year, the random and prediction strategies data are each the mean of 500 random simulations. The lines represent the fitted values for the local and prediction strategy using a non-linear least square model: control efficiency $_{\text{local}}$ = $1-\exp(-a*\text{net survey area}^b)$ and control efficiency $_{\text{predictions}} = 1-\exp(-c*\text{net survey area}^d)$, where $a_{2011} = 0.004$, $b_{2011} = 0.358$, $a_{2012} = 0.018$ and $b_{2012} = 0.287$ (P-values < 0.001 for the null hypotheses $a = 0$ and $b = 1$, df = 17) for the local search, $c_{2011} = 2.25^{-6}$, $d_{2011} = 0.884$, $c_{2012} = 3.65^{-5}$ and $d_{2012} = 0.709$ (P-values = 0.309 and 0.164 respectively for the null hypotheses c_{2011} and $c_{2012} = 0$, and P-values < 0.001 for the null hypotheses d_{2011} and $d_{2012} = 1$, df = 17) for the predictions strategies. For the random search, we used a linear regression: control efficiency $_{\text{random}} = e * \text{net survey area}$, if net survey area $\leq$ park area or 1 if net survey area $>$ park area, where $e = 5.31^{-9}$ (P-value < 0.001 for the null hypothesis $e = 0$, $R^2 = 0.999$, df = 37). The striped bars represent the percentage of park area covered by the survey.

trol efficiency for the random search at the same survey area. For survey areas larger than those in the current strategy ($\sim$ 2 200 000 m^2), the predictions control efficiency is higher than the local control efficiency (Fig. 2.4). For example, for a survey area corresponding to 70-meters radius for the local search ($\sim$ 3 900 000 m^2), the control efficiency is 60.6% to 73.7% for the local search and 81.9% to 84.4% for the predictions strategy.

The management cost increases linearly with the net survey area (Fig. 2.5). We numerically obtain the net survey area values corresponding to the minimum management cost per controlled tree over the extent of net survey area values studied for the local and predictions strategies for 2011 and 2012: 2 178 332 to 2 225 780 m^2 (Fig. 2.6a). We obtain the matching radius 50 meters using Eq. 2.2 for the local search. However, it is highly probable that the cost of missing a green infestation θ is non-negligible. As the management cost increases with the survey area and the avoided cost decreases, the total cost shows a minimum value larger than zero (Fig. 2.7 for $\theta = 1000$). Therefore, the minimum total cost per controlled tree with $\theta = 1000$ gives survey area values ranging from 3 010 378 to 5 062 968 m^2 and corresponding to the radius 60 to 65 meters using Eq. 2.2 for the local search (Fig. 2.6b).

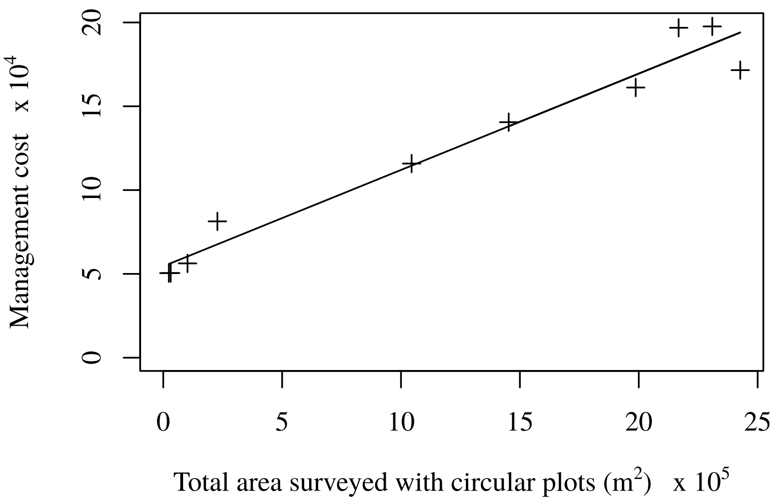

Figure 2.5 – Cost of aerial survey, control and circular survey plots in relation to the total area surveyed using circular survey plots from 2006 to 2015. The line represent the fitted values using a linear regression: management cost $= k + l *$ net survey area, where $k = 54\ 540.00$ and $l = 0.057$ (P-values < 0.001 for the null hypotheses $k = 0$ and $l = 0$, $R^2 = 0.961$, df $= 8$).

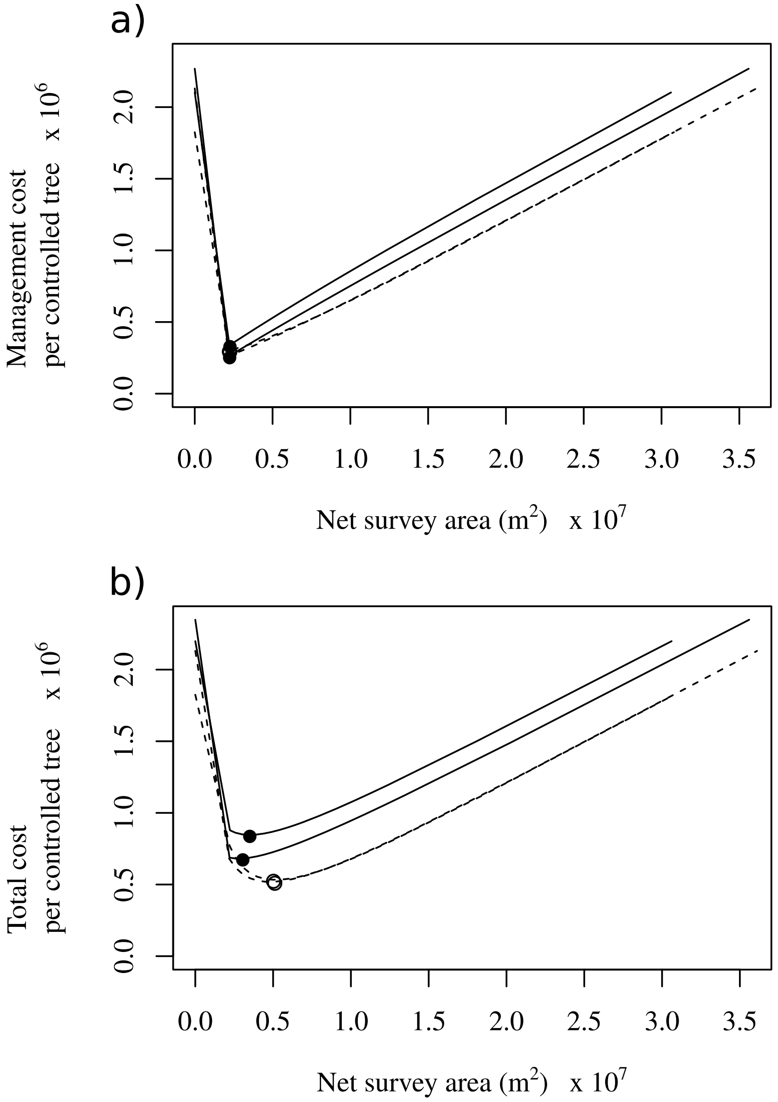

Figure 2.6 – Management cost per controlled tree (a; Eq. 2.4) and total cost per controlled tree (b; from Eq. 2.5 using $\theta = 1000$) in relation to the net survey area. Solid lines represent the local search around red-top trees for each 2011 and 2012. Dashed lines represent the search at locations chosen from model predictions for each 2011 and 2012. Black circles correspond to the minimum cost for the local search whereas white circles correspond to the minimum cost for the model predictions strategy.

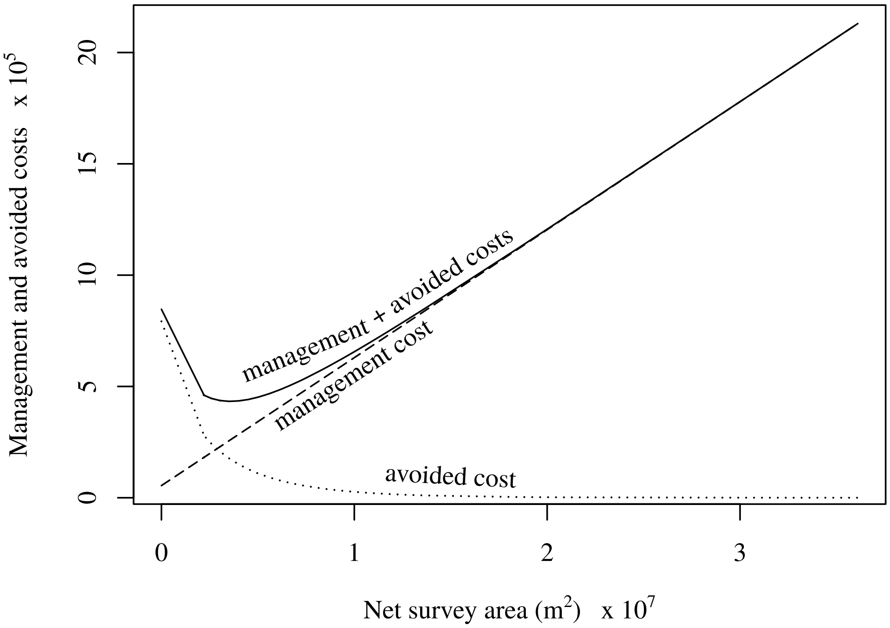

Figure 2.7 – Management cost (dashed line), avoided cost with $\theta = 1000$ (dotted line) and management plus avoided costs (= total cost; solid line) in relation to the net survey area for the model predictions strategy. The local search values, not presented here, display similar patterns.

2.4 Discussion

MPB infestations can be well predicted in space using a generalized boosted classification tree and variables related to the location of previous year infestations. A detailed analysis of the impact of survey areas on the control efficiency shows that combining an increase in survey area with a change in detection strategy leads to more cost-effective control.

2.4.1 MPB predictions

Generally, generalized boosted classification approaches often give better predictive accuracies than generalized linear approaches (Marmion *et al.*, 2009; Youssef *et al.*, 2016). Here, the percentage of correctly classified cells, $1-$ misclassification rate, is 84.9%. In comparison, Aukema *et al.* (2008) reported a predictive accuracy of 78% for a one-year ahead forecast using a spatial-temporal autologistic regression model on similar variables. At large scales (respectively 12x12 km and 1x1 km grid cell size in Aukema *et al.*, 2008; Preisler *et al.*, 2012), beetle pressure has a great impact on new infestations so it is not surprising to find indications that this is also the case in our results at a smaller scale.

While classification tree approaches can be used for prediction, they cannot be used to determine the actual impact of covariates on the response. Indeed, a classification approach, such as decision trees or boosted classification trees, often provide a relative importance index for each covariate, but this relative importance is an index of performance that depends highly on tree structures. A classification method does not test the impact of a covariate on the response like a traditional statistical method would, but rather attempts to explain the response by a sequence of binary choices among covariate values. However, it makes sense that environmental variables have less impact on the MPB presence than beetle pressure given that a small-size study area is usually relatively homogeneous.

Machine learning algorithms are widely used to detect/predict species locations (Marmion *et al.*, 2009) but few quantitatively compare the result to non-modelling/expert-knowledge methods like we did in this study (*e.g.* Boissard *et al.*, 2008).

2.4.2 Assessing management

The management assessment results show that the current detection strategy (searching in a 50 meter-radius plot around previous infestations) is efficient, but that using a larger survey area and a different strategy would improve efficiency. Robertson *et al.* (2007) found that 20 to 50 meters is the most common dispersal range but that MPB can go farther. These few individuals that go farther, and therefore are not removed during control, might be sufficient to sustain the population in the stand. MPB is subject to a strong Allee effect (Logan *et al.*, 1998; Goodsman *et al.*, 2016): at low beetle densities, a certain number of individuals is needed for a successful mass attack. Below this threshold, the attack is unsuccessful and the beetles either do not survive or fall back into the endemic population phase. The transition between endemic and epidemic population phases highly depends on both intrinsic and extrinsic factors which are subjected to a lot of uncertainty, making the transition forecast problematic (Cooke & Carroll, 2017).

Because of the existence of this threshold, local densities of beetles are important to infestation success. For that reason, Strohm *et al.* (2016) found that increasing search radius is more important than increasing search effectiveness, which is the percentage of infestations found within a survey area. Indeed, search effectiveness does not need to be flawless to decrease the beetle number below the Allee threshold. However, if the search radius is too small, enough beetles can disperse from neighbouring locations and successfully infest trees. For a search effectiveness of approximatively 80%, Strohm *et al.* (2016) show that MPB population size would decrease only if the search radius increases despite increases in search effectiveness. In Cypress Hills, for 2011 and 2012, we estimated the search effectiveness at 89%. This supports our recommendation to increase the survey area. Overall, Strohm *et al.* (2016) show that the search plot size of the Alberta management strategy (similar to Saskatchewan's strategy) was not large enough to reach the desired goal of reducing MPB population by 80% (Alberta Sustainable Resource Development, 2007) and the present study shows results consistent with this conclusion.

Local search around red-top trees, associated with short-distance dispersal, is a more efficient method than the random search, associated with random events from long-distance dispersal. This suggests that, despite intensive management, short-distance dispersal is still the main MPB dispersal strategy in Cypress Hills. However, a mechanistic model, such as the ones developed in Heavilin & Powell (2008), Rodrigues *et al.* (2015) and Goodsman *et al.* (2016), or the method described in Chen & Walton (2011), adapted for this area could likely give more insights on the subject by, in particular, quantifying the importance of both dispersal strategies.

An alternative to the local search around red-top trees is to survey locations with high probabilities of infestation. For a survey area larger than the one corresponding to the current strategy, it becomes more efficient to use the predictions strategy rather than the local strategy. This could be explained by the spatial scale of our model predictions. One 100×100m grid cell area and one 50 meter-radius circular plot area have the same order of magnitude. For a similar number of plots, the previous infestation at the same location decides for half of the model predictions results according to the relative importance whereas a red-top tree is always at the center of a circular plot. As the survey area increases, more of the red-top trees are included in the predictions survey in addition to other susceptible locations whereas the number of red-top trees included in the local survey does not change. Therefore, while in the local survey fewer and fewer green infestations are present the further away from the red-top tree, the predictions survey focuses on additional high risk locations chosen according to other variables, mainly the distance to the southern infested border, increasing the chance of finding more green infestations. One could combine both strategies: surveying first around red-top trees than adding extra survey plots in predicted areas that were not already surveyed until the alloted budget is reached.

Introducing a management cost allows for more informed decisions upon which to choose survey area size and detection strategy. Indeed, there is a minimum cost per controlled tree that corresponds to an optimal survey area larger than zero. This optimal survey area varies with the cost of missing a green infestation which can be calculated, for example, by the cost of a circular survey plot plus the cost of removing a certain number of new green infestations due to this red-top tree.

2.4.3 Limitations

A potential limitation of this work is the assumption that the cost associated with several 50 meter-radius plots is equivalent to the cost of one much larger plot of the same total area, and that this relationship is linear, even for areas as large as 20% of the park surface. One could also assume that the relationship's slope would decrease as survey locations are closer in space and managers spend less money and time travelling between locations. These savings seem negligible, nonetheless, it would decrease the slope of the relationship between cost per controlled tree and survey area at larger survey areas. However, it would probably have little impact on the location of the minimum cost and thus the optimal survey area size.

Another limitation is that we only undertook the analysis for years with a number of red-top trees approximately equal to 300 as only data for these years were available. The survey area values are directly linked to the number of survey plots and, thus, the number of red-top trees for each year. Therefore, the survey area values are not directly applicable to years with a different number of red-top trees, although the curve patterns would be similar. The results also vary with the ratio total number of green infestations to number of red-top trees. This ratio was larger in 2012 than 2011. However, we scaled most of the results by the total number of green infestation to allow a fair comparison of both years.

Furthermore, the selection of only two consecutive years of data makes the analysis potentially susceptible to bias due, for example, to particular weather conditions or to the specific details of implementation of management work for these two years. To minimize the latter, however, a detailed survey protocol is implemented.

2.4.4 Conclusion

The control efficiency in Cypress Hills could be slightly increased for a smaller cost, which includes the future savings made by controlling an infested tree now rather than several ones the following year. This would be done by engaging more management resources, such as survey plot radius larger than 50 meters, in combination with using a search strategy that exploits criteria other than the location of red-top trees, such as weather and stand characteristics.

Chapter 3

When managers forage for pests: Implementing the functional response in pest management

3.1 Introduction

The functional response framework characterizes predator-prey interactions by the relationship between prey density and the number of prey captured by a predator (Holling, 1959a). In this well-defined framework, time acts as a limitation: authors typically consider temporal costs, such as handling time, which, when large, decreases the number of occasions for prey capture. In the functional response literature, researchers usually study pest management in two ways. First, in the context of biological control, a predator is a means of management (*e.g.* Mills & Getz, 1996; Van Den Berg *et al.*, 1997). Functional responses are then used in the usual way to describe a biological control response to the pest level. Second, management is added to a predator-prey system by removing either predator or prey at a given rate. When management is added to models in this way, the functional response usually does not vary except if the control method affects predators or prey behaviour or if the ratio between prey and predator abundance affects the number of prey captured (*e.g.* Liu *et al.*, 2006; He *et al.*, 2012).

An alternate way to model the direct influence of management on pests using functional responses would be to consider managers physically removing or controlling a pest as predators. Although they deal with a resource and not a pest, fisheries models, using harvesting effort or harvest rate, are a first step into describing the effect of management on a resource using a linear relationship (Sutherland, 2001). However, to our knowledge, the theory of manager functional responses displaying various shapes has not been previously proposed or tested. Applying functional response literature to human-pest systems would be powerful as there is an extensive functional response theoretical background.

Could human-pest interactions be treated as predator-prey interactions in a functional response framework? Managers and predators tend to have similar behaviours in this context: both wish to remove the maximum possible number of prey/pests; both have limitations, whether temporal or monetary, preventing them from maximizing the number of prey/pests they capture (Hassell, 1978); both face complex spatial distribution of prey/pest population; and both may employ a variety of search strategies. In this study, we will explore the functional response types obtained under the assumption that managers take the role of predators, pests take the role of prey, and monetary costs constrain managers in a similar way that temporal costs constrain predators. For example, the pest handling cost for a manager can be shown to be equivalent to the prey handling time for a predator in the functional response formulation. Although time plays a major role in management, we argue that, at the scale of a management season, monetary costs are the most important limiters of management success; mostly because the time alloted to management directly depends on budget.

Although functional response theory usually concerns a random search in an homogeneous domain, spatial heterogeneity is usually considered as a norm in the environment (Levin, 1992; Gustafson, 1998) and should be studied for realism. Spatial patterns are usually classified as regular, random or clustered (Hopkins & Skellam, 1954). A regular spatial pattern would look like points on a grid whereas a clustered, or aggregated, spatial pattern would feature isolated groups of points in space. For example, black flies (Diptera: Simuliidae) display a clustered, random, or regular spatial distribution depending on their species (McCreadie & Adler, 2019). In each case, the observed pattern depends on the scale of the study. Indeed, a population could appear randomly distributed when viewed on a small scale, but clustered on a large scale. Depending on the spatial distribution of resources the species depends on, observed patterns could also change from one area to another. Not taking into account realistic pest spatial distribution in functional response studies may include bias and lead to inaccurate determination of the functional response shape (Ives *et al.*, 1999; Hochberg & Holt, 1999). In this study, we examine the impact of random, clustered and regular pest spatial patterns.

The predator/manager searching strategy could also have an impact on the type of functional response observed. Functional response studies usually assume a random search but other strategies are possible. The adaptive cluster sampling strategy is an established alternative to random searching (Thompson, 1990). In this case, after the random sample of a first set of locations, if the variable of interest in any location is bigger than zero, additional nearby locations are added to the survey. When the pest population is clustered, the adaptive cluster sampling will efficiently find most of the individuals in a given cluster. Maxwell *et al.* (2012) compared adaptive cluster sampling to traditional transect designs and found out that the former was more efficient than the latter when the purpose is to survey as many individuals as possible in well-established populations away from roads in the shortest amount of time. This was found to be especially true for clustered populations. The adaptive cluster sampling strategy is thus an efficient survey strategy that can be easily simulated in our theoretical study. Therefore, we employed two simple contrasting search strategies: random sampling and adaptive cluster sampling.

The objectives of this theoretical work are: 1) to create simulation models of pest management using a spatially-explicit individual-based model and spatially-implicit

Monte-Carlo simulations, 2) to derive simple functions describing the number of pests that managers control as a function of pest density and corroborate the simulation models, 3) to characterize the impact of the different components on the functional response shape and compare with predator functional response components and shapes, and 4) to validate this new framework by applying it to data using the example of mountain pine beetle in Cypress Hills, Saskatchewan. Using monetary cost for the manager as an equivalent of temporal cost for a predator, we are able to make the analogy between functional responses in predator-prey interactions and human-pest management systems. This opens the door to applying functional response to manager-pest relationships.

3.2 Methods

We simulated management strategies of pest removal using two approaches: a spatially-explicit individual-based model and spatially-implicit Monte-Carlo simulations, and confirmed the results in the simplest cases by deriving analytical solutions to mathematical models representing the simulation scenarios. First, we presented the computational (§3.2.1 and §3.2.2) and mathematical (§3.2.3) models corresponding to the first two objectives. Second, we explained how the different parameters defined in the models help in connecting our ideas to the functional response framework (§3.2.4). Finally, we applied our models to the mountain pine beetle example (§3.2.5).

In this study, we explored various scenarios so as to capture different functional response behaviours. Each scenario has two elements: 1) a pest spatial pattern in a 2-dimensional domain, and 2) a management strategy. We divided the domain into cells of varying size on a 16×16 grid. Depending on the scenario, the cell size was 1×1, 2×2 or 8×8 distance units. Each cell was characterized by its spatial position in the domain and by the number of pest items it contained. Fig. 3.1 represents the process for each simulation run and sections 3.2.1 and 3.2.2 provide details on each component. The symbols used in this study are described in Table 3.1.

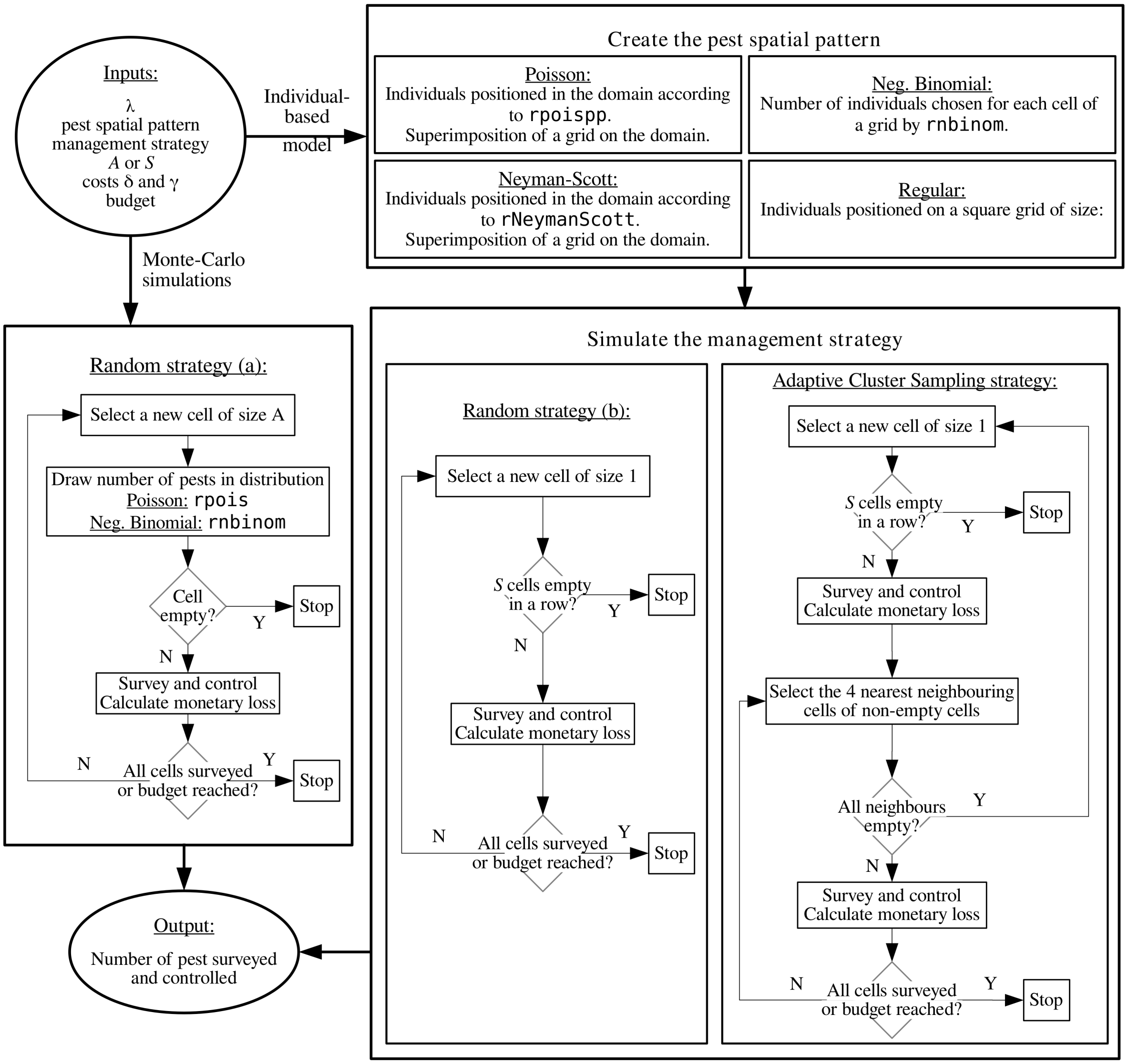

Figure 3.1 – Conceptual diagram representing the process determining each simulation run. Start at the top left of the diagram.

Table 3.1 – Description of the symbols used in the text.

Symbol	Description
λ	Pest intensity or Average number of individuals in a size-1 grid cell
δ	Cost of surveying a size-1 cell
γ	Cost of removing a pest individual
A	Cell size/area or Number of contiguous size-1 cells
S	Nb. of empty cells to survey in a row before stopping the search process
search area	Search area refers indistinctively to A or S
M	Domain area where management takes place
ϕ	Probability of finding at least one individual in a grid cell of area A
D	Expected area explored by managers
E	Expected number of individuals that are surveyed and controlled
B	Management budget
k	Amount of clustering in the negative binomial distribution
ψ_A	Probability of finding zero individuals in a cell of area A

3.2.1 Simulating Pest Distributions

We examined four types of spatial patterns: random, clustered within-cells, clustered between-cells and regular point patterns.

Random point process

The random point process employed a homogeneous Poisson process using the `rpoispp` function of the `R` package `spatstat` (Baddeley *et al.*, 2015; R Core Team, 2016). A grid with rectangular cells of constant area (A) was then superimposed over the whole spatial domain. Each cell had the same probability of having a pest, thus, they were independent of each other.

Within-cell cluster point process

The within-cell cluster point process sampled a Negative binomial distribution for the number of pests in each grid cell using the `rnbinom` function of the `R` package `stats` (R Core Team, 2016). We used the overdispersion parameter k fixed to 1 to describe the amount of clustering in the distribution. With this method, clusters happened within a cell and thus, grid cells were independent of each other. This spatial distribution led to no spatial correlation in the density of pests between cells but greater variability in the density of pests from one cell to another than for between-cell clustering. The spatial resolution of the simulated pest data depended on the size of the square cells in the grid that we overlaid on top of the spatial domain.

Between-cell cluster point process

The between-cell cluster point process employed a Neyman-Scott process using the `rNeymanScott` function of the `R` package `spatstat`. This method distributed "parent" points in the domain according to a Poisson point process. Then, it drew the actual points from a Gaussian distribution around each parent. The final step consisted of removing the parents. A grid with rectangular cells of constant area (A) was then superimposed over the whole spatial domain. Thus, there was a spatial correlation of the density of pests between cells.

Regular point process

The regular point process distributed individuals regularly on a square grid of size equal to the square root of the total number of individuals in the domain rounded down to the nearest integer. We positioned remaining individuals using the same process. We added some noise proportional to the distance between individuals so as to reproduce a more biologically relevant spatial pattern. In order to be distributed uniformly in the domain, each point depended on the location of its neighbours. Thus, the number of individuals in each cell was dependent on each other.

Fig. 3.2 shows an example of each spatial distribution of points for an average intensity of 0.2 individuals per grid cell.

3.2.2 Simulating Manager Search Strategies

We simulated three different management strategies. The first two were both random searches that differed in the implementation of the search area parameter: a random search strategy (a) in which management stops after the manager discovers one empty cell of variable size and a random search strategy (b) in which the manager stops searching after a specified number of consecutive empty cells of size 1. The third strategy was the adaptive cluster sampling strategy.

Random Strategy (a): stop after single empty cell of area A

Using a spatially-implicit Monte-Carlo approach, we simulated the random strategy (a) (stop after single empty cell of area A) for the spatially uncorrelated pest spatial distribution (Poisson and negative binomial) with intensity λ on a bounded rectangular spatial domain of area M. This approach could only work with spatially uncorrelated distribution given the spatially-implicit nature of the process. The management then proceeded as follows. 1) One cell in the grid was randomly selected. The number of pest in the cell was drawn from a Poisson or a negative binomial distribution using,

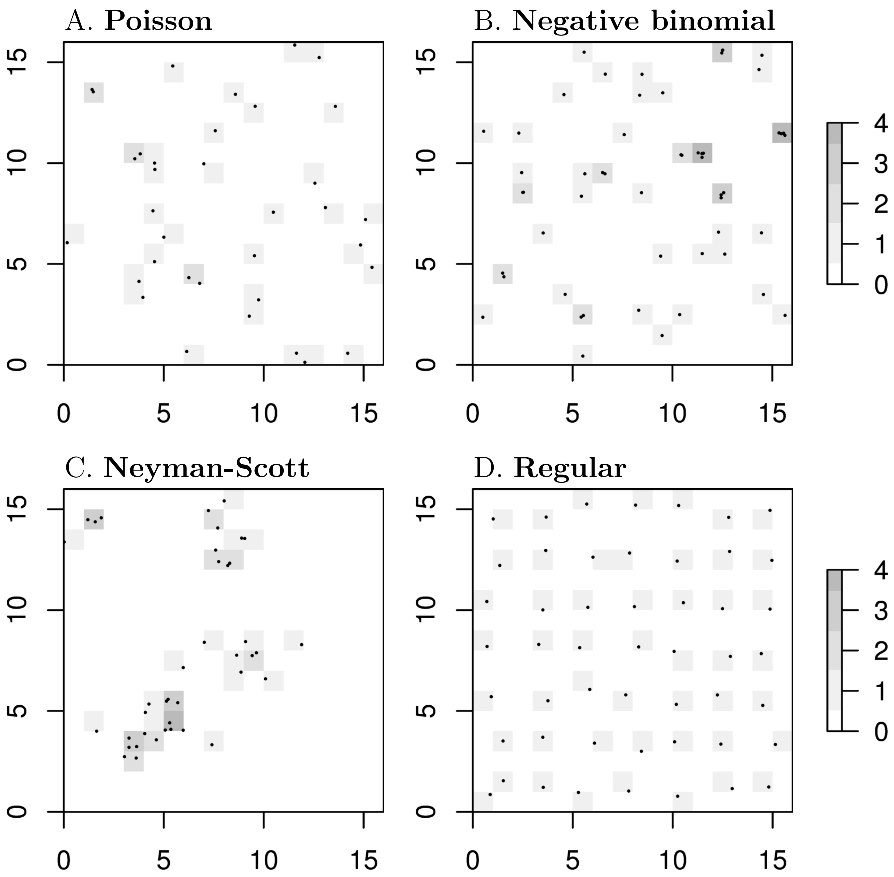

Figure 3.2 – Spatial distributions of pests on a 16×16 lattice with an intensity of 0.2 individual per grid cell: A. Poisson spatial distribution, B. Negative binomial (within-cell cluster) spatial distribution with an overdispersion parameter $k = 1$, C. Neyman-Scott (between-cell cluster) spatial distribution and D. Regular spatial distribution. The black dots represent the actual point patterns from which the lattice spatial distributions are derived. The shades of grey represent the number of pests per grid cell: white (zero individuals) to the darkest grey (four individuals).

respectively, the R functions `rpois` and `rnbinom`. If the initial cell contained no individuals, the survey stopped. If it contained at least one individual, all individuals inside it were surveyed/controlled/eradicated and another cell was chosen at random. 2) This same process was repeated for the next cell and the procedure was repeated until either no individuals were found in a quadrat, all the cells in the management area were surveyed, or the budget was reached. Note that the random selection of a new cell was from the cells that had not already been surveyed. Since the survey area A could be considerably large, when a step would put managers over the budget, only the fraction of the area allowed by the budget was managed.

Random Strategy (b): stop after S empty cells

For the random search strategy (b), we used a spatially-explicit individual-based model that we simulated on all four pest spatial distributions. This model was computationally intensive compared to the Monte-Carlo approach so we did not use it for strategy (a). The random search strategy (b) consisted of randomly choosing cells without replacement and removing any pest contained in it. The process stopped when it reached the budget or the maximum number of empty cells surveyed in a row (S), whose value depended on the scenario. We depleted the budget in two ways depending on the scenario: by deducting the cost of surveying a cell (δ) $*$ the number of cells surveyed at this step, and by deducting the cost of removing a pest item (γ) $*$ the number of pest items removed at this step. We did not take any action/step that would put managers over the budget.

Adaptive cluster sampling strategy

In the adaptive cluster sampling strategy, we chose an initial cell to survey. If at least one pest was present in the cell, we added the 4 adjacent neighbouring cells to the survey (Von Neumann neighbourhood). We repeated this around the new cells that contained pests. At each step, we removed pests found in the surveyed cells. If we found no pest in the initial cell or in all of the neighbouring cell at some step, we chose a new initial cell randomly and the process resumed. The process stopped when it

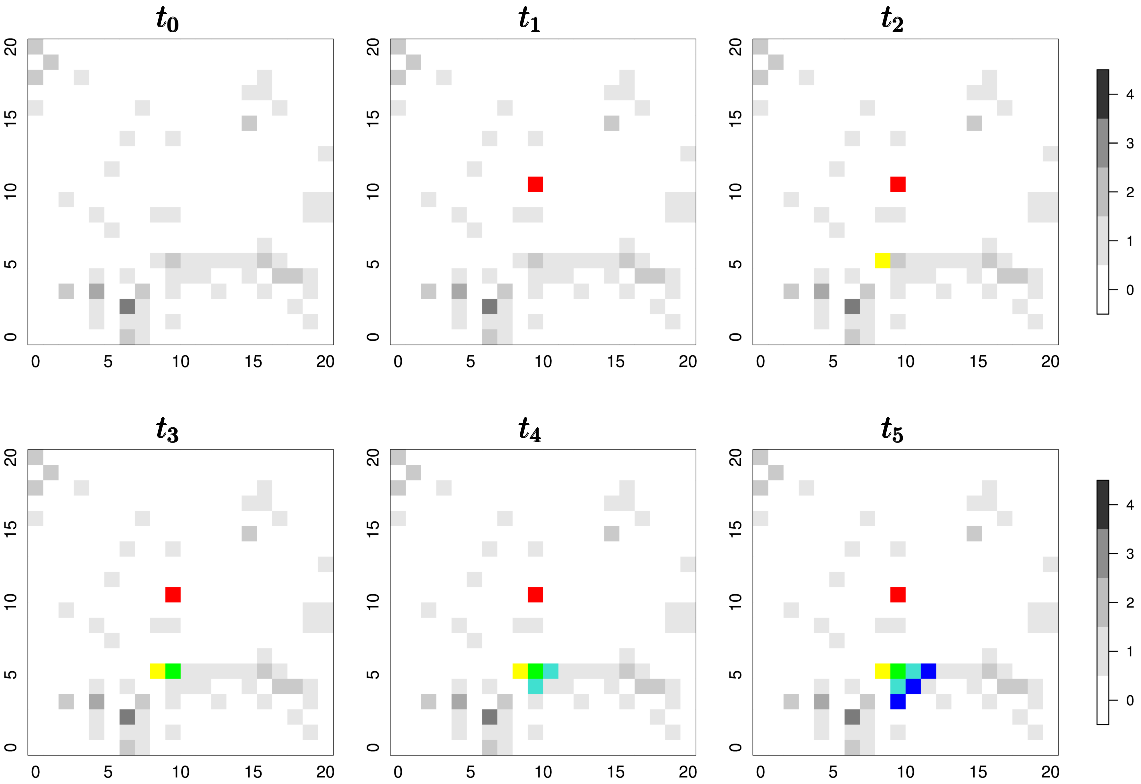

Figure 3.3 – Step by step images of the adaptive cluster sampling strategy inspecting a clustered (Neyman-Scott process) spatial distributions of pests (in shades of grey: from white, zero individual, to the darkest grey, four individuals) on a 20×20 lattice with an intensity of 0.2 individual per grid cell. The colors represent the steps taken by the algorithm. The first step (in red) hit an empty cell so another cell is chosen at random (in yellow). The yellow cell contains one individual so it is removed and the 4 cells of the Von Neumann neighbourhood are surveyed. One of them contains individuals (in green). Individuals are removed and the neighbourhood of the green cell is surveyed. In this neighbourhood, two cells contain individuals (in turquoise) so they are removed and the Von Neumann neighbourhood of both cells is surveyed. In this new neighbourhood, three cells contains individuals (in blue). The process goes on.

reached the budget or the maximum number of empty cells surveyed in a row (S). See Fig. 3.3 for an illustration of this process. Again, we depleted the budget allowance by an amount defined by the cost of surveying a cell (δ) $*$ the number of cells surveyed at this step, and the cost of removing a pest item (γ) $*$ the number of pest items removed at this step. Since the added neighbourhood could be of considerable size, when a step would put managers over the budget, only the fraction of the area allowed by the budget was managed.

The random strategy (b) (stop after S empty cells) and the adaptive cluster sampling strategy were deployed on all four pest distributions and the grid applied on the domain had cells of constant size 1×1.

For the strategies simulated with a spatially-explicit model (random strategy (b) and adaptive cluster sampling), we defined, for simplicity, the domain as a torus, which means that when a manager traversed the border of the domain, the manager reappeared on the other side of the domain.

3.2.3 Mathematical Models

We derived mathematical models of the random sampling in which management stops after the manager discovers one empty cell of area A (random strategy (a) described in §3.2.2) and the manager random sampling in which the manager stops searching after S consecutive empty size-1 cells (random strategy (b) described in §3.2.2) for pest distributions that are spatially uncorrelated from one cell to another (Poisson and Negative binomial spatial distributions).

Random Strategy (a): stop after single empty cell of area A

The probability of finding at least one individual in a grid cell of area A ($A < M$) with a Poisson spatial distribution of individuals is

$$Pr(X > 0; \lambda, A) = 1 - \exp(-\lambda A) = \phi. \tag{3.1}$$

Note that $\phi < 1$. We can then write an expression for the expected area explored by managers (D) using a strategy like the one described in the previous section. We can call this the *'area of discovery'* to connect it to the functional response ideas.

$$\begin{aligned} D(\lambda, A, M) &= A(1 + \phi + \phi^2 + ... + \phi^{M/A-1}), \\ &= A\left(\sum_{i=0}^{M/A-1} \phi^i \right), \\ &= A\left(\frac{1 - \phi^{M/A}}{1 - \phi} \right). \end{aligned} \tag{3.2}$$

Note that in Eq. 3.2 managers are unable to explore more area than there is in the management area M and so the maximum number of quadrats is M/A. The number of pests that are eradicated/controlled/surveyed (E) is proportional to the area explored

$$E(\lambda, A, M) = D(\lambda, A, M)\lambda, \tag{3.3}$$

Therefore, the number of pests that are eradicated/controlled/surveyed is

$$E(\lambda, A, M) = A\lambda\left(\frac{1 - \phi^{M/A}}{1 - \phi} \right). \tag{3.4}$$

If there is a limited management budget B, each quadrat costs δ, and the cost γ is associated with eradicating each pest in addition to treating an infested cell, then the maximum number of quadrats of area A that can be explored is $\frac{B}{A(\lambda\gamma+\delta)} \leq \frac{M}{A}$. We assume that the budget restricts exploration such that an area less than the full area of the management zone can be explored. The idea of adding a budget, a cost per quadrat, and a cost per individual is similar to dividing the total foraging time into searching time and handling time in the Holling's disc equation. The number of pests that are eradicated/controlled/surveyed becomes

$$E(\lambda, A, B, \delta, \gamma) = A\lambda\left(\frac{1-\phi^{\frac{B}{A(\lambda\gamma+\delta)}}}{1-\phi}\right). \tag{3.5}$$

If instead we assume that individuals are distributed according to the Negative binomial distribution and that the number of pests in any cell is independent of the number in any other cell, we write for a cell of area 1

$$Pr(X > 0; \lambda, k) = 1 - \left(1 + \frac{\lambda}{k}\right)^{-k} = \phi. \tag{3.6}$$

In Eq. 3.6, k represents the amount of clustering from one cell to another with small k corresponding a high degree of clustering.

For cells of area $A > 1$, the Negative binomial distributed pest data would have a different spatial resolution as we sample on cells of area A. Eq. 3.6 becomes then:

$$Pr(X > 0; \lambda, A, k) = 1 - \left(1 + \frac{\lambda A}{k}\right)^{-k} = \phi. \tag{3.7}$$

As $k \to \infty$, Eq. 3.7 converges to the Poisson case given by equation (3.1). The derivations for the numbers of individuals eradicated (Eqs. 3.4–3.5) are still valid for this case.

Table 3.2 summarizes the equations used for the expected number of pests controlled depending on the pest spatial distribution for random strategy (a).

Random Strategy (b): stop after S empty cells

In strategy (a), A represents the cell area or the survey of A contiguous size-1 cells and the process stops after one empty step. In strategy (b), however, the manager surveys an integer number S of empty cells before stopping. The main difference between random strategy (a) and random strategy (b) is that the expected number of individuals removed by exploring A contiguous size-1 cells would, in theory, be affected by the

Table 3.2 – Expected number of pests controlled (E) for random strategy (a).

Spatial distribution	E	ϕ
Poisson	Eq.(3.5): $A\lambda\left(\frac{1-\phi^{\frac{B}{A(\lambda\gamma+\delta)}}}{1-\phi}\right)$	Eq.(3.1): $1-\exp(-\lambda A)$
Negative binomial	Eq.(3.5): $A\lambda\left(\frac{1-\phi^{\frac{B}{A(\lambda\gamma+\delta)}}}{1-\phi}\right)$	Eq.(3.7): $1-\left(1+\frac{\lambda A}{k}\right)^{-k}$

spatial distribution of individuals while exploring S randomly picked size-1 cells would not. Below we compare mathematical models of the management functional response for random manager searches when pests are distributed according to a Poisson point process and when pests are distributed according to a negative binomial point process. Our derivations illustrate that the expected number of individuals removed can depend on the pest distribution when multiple empty cells are required before the manager stops looking for additional pests.

In the case of a Poisson distributed resource, from Eq. 3.1, we get the probability that a cell i of area A has zero individuals:

$$Pr(X_i = 0; \lambda, A) = \exp(-\lambda A) = \psi_A. \tag{3.8}$$

The probability that S different cells of area A have zero individuals in each cell is:

$$\begin{aligned} Pr(X_{1,\ldots,S} = 0; \lambda, A, S) &= (\psi_A)^S \\ &= (\exp(-\lambda A))^S \\ &= \exp(-\lambda A S). \end{aligned} \tag{3.9}$$

A and S being multiplied, the probability that a cell ($S = 1$) of area $A = x$ has zero individuals (ψ_A^1) is the same as the probability that x different size-1 cells ($S = x$, $A = 1$) have zero individuals in each cell (ψ_1^S). So in this context, $\psi_A^1 = \psi_1^S$. For this reason, the mean number of individuals managed obtained from the simulation process described above should match the analytical solution provided in Eq. 3.5 using Eq. 3.1.

In the case of a resource distributed with a negative binomial distribution, from Eq. 3.7, we get the probability that a cell i of area A has zero individuals:

$$Pr(X_i = 0; \lambda, A, k) = \left(1 + \frac{\lambda A}{k}\right)^{-k} = \psi_A. \tag{3.10}$$

The probability that S different cells of area A have zero individuals in each cell is:

$$\begin{aligned} Pr(X_{1,\ldots,S} = 0; \lambda, A, S, k) &= (\psi_A)^S \\ &= \left(\left(1 + \frac{\lambda A}{k}\right)^{-k}\right)^S \\ &= \left(1 + \frac{\lambda A}{k}\right)^{-kS}. \end{aligned} \tag{3.11}$$

In this case, A and S are not multiplied, so the probability that a cell ($S = 1$) of area $A = x$ has zero individuals (ψ_A^1) is generally not equal to the probability that x different size-1 cells ($S = x$, $A = 1$) have zero individuals in each cell (ψ_1^S). So in this context, $\psi_A^1 \neq \psi_1^S$. The mean number of individuals managed obtained from the simulation process described above should generally not match the analytical solution provided in Eq. 3.5 using Eq. 3.7. In this case, instead of Eq. 3.7, we can use

$$Pr(X_S > 0; \lambda, A, S, k) = 1 - \left(1 + \frac{\lambda A}{k}\right)^{-kS} = \phi. \tag{3.12}$$

Table 3.3 summarizes the equations used for the expected number of pests controlled using the random strategy (b) depending on the pest spatial distribution.

3.2.4 Correspondence with functional responses types

To compare human-pest management to functional response in predator-prey interactions, we considered that pests to managers were like prey to predators in the functional response framework. We then assumed that the budget allocated to a manager corresponded to the foraging time for a predator. Keeping the comparison between monetary and temporal costs in mind and with the assumption that the behaviour "processing food/pest" is independent from the behaviour "searching for food/pest", we could make the correspondence between the time needed to search for a prey and the monetary cost requested to search for a pest (survey), and between the time needed to consume a prey and the monetary cost requested to remove a pest.

Table 3.3 – Expected number of pests controlled (E) for random strategy (b).

Spatial distribution	E	ϕ
Poisson	Eq.(3.5): $A\lambda\left(\frac{1-\phi^{\frac{B}{A(\lambda\gamma+\delta)}}}{1-\phi}\right)$	$1-$ Eq.(3.9): $1-\exp(-\lambda AS)$
Negative binomial	Eq.(3.5): $A\lambda\left(\frac{1-\phi^{\frac{B}{A(\lambda\gamma+\delta)}}}{1-\phi}\right)$	Eq.(3.12): $1-\left(1+\frac{\lambda A}{k}\right)^{-kS}$

In a Type I functional response scenario, the predator is not affected by its capacity to consume a prey so the number of prey consumed increases linearly with the intensity of prey in the domain. Thus, by setting the cost of removing a pest item (γ) to 0, we expected to obtain a linear relationship (Holling Type I) between the number of pest removed and the intensity of pests in the domain.

In a Type II functional response scenario, the predator is limited by its capacity to consume a prey so as the intensity of prey in the domain increases, the predator cannot consume more than a certain number of prey and the curve saturates. In a pest management context, satiation is represented by time or monetary limitations. At high pest density, a manager cannot remove more pests from the environment than allowed by their budget. Thus, by setting the cost of removing a pest item (γ) to be larger than 0 while maintaining a large search area (A and S both impact the search area), we expected to obtain a hyperbolic relationship (Holling Type II) between the number of pest removed for a fixed cost and the intensity of pests in the domain.

In a Type III functional response scenario, at low intensities of prey, some mechanisms make it more difficult for the predator to find and consume a prey than at higher intensities. Some researchers explain this concept using predators' behaviour (Turchin, 2013). A generalist predator might switch to another prey when the focus prey density is too low leading to a Type III whereas a specialist would have no choice but to continue searching for the focus prey leading to a Type II. In a pest management context, a manager might act like a specialist by wanting to remove all pest or like a generalist by being satisfied with a low pest number and switching to another management activity. Thus, by setting the maximum number of empty cells surveyed in a row before stopping (S) (or its equivalent A: number of contiguous size-1 cells, see previous section) to a low number and by setting $\gamma > 0$, we would expect that Eq. 3.5 takes the sigmoid shape of a Type III functional response. We thereby simulate a manager estimating that the probability of encountering a pest is too low to be worth the search effort at a low pest density.

3.2.5 Application

We applied our modelling framework to the mountain pine beetle (*Dendroctonus ponderosae*, Hopkins 1902) management in Cypress Hills, Saskatchewan, Canada. The mountain pine beetle is a bark beetle that infests and kills pine trees in western North-America. Mountain pine beetles have two main population stages: an endemic stage in which there are not enough individuals to overcome healthy trees, so populations persist by attacking stressed trees in association with other bark beetles, and an epidemic stage in which mountain pine beetles are a threat to vast stands of healthy pines. There are currently several epidemic populations, including one in the Cypress Hills park. Mountain pine beetle populations exhibit a type of Allee effect (Allee, 1931; Stephens *et al.*, 1999): below a certain number of individuals, a local epidemic population cannot persist and will either go extinct or turn to an endemic population stage.

Pest spatial distribution

During an epidemic, the attacked trees are usually spatially aggregated due to beetle offspring emerging from one previously attacked tree and attacking susceptible trees within several hundred meters (Safranyik & Carroll, 2006; Robertson *et al.*, 2007). This pattern resembles the Neyman-Scott process described in section 3.2.1. Therefore, we fitted a Neyman-Scott process to attacked trees locations in 2011 and 2012 from several portions of the park with different pest densities using the `vargamma.estpcf` function of the `R` package `spatstat`. This function uses the pair correlation function to fit the point process to a point pattern by the method of minimum contrast.

In addition to the parameter values, we modified the Neyman-Scott process in our simulations to draw points around parents from a fat-tail distribution (here an exponential distribution) which corresponds better to the mountain pine beetle behaviour than a Gaussian distribution due to the occurrence of long-distance dispersal events (Safranyik & Carroll, 2006; Robertson *et al.*, 2007; Goodsman *et al.*, 2016).

Management strategy

The management strategy implemented in Cypress Hills consists of locating previously attacked trees (red-top trees easily detectable during aerial surveys) and surveying the surrounding neighbourhood for new attacks. We simulated this process by adapting the adaptive cluster sampling strategy. The strategy is deployed at the location of previously attacked trees instead of random locations. The process continues until all cells containing previously attacked trees are surveyed. We set the managers ability to detect newly attacked trees within the survey area to 89% (value obtained for Cypress Hills from M. Kunegel-Lion, unpublished data).

Simulations

Using the fitted pest spatial distribution and the modified adaptive cluster sampling strategy, we simulated the management process 10 000 times for each pest density on a 6.5×9.9km rectangular domain with grid cells of size 100×100m. Thus, the grain and extent of the domain is the same order of magnitude as the grain and extent of the management area in Cypress Hills. The management costs (δ and γ), and the total budget (B) values were chosen as proportional to the actual costs and budget in the park for 2011 and 2012. We then compared the curve obtained to actual numbers of attacked trees controlled in several areas of the park presenting different densities of attacked trees during 2011 and 2012.

Management goal feasibility

Knowing about management functional responses shape can help us assess the feasibility of a management goal. We chose two ways of expressing a management goal: 1) managers would like to remove at least $x\%$ of all the pests, and 2) managers would like to leave no more than x individuals per unit of area. This second management goal makes sense for populations exhibiting an Allee effect, as the mountain pine beetle does. Indeed, it is not necessary to put more effort into control when the pest population will not persist below a certain number. We simulated the management process described above 1000 times for increasing budget values and compared the management functional response curves obtained to the two management goals. To place the functional response curves and the management goals on the same scale, we divided the number of individuals removed by the domain area, thereby graphing the density of pest controlled by the total pest density. Note that this only changed the y-axis scale and therefore the interpretation of the curve did not change.

3.3 Results

For each combination of pest distribution, management strategy, and parameter values, we compared the means of 2000 simulations to the analytical solutions, when calculated. The budget value used in the simulations and analytical solutions was 300, limiting the number of cells managers can explore. The values of δ and γ were either 10 or 0 and the values of A and S either 64 (one 8×8-cell or 64 1×1-cells), 4 (one 2×2-cell or four 1×1-cells) or 1 (one 1×1-cell) depending on the scenario. We generated all negative binomial point processes using the parameter $k = 1$ in both the simulations and the analytical solution.

3.3.1 Varying the management costs

When the cost of surveying a cell δ increases, the functional response slope decreases and when the cost of removing a pest item γ increases, the functional response curve saturates as seen on Fig. 3.4 for the Poisson distribution. This is true for all the management strategies and their analytical solutions on all pest distributions. Note that the curve on Fig. 3.4a) continues to increase linearly above 50 individuals removed. See appendix C, Fig. C.1 for the non-truncated graph.

3.3.2 Varying the management strategy and the pest spatial distribution

For a Poisson pest spatial distribution, the functional response curves for the simulations of the random strategies (a) and (b) are the same as predicted by the derivations of

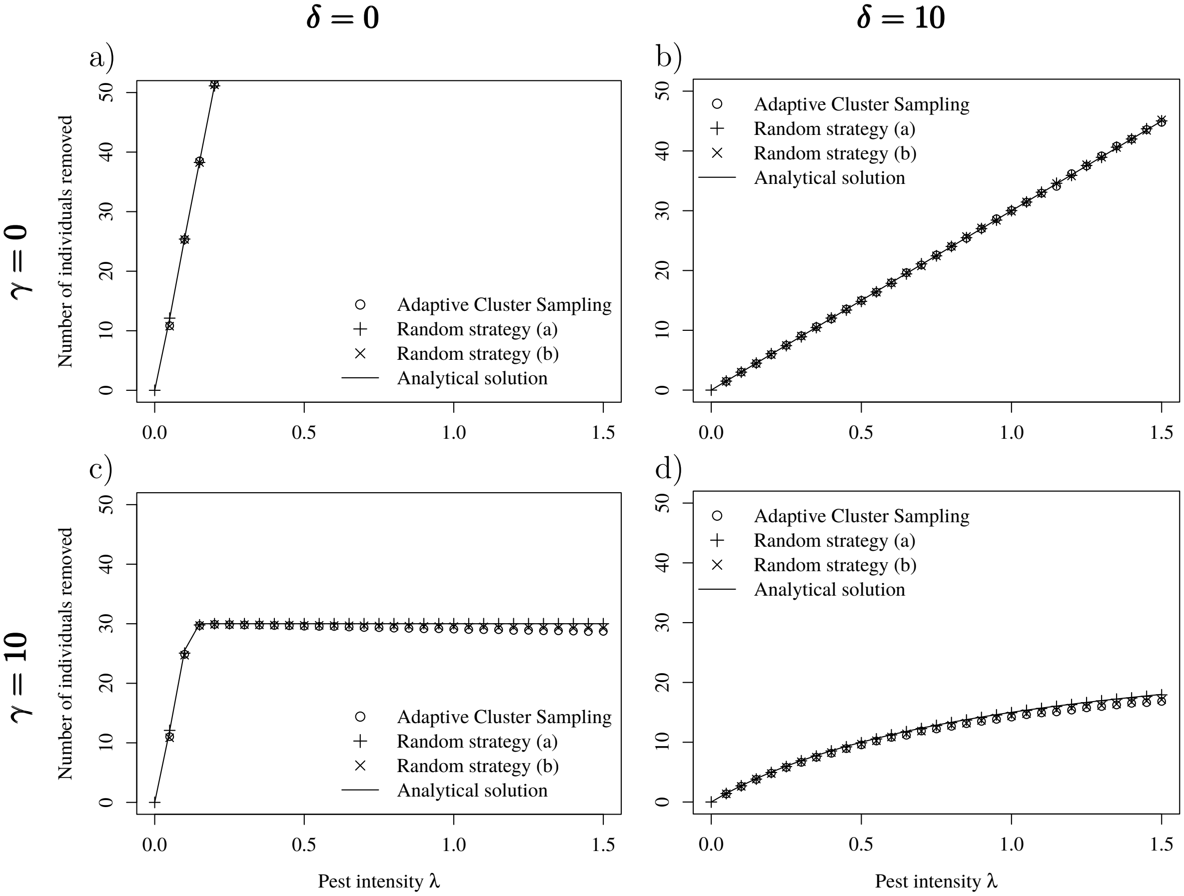

Figure 3.4 – Functional response curves when varying the management costs (δ the cost of surveying a cell and γ the cost of removing a pest) for fixed values of $A = 64$ and $S = 1$ for the random strategy (a) and $A = 1$ and $S = 64$ for the random strategy (b) and the adaptive cluster campling strategy, and for a Poisson pest spatial distribution. Circles represent the means of 2000 simulations of the adaptive cluster sampling strategy, pluses of the random strategy (a), and crosses of the random strategy (b). Solid lines represent the values for the analytical solution of random strategy (a) and (b).

analytical solutions. For a negative binomial pest spatial distribution, the functional response curves for the simulations of the random strategies (a) and (b) are extremely similar even though the expected means differ due to a different pest spatial resolution. This can be seen on Fig. 3.5. Unexpectedly, the pest spatial distribution does little to change the functional response curves. The adaptive cluster sampling and the random strategies present the same pattern with little variation from one strategy to another except for Neyman-Scott and regular pest spatial distribution. Indeed for a Neyman-Scott distribution, the adaptive cluster sampling strategy does often better than the random strategy (b) especially for the parameters values $\delta = 10$, $\gamma = 0$ (Fig. 3.5). For a regular distribution, the adaptive cluster sampling strategy results in oscillations around the random strategy mean especially for the parameters values $\delta = 10$, $\gamma = 0$ (Fig. 3.5).

3.3.3 Varying the search area

As expected, decreasing the search area, as represented by A or S (see Methods section), decreases the number of pests found and controlled at low pest intensity levels for both random strategies and leads to a curve resembling a Type III when $\gamma = 10$ (Fig. 3.6). In the Negative binomial case, there are differences between random strategies which might be due to the difference between A and S as shown in Eq. 3.12 or to the difference in the pest spatial resolution when $A = 1$ and when $A > 1$. Those two cases cannot be distinguished given our parameter values.

3.3.4 Theoretical results summary

Additional simulations (see Appendix D, Fig. D.1 and D.2) show that when A or S are large, the cost per cell to survey $\delta > 0$ and the cost per pest to remove $\gamma \to 0$, the functional response tended to a Type I. When A or S are large, $\gamma > 0$ and $\delta \to 0$, the functional response tends to a Type I/II which is linear like a Type I at first with a progressive saturation like a Type II (Jeschke *et al.*, 2004). When γ and $\delta > 0$, the functional response tends to a Type II. To summarize, to obtain a Type I, we set $\delta > 0$

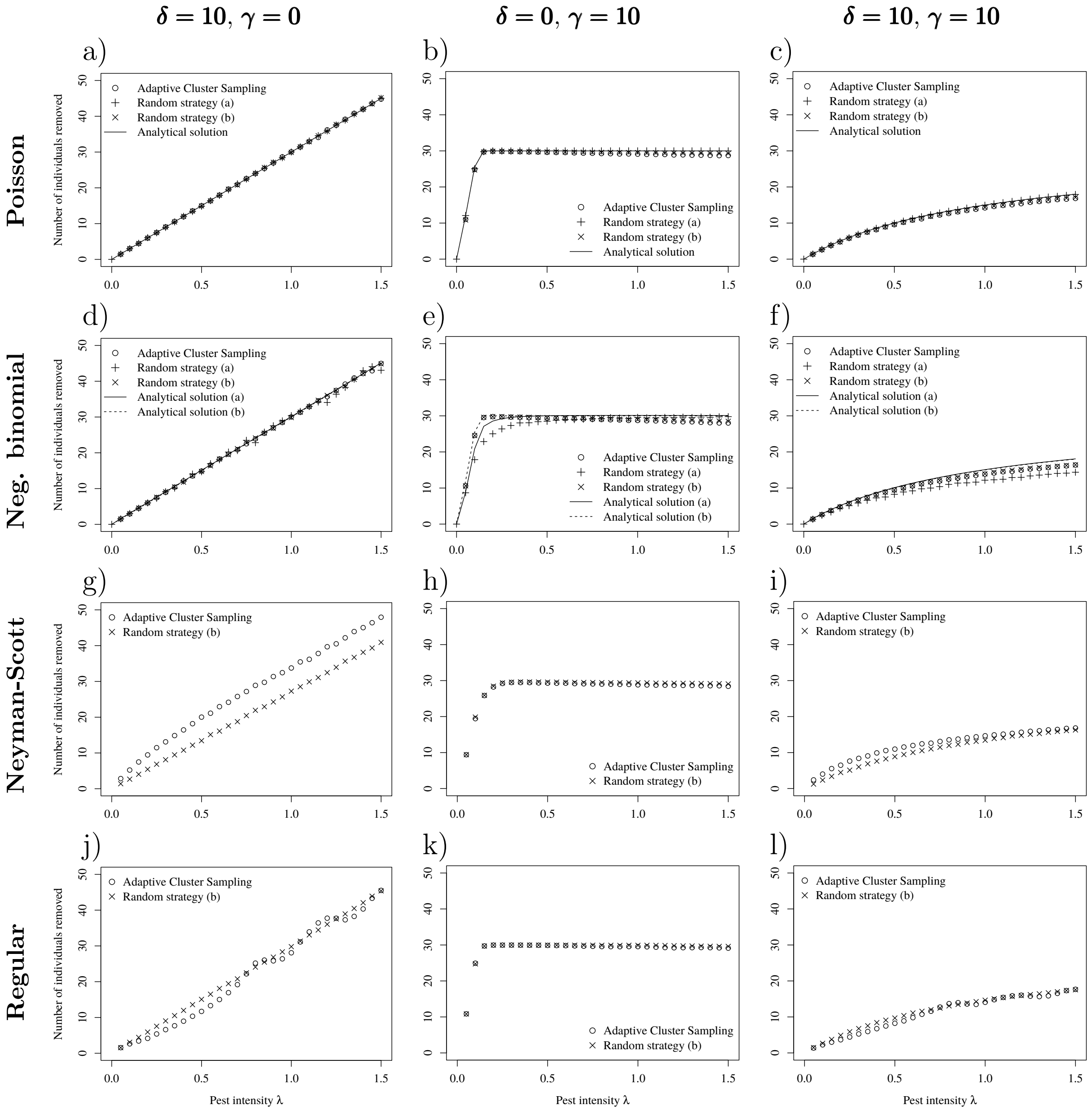

Figure 3.5 – Functional response curves when varying the management strategy and the pest spatial distribution for fixed values of $A = 64$ and $S = 1$ for the random strategy (a) and $A = 1$ and $S = 64$ for the random strategy (b) and the Adaptive Cluster Sampling strategy. Circles represent the means of 2000 simulations of the Adaptive Cluster Sampling strategy, pluses of the random strategy (a), and crosses of the random strategy (b). Solid lines represent the values for the analytical solution of random strategy (a) and (b) in the Poisson case and for the analytical solution of random strategy (a) in the negative binomial case. Dotted lines represent the values for the analytical solution of random strategy (b) in the negative binomial case.

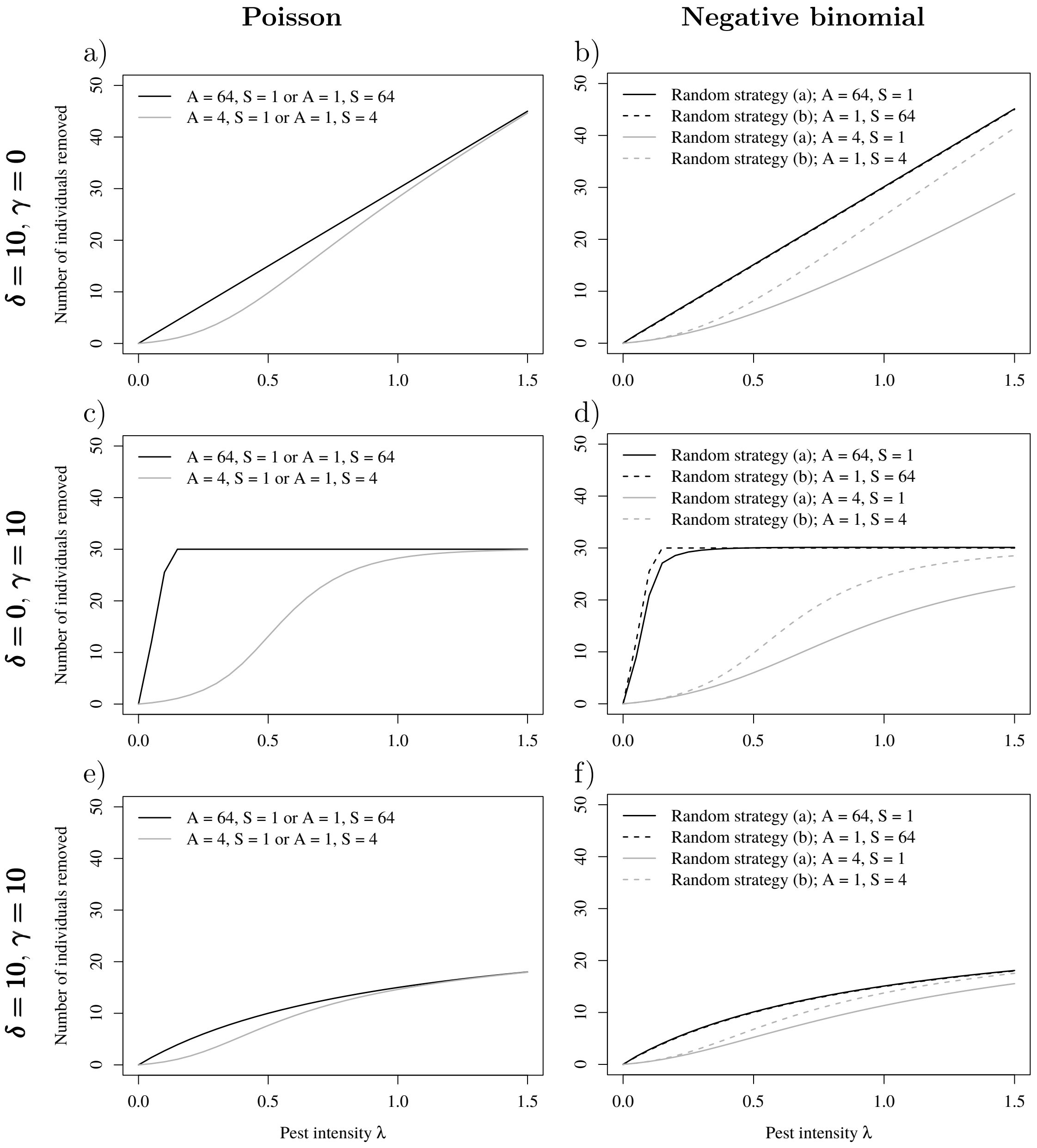

Figure 3.6 – Functional response analytical solutions when varying the cell area A or the number of empty steps to survey in a row S for the random strategy (a) and (b). Black lines represent $A = 64$, $S = 1$ or $A = 1$, $S = 64$ while grey lines represent $A = 4$, $S = 1$ or $A = 1$, $S = 4$. Solid lines represent the values for the analytical solution of random strategy (a) and (b) in the Poisson case and for the analytical solution of random strategy (a) in the negative binomial case. Dotted lines represent the values for the analytical solution of random strategy (b) in the negative binomial case.

and $\gamma = 0$. To obtain a Type I/II, we set $\delta = 0$ and $\gamma > 0$. To obtain a Type II, we set $\delta > 0$ and $\gamma > 0$. To obtain a Type III, we set A or S small, δ small and $\gamma > 0$ (Fig. 3.6).

In addition, variations in budget did not change the qualitative shapes of functional responses.

In conclusion, the pest spatial distribution did not have a large impact on the functional response shape and neither did the search strategy. However, the costs associated with the management strategy as well as the search area had a great impact on the functional response type, as expected. The analytical solutions generally corroborated the simulations although a slight mismatch is apparent, especially for the negative binomial case: the mean values from the simulations of random strategies (a) and (b) are below their respective expected means.

3.3.5 Correspondence with predator functional response

Using Holling's disk equation (Holling, 1959b):

$$N_a = \frac{a'TN}{1 + a'T_hN}, \tag{3.13}$$

we can compare the parameters from predator and manager functional responses. Table 3.4 summarize the parameter equivalences. As stated before, the cost of pest removal (γ) is the manager equivalent of the handling time (T_h). This is corroborated by the similar effect of γ and T_h on functional responses. Indeed, as seen on Fig 3.4, introducing this cost causes the apparition of an asymptote in the curve. Likewise, the survey cost (δ) is the manager equivalent of the searching time. At low densities of prey/pest, the predator/manager spends more of its time/budget on searching rather than handling/removing. This is presented in Eq. 3.13 by the encounter rate a', which represents the searching efficiency or the number of prey items attacked per unit of time at low prey densities. The manager equivalent to a' would then be $1/\delta$. Additionally,

Table 3.4 – Correspondence between predator and manager functional response components.

Predator-prey		Manager-pest	
Nb. of prey items attacked	N_a	Nb. of pest items removed	E
Prey density	N	Pest density	λ
Total foraging time	T	Budget	B
Encounter rate	a'	Detection rate	$1/\delta$
Handling time	T_h	Cost of pest removal	γ
Slope	$a'T$	Slope	B/δ
Asymptote	T/T_h	Asymptote	B/γ

the total foraging time (T) in predator functional response would be the manager's budget (B). Therefore, the slope of the response, corresponding to $a'T$, will be B/δ and the asymptote, corresponding to T/T_h, B/γ (Table 3.4). For $\delta = 10$ and $B = 300$ (Fig. 3.4b) and d)), the slope at low densities is 30 which corresponds to $B/\delta = 300/10$. For $\gamma = 10$ and $B = 300$ (Fig. 3.4c) and d)), the asymptote is 30 which corresponds to $B/\gamma = 300/10$. These comparisons hold for different values of $\delta > 0$ and $\gamma > 0$ (Appendix D, Fig. D.1 and D.2).

These equivalences allow us to use Eq. 3.13 as an approximation for Eq. 3.5 as long as A is large, $\delta > 0$, the management strategy random, and the pest spatial distribution Poisson. See Appendix E, Fig. E.1 for the comparison between the results of Eqs. 3.5 and 3.13.

3.3.6 Application

The fitted pest spatial distribution of the mountain pine beetle infested trees has the mean number of points per cluster 67 ± 55 (standard deviation) and the mean cluster size 266 ± 131. The management strategy has costs of surveying a cell $\delta = 231$ units and cost of removing a pest $\gamma = 181$ units for a budget set to 144 000 units. All actual numbers of attacked trees controlled but one are within 95% of the simulations' distribution as seen on Fig. 3.7. The remaining number is within 99% of the simulations' distribution. The simulations' mean shows a Type II functional response.

According to Table 3.4, we expected that an increase in budget would lead to an increase in slope and asymptote. However, our simulations show that the asymptote increases with an increasing budget but the slope does not (Fig. 3.8).

Applying the adaptive cluster sampling search to Cypress Hills instead of the search described in section 3.2.5 shows that it is the type of search strategy and not the parameter values that is responsible for the constant slope at low pest intensity (see Appendix F, Fig. F.1). Indeed, the adaptive cluster sampling process searches cells until the budget is reached whereas the process described in section 3.2.5 only searches cells around a fixed number of previously infested trees which limits the number of

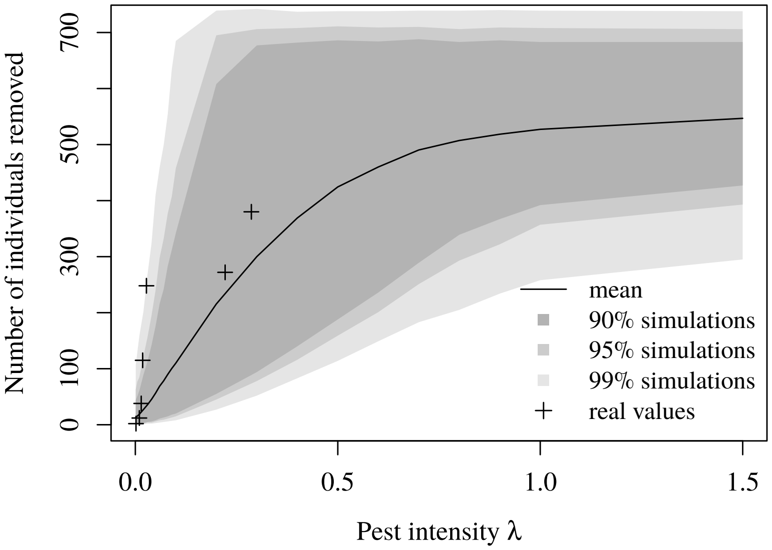

Figure 3.7 – Functional response of the mountain pine beetle management in Cypress Hills, Saskatchewan. The black line represents the mean of 10 000 simulations of the management process. The grey areas represent 90%, 95% and 99% of the simulations' distribution. The points represent the actual values obtained by managers in Cypress Hills in 2011 and 2012.

individuals managers could control with a certain budget.

Fig. 3.8 also shows the management goals. If the functional response curve is on or above the management line, the goal is reached, otherwise not enough individuals are controlled to meet the management goal. If the desired management efficiency is above 20% removal, the goal is unattainable regardless of the budget (Fig. 3.8). If, however, the goal is the removal of all individuals above a certain threshold, it is possible to have an effective management at low pest intensity depending on the budget and the threshold value.

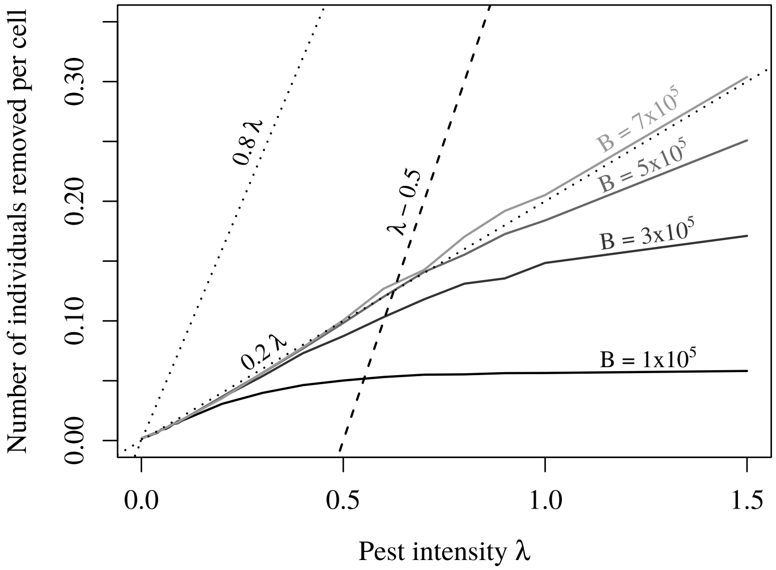

Figure 3.8 – Functional response of the mountain pine beetle management in Cypress Hills, Saskatchewan. Solid lines represent the mean of 1000 simulations of the management process for different budget values: 1×10^5, 3×10^5, 5×10^5, and 7×10^5. Dotted lines represent the management goal 1): removing 80% and 20% of all the individuals. The dashed line represents the management goal 2): removing all individuals above the threshold 0.5 individual per cell.

3.4 Discussion

Considering that pests to managers were like prey to predators, we were able to draw an analogy between functional responses in predator-prey interactions and human-pest management. The searching and handling time were replaced by the surveying and removal costs. The costs associated with the management strategy as well as the search area were the main factors influencing the functional response type as expected. The point process and the search strategy impacted the manager efficacy to a lesser extent. Our framework was applied to the mountain pine beetle epidemic in Cypress Hills, Saskatchewan, Canada.

There is a slight mismatch between the analytical solutions of the two random strategies and their simulations. Indeed, the mean values from the simulations are below their respective expected means. In the simulations, if a step would put managers over the budget, only a fraction of the area would be actually surveyed and controlled (random strategy (a)) or the step would not be taken (random strategy (b)). Therefore, the whole budget might not be used due to rounding in the calculations and the number of pest controlled would then be slightly lower than what it could have been without this process. This could explain the differences between simulation means and analytical solutions of the random strategies (a) and (b). However, the functional response shape does not change between the simulations and the analytical solutions. Accordingly, we can assume that the functional response shapes in the case of the adaptive cluster sampling and in the case of the Neyman-Scott and regular pest distributions are not biased by the budget handling process of the simulations although the values might be slightly underestimated.

The pest spatial resolution was chosen at the scale of an area-A grid cell. This resolution does not impact the Poisson results when the area changes as the sum of Poisson-distributed variables is also Poisson-distributed. However, this is not the case for negative binomially distributed pests. Changing the cell area changes the distribution grain. Therefore, random strategies (a) and (b) consider a different distribution grain for the negative binomial distribution and the results cannot be compared. For management simulations involving the Neyman-Scott and regular distributions, we only consider random strategy (b) and the adaptive cluster sampling strategy which have the same cell size $A = 1$ and, therefore, can be readily compared.

In the functional response theory, the predator distribution is assumed homogeneous. The impact of a clumped predator population on the kill rate lead to the establishment of the aggregative response theory (Cosner *et al.*, 1999). However, the functional response theory originally assumes also a homogeneous distribution of prey in the domain. This could weaken the application of this framework to real systems where the prey distribution is often heterogeneous as well as dynamic in time (Arditi & Ginzburg, 1989; Ives *et al.*, 1999). Several studies mention this issue. Nachman (2006) found that switching the prey spatial distribution from random to aggregated changed the functional response type from a Type II to a Type III as predators adopt a non-random searching behaviour and aggregate in prey clusters. Hossie & Murray (2016) found that for ratio-dependent functional responses, *i.e.* functional responses depending on the density of predators, the pest spatial distribution changed the functional response shape. Rincon *et al.* (2017) found that the functional response shape differs with the difference between the predator and the prey distributions and with the predator foraging strategy. Those studies differ from ours by the fact that several predators interact and the interaction between predator and prey aggregation patterns lead to a change in the functional response shape. However, in our case of a single predator/manager, we highlight the minimal impact of the point process on the functional response shape. Thus, for prey-dependent functional responses, in opposition to ratio-dependent functional responses, the prey spatial distribution seems to have little significance.

Similarly, our results show the minimal impact of the management strategy on the functional response shape. This agrees with Berec *et al.* (2015) who found that survey spatial arrangements (random or regularly spaced) have little impact on the pest detection probability. This is an important information for pest managers who can then minimize costs by choosing a cheaper strategy. However, we should keep in mind that when the pest is spatially clustered, a method resembling the adaptive cluster sampling provide slightly better results.

Unlike other studies focused on predator functional responses when the predator uses a random search strategy (*e.g.* Avgar *et al.*, 2011), our results show that the mean number of individuals controlled varies, to some degree, with the spatial distribution of pests. Avgar *et al.* (2011) found that when employing a random search strategy, the expected number of individuals managed should be the same regardless of the individual spatial distribution because the probability to have at least one individual in a random cell becomes independent of the adjacent cells. If we take into account the specifics of our study, we can explain the difference. Considering random search strategy (a): multiple contiguous size-1 cells are randomly selected so the number of individuals controlled is spatially correlated for the Neyman-Scott and regular processes. Therefore, the expected number of individuals managed in those cases will be different than the expected number of individuals managed from a Poisson process or a negative binomial distribution. The Poisson and negative binomial cases should have the same means since the cells are not spatially correlated. Considering random search strategy (b): with a Neyman-Scott spatial pattern it is more likely to get multiple consecutive empty cells before stopping than for a Poisson spatial pattern. For a regular pattern, it is less likely. The expected number of individuals managed for a Neyman-Scott and regular processes would be different than for a Poisson process or a negative binomial distribution. Again, the Poisson and negative binomial cases should have the same means since the cells are not spatially correlated.

We show that non-random search (adaptive cluster sampling) on an aggregated pest spatial distribution such as the one produced by a Neyman-Scott process leads to higher numbers of pest controlled than on a random pest distribution. However, this is not the case on a Negative binomial distribution. This difference is due to the details of the adaptive cluster sampling process. Indeed, in this strategy, the managers make decisions on which cells to survey. Considering that the Neyman-Scott process produces aggregation among cells whereas the Negative binomial process produces aggregation within cells with the among-cell pattern resembling a random pattern, it makes sense for the random and adaptive cluster sampling strategies to be similar on a Negative binomial pest spatial distribution but different on a Neyman-Scott distribution. Managers following a non-random search strategy on a Neyman-Scott pest spatial distribution are more efficient than managers following a random search since they explore the neighbouring cells when they find pests in a cell. This result agrees with Nachman (2006) who found in the analysis of their predator-prey model that predators searching non-randomly on an aggregated prey distribution have higher predation rates than random search.

Functional response theory is widely used to assess the impact of a predator on a prey population (*e.g.* Messier, 1994; Finke & Denno, 2002). This framework provides a well-studied and reliable method to assess the impact of management on a pest population. Furthermore, knowing aspects of the functional response can inform pest management in several different ways. For instance, human management could be efficiently and quantitatively compared to other control methods such as biological control. This idea would be extended to give comparisons across different methods and different locations. For instance, different methods could be compared at the same location or the same method could be compared at different locations by contrasting functional response curves obtained analytically or by simulation.

To give further insight, manager functional responses could be included in dynamical systems to represent human impact on a pest population. For example, in the Cypress Hills case, future steps could include using this functional response in a population dynamics model to study the evolution of the beetle population over the years with a realistic incorporation of the management effect on the population. In the same way functional response curves inform about stability of predator-prey interactions (Dick *et al.*, 2013), in our study, they could tell managers whether control would leave a refuge for the pest or lead to the pest extinction. Indeed, in a Type III functional response scenario, management is less efficient at low pest densities and, thus, small pest populations could persist (Murdoch & Oaten, 1975). However in a Type II functional response scenario most pests would be eradicated, even if they occur at low density (Hassell, 1978). With respect to control of mountain pine beetle in Cypress Hills, the Type II functional response suggests a high efficiency of managers at low pest densities. This is consistent with the strict management policy in Cypress Hills to try to control all infested trees. The fact that true values for the number of individuals controlled lie above those simulated in Fig. 3.7 suggests that managers in Cypress Hills are likely finding the means to make the process even more efficient than our model would predict.

From a practical perspective, managers are typically interested in controlling a certain proportion of pests, given an environmental context and a particular strategy. For example, in some areas of the province of Alberta, Canada, infested by the mountain pine beetle, the management goal is to reduce populations by 80% (Alberta Sustainable Resource Development, 2007) using individual tree removal. Our model results, as shown in Fig. 3.7 and 3.8, allow for such a calculation to be made. By comparing this goal to the detection and control rate (= slope of the functional response curve; equivalent to the attack rate), the management functional response can give direct insights on the goal feasibility. We illustrated this point using the Cypress Hills case study. The results showed that an 80% control efficiency, such as the target used in some areas in Alberta, would not be feasible in Cypress Hills given the current management strategy and parameters values. If the functional response resembled a type III instead of a type II, there could be cases where management would be efficient only at intermediate pest intensity. However, decreasing the pest population below a certain threshold would be more appropriate than a removal percentage in the Cypress Hills case. Goodsman & Lewis (2016) found Allee threshold values of 3.789×10^{-4} and 5.311×10^{-5} infested stems per m^2 for two times series of mountain pine beetle infestation in central Idaho. Therefore, a management threshold of 0.5 individuals per 100×100m cell such as the one used in the Cypress Hills example would be below such an Allee threshold. Moreover, we showed that the pest intensity in Cypress Hills is always below 0.5 individual per cell which means that managers are probably efficiently reducing local mountain pine beetle populations below the Allee threshold. However, a non-spatial Allee threshold could be an issue as the within-tree productivity of the beetle varies in time and space (Goodsman & Lewis, 2016). As an alternative, varying the management strategy, such as adding search locations beyond the neighbourhood of previously infested trees until the budget is reached, would allow managers to meet a certain removal percentage depending on the budget amount.

In summary, functional response is a tool to help pest management, for example by providing a means to assess current strategies, to compare with alternative strategies, to test various strategies *in silico* before implementation, to provide a realistic control component in a population dynamics model, and to assess the feasibility of a management goal.

One constraint to the application of this method, as in regular functional response studies, is the need for several levels of pest density to be able to draw a functional response curve. In our application to mountain pine beetle, we got around this issue by dividing the infested area in portions of different pest densities. This might not be sufficient to draw an entire functional response curve but simple simulations could help complete the curve in this case.

Researchers often complement the functional response with a numerical response in predator numbers and sometimes an aggregative response describing the distribution of predators (Turchin, 2013). We could argue that in a human-pest management scenario, the numerical and aggregative responses are negligible or even non-existent. Indeed, the number or aggregation level of managers might not be relevant for two reasons: 1) managers usually group as a team/unit, in our study, this unit represent "the manager" and 2) there is no competition between managers or teams, when enough persons are present to make several teams they often do not survey the same area but rather divide the entire management domain between themselves in order to be efficient. However, one could study the social impact of pest management and thus be interested in new hires in response to an increasing pest abundance. This could be modelled by a growing manager population and be analogous to predator numerical responses. Finally, the cost of moving from one location to another was not taken into account in our study for simplicity but it could be added easily.

Chapter 4

Factors governing outbreak dynamics in a forest managed for mountain pine beetle

4.1 Introduction

The mountain pine beetle (MPB; *Dendroctonus ponderosae*, Hopkins 1902) epidemic behaviour in North American pine forests is causing massive ecological consequences and losses to the timber industry (Walton, 2013) as well as threatening cultural and tourism activities (Dhar *et al.*, 2016). Although MPB is native to part of North America, its spread is considered human-induced as it is a function of climate change and fire suppression (Bentz *et al.*, 2010). Because of the influence of human actions on MPB spread combined with the ecological and economic impact of MPB epidemics, this species fits the description of an invasive species (Lockwood *et al.*, 2013). As a consequence, MPB outbreaks are actively monitored and heavily controlled in Canadian pine forests (Alberta Sustainable Resource Development, 2007; Saskatchewan Ministry of Environment, 2016). Managers face several challenges related to detection and control. An efficient control is direct, early, aggressive, and continuous until the outbreak is suppressed (Carroll *et al.*, 2006). To be able to implement such control, managers need to have efficient detection methods. In turn, detection could be improved by including different ecological and environmental factors depending on the outbreak phase.

From the perspective of the biology of MPB outbreaks, four major phases have been described: endemic, incipient-epidemic, epidemic, and post-epidemic (Shore *et al.*, 2004). Managers, tracking the rise and fall in numbers of infested trees may not easily be able to identify the endemic or early incipient-epidemic phase. This leads to an alternative categorization based on infested tree numbers, which we describe here and use in this paper: onset (increasing number of infested trees; typically late incipient-epidemic and early epidemic), peak (high and constant number of infested trees; late epidemic) and collapse (decreasing number of infested trees; post-epidemic). See Fig. 4.1 for a representation of the two approaches to categorizing MPB outbreaks.

We now briefly review the biology related to the major phases of MPB outbreaks, starting with the endemic phase. Beetles in the endemic phase attack weak or stressed pines in the summer, sometimes with the help of other bark and wood boring beetles, by drilling galleries under the bark and laying their eggs (Safranyik & Carroll, 2006). Over the fall and early winter, the eggs become larvae then pupae before emerging as adults in the summer of the following year. Adults do not usually survive the winter. The pine hosts are typically killed by the MPB development process and their crowns fade to a red colour within one year after the attack. The range of MPB endemic populations is widespread in north America, covering north-west United States and western Canada.

When a MPB population has enough individuals to successfully attack large and healthy trees on its own, the population transitions from an endemic phase to an outbreak. Outbreaks include the three phases: incipient-epidemic, epidemic, and post-epidemic. Beetles in the incipient-epidemic phase reach a number large enough to successfully attack large trees, although spread is slow and limited. Then, as population size increases, and single-tree infestations become patches containing multiple infested trees, the MPB population enters the epidemic phase. The post-epidemic phase is characterized by population decline in MPB. An epidemic usually lasts several years if sufficient

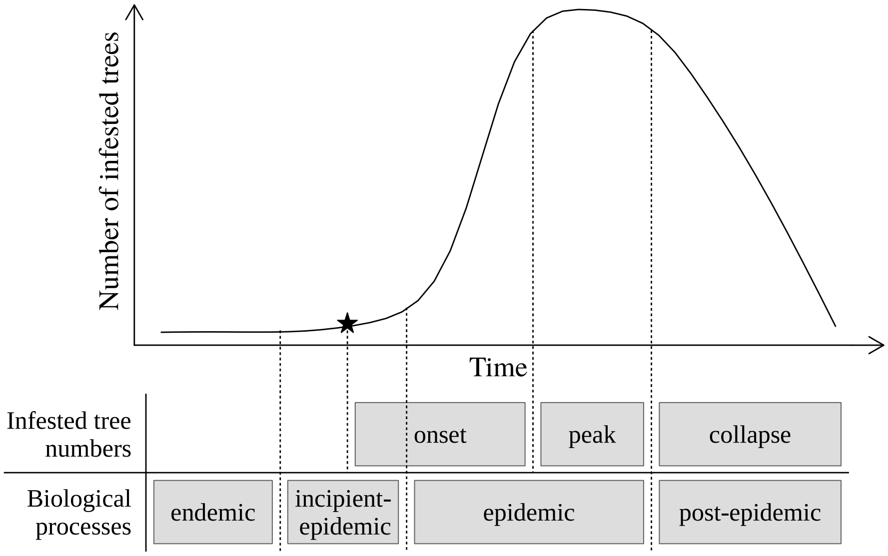

Figure 4.1 – Representation of the biological processes and infested tree numbers approaches to categorizing MPB population phases. The star represents the first detection of the MPB outbreak. The endemic phase has less than one mass-attacked tree per stand.

host pines are present, with an average of approximately 10 years in British Columbia (Safranyik, 2004).

Commonly, red-top trees, which are the distinct dead pines infested by MPB individuals from the previous year, are used to estimate the presence of new infestations in the area (Alberta Sustainable Resource Development, 2007; Saskatchewan Ministry of Environment, 2016). While single-tree infestations may escape detection, the transition to infested patches make the infestations more likely to be found. As more single trees become infested, MPB is detected and enters the outbreak onset. At this point, managers face three outbreak phases, based on infested tree numbers: onset, peak, collapse.

During the outbreak, the locations of infested trees can be georeferenced to characteristic ecological and environmental factors. However, these factors have different roles on MPB population dynamics at the various stages of an outbreak (Aukema *et al.*, 2008; Preisler *et al.*, 2012). Understanding these roles provides an opportunity to improve detection methods through a systematic evaluation of cues from ecological and environmental factors.

We now turn to a review of ecological and environmental factors associated with the spatial locations of infested trees in different outbreak phases. We characterize this review according to host tree properties, beetle pressure, and weather factors. Host tree properties affect susceptibility to infestation. Beetle pressure provides the source of infestation. Weather factors impact the details of life-history and environmental stress of both the beetles and the trees. Collectively, these factors determine the outbreak level and duration.

Host tree abundance, resistance, and size impact MPB infestation differentially, depending on the phase of an outbreak (Safranyik, 2004). Indeed, an MPB endemic population first needs sufficient small and weak/stressed trees in order to increase the population size to outbreak levels and attack larger and healthier trees (Safranyik, 2004; Nelson *et al.*, 2018). MPB population decline happens in the post-epidemic phase when the number of susceptible pines decreases and MPB switches back to weaker and smaller trees. This decline in the epidemic that is associated with the reduction in the susceptible population is a common feature characterizing epidemic processes (Kermack & McKendrick, A. G., 1927).

Beetle pressure is needed for MPB presence at all outbreak phases (Preisler *et al.*, 2012). It describes the source of a new beetle generation. Outbreak onset relies on local endemic population increase and/or contributions from outside sources via dispersal whereas established outbreaks relies on adjacent sources (Safranyik, 2004; Aukema *et al.*, 2008; Nelson *et al.*, 2018).

Among weather factors, temperatures have the greatest impact on MPB population life-history. Warm winter temperatures allow MPB individuals in endemic phase to survive the cold season in greater proportions, thereby potentially increasing their population size to outbreak levels (Aukema *et al.*, 2008; Preisler *et al.*, 2012; Sambaraju *et al.*, 2012; Creeden *et al.*, 2014). However, cold snaps in fall or early spring, or generally lower winter temperatures, can lead to outbreak collapse (Safranyik, 2004; Creeden *et al.*, 2014). Average temperatures over the year as well as spring and summer temperatures are indicators of beetle development rate and timing. Warmer temperatures allow individuals to complete their life-cycle (Preisler *et al.*, 2012; Creeden *et al.*, 2014) and high temperatures during flight periods increase attack success rate by increasing spatial synchrony (Aukema *et al.*, 2008). However, excessively high temperatures during the summer can decrease emergence rate as well as dispersal success (McCambridge, 1971; Safranyik & Carroll, 2006). In summary, warm temperatures are crucial to MPB development but this positive effect on MPB infestation can be counteracted when temperatures become too high for successful emergence and dispersal.

Rainfall is also an essential weather factor governing MPB populations. Water deficit lowers pine defenses against MPB attacks (Erbilgin *et al.*, 2017b). Therefore, drought can help MPB endemic populations successfully attack sufficient trees to increase their population size to outbreak levels (Safranyik, 2004; Aukema *et al.*, 2008). However, it is also necessary to have abundant vigorous trees for a successful outbreak (Safranyik, 2004; Nelson *et al.*, 2018) and tree vigour can be reduced by drought (Matthews *et al.*, 2018). Nonetheless, drought may not be sufficient to decrease vigour levels to the point of a suppressed outbreak. For example, entire outbreaks in western US were exposed to drought (Creeden *et al.*, 2014). Therefore, the timing and intensity of drought can either help or hinder MPB populations.

Much of our understanding of the ecological and environmental factors governing MPB outbreaks come from the conceptual and observational work synthesized by Safranyik, Carroll and coworkers (Safranyik, 2004; Safranyik & Carroll, 2006). However, a deeper quantitative analysis requires detailed connections between models and data via statistical inference. Such analyses exist, but, to date, have employed large spatial scales, typically with different sub-regions in different outbreak stages (Aukema *et al.*, 2008; Preisler *et al.*, 2012). To the best of our knowledge, there has been no local-scale statistical analysis of ecological and environmental factors governing MPB outbreaks, from onset to collapse, where the outbreak phases are relatively synchronized across the study site. The recent Cypress Hills MPB outbreak in Saskatchewan provides a unique opportunity to do this very thing. Located far from the main lodgepole pine range, the Cypress Hills MPB infestation is isolated from other outbreaks. The Cypress Hills park spatial scale (184 km^2) is such that the outbreak phases have been relatively synchronized spatially. The data set is very high quality as the region was completely censused for MPB infection yearly from the onset in 2006 up to the current collapse in 2018. This provides a unique opportunity to follow a single outbreak in a fixed location at a small spatial scale, and to perform a comprehensive statistical analysis of ecological and environmental factors influencing each outbreak phase.

Our study aims to 1) determine the local ecological and environmental factors driving MPB presence for the different phases of an outbreak (onset, peak, and collapse) in a forest managed for MPB, 2) assess the degree to which the models predict MPB presence for each outbreak phase, and 3) show how selected factors have differing impacts on MPB presence depending on the outbreak phase. For each phase, we hypothesized that MPB presence depends on a combination of weather, beetle pressure, control, and host-related variables. We model those relationships using logistic regressions during the onset, peak, and collapse phases of one MPB outbreak studied in the Cypress Hills interprovincial park in Saskatchewan, Canada.

4.2 Material and methods

4.2.1 Study area and data

We use data from the Saskatchewan portion of the Cypress Hills interprovincial park, located at the border between the provinces of Alberta and Saskatchewan, Canada. This portion of the park covers 184 km^2. MPB infestation data and ecological and environmental covariates from this region provide an opportunity to connect local factors to outbreak phases. These data were studied in Kunegel-Lion *et al.* (2019), so our discussion of the study area follows this reference closely. The Forest Service Branch of the Saskatchewan Ministry of Environment has implemented a "zero-tolerance" policy as of 2006, designed to catch and control as many newly infested trees as possible. The policy operates according to the following procedure. In early fall, after MPB have colonized new trees, an aerial survey of the park extent is conducted to collect geo-referenced data on potential red-top trees, which are dead or dying trees infested by MPB the previous year. These are later ground-truthed for MPB attacks. Then, 50 meter-radius circular survey plots are drawn around each of the red-top trees confirmed to have been killed by MPB. The survey plots are searched for infested trees, which are trees recently attacked by MPB during the summer. These are later controlled in late fall/winter, which usually consists of felling and burning massively infested trees or peeling the bark of trees presenting lower numbers of MPB larvae, ensuring that beetles are killed. In addition to these measures, areas presenting high densities of red-top trees are entirely surveyed and controlled. Although a small number of bait trees were used during the study period, our understanding is that this process would likely not change which stands are attacked as the area of effect of the bait is within a stand.

The ecological and environmental covariates and the infestation response values were distributed discretely in space and time. We applied a grid of 18 317 cells of size 100×100 meters to the Cypress Hills park extent. This cell size was chosen to match the size of the management surveys. The fact that a cell's area (10 000 m^2) and a search plot's area (7 854 m^2) are the same order of magnitude make the analogy between grid cell and survey plot possible. For each cell for each year, the observation consisted of a set of ecological and environmental covariates plus the response variable. The response variable was the presence/absence of MPB derived from the presence/absence of infested trees in each cell of the grid based on data from the Forest Service ground survey (Saskatchewan Ministry of Environment, 2016). From these Forest Service surveys, we obtained the locations of infestations controlled by managers and we deduced which trees had been infested in the previous year using the red-top trees. Part of the red-top tree data consists of infested zones and not actual tree locations. Therefore, we deduced the presence/absence of MPB in each cell from the location of infested trees or the overlap of an infested zone with the cell.

At the time of this study, the Forest Service had consistently recorded the mountain pine beetle outbreak from 2006 to 2018. We divided these data into outbreak phases and trained and validated the models separately for each phase. The phases were defined as follow: 1) outbreak onset from 2006 to 2011, the number of infested cells increases; 2) outbreak peak from 2012 to 2013, the number of infested cells is high and relatively constant; and 3) outbreak collapse from 2014 to 2018, the number of infested cells decreases (Fig. 4.2).

We chose the ecological and environmental covariates to represent as much as possible each of the processes described in the introduction. To consider the fact that the MPB's emergence and dispersal is reduced with high temperatures and its survival is compromised if the early stages–developing in fall and winter–are exposed to extreme temperatures (Cole, 1981; Safranyik & Carroll, 2006; Régnière & Bentz, 2007), we included the covariates highest maximum daily temperature during July and August and cold tolerance from Régnière & Bentz (2007). The number of degree-days above

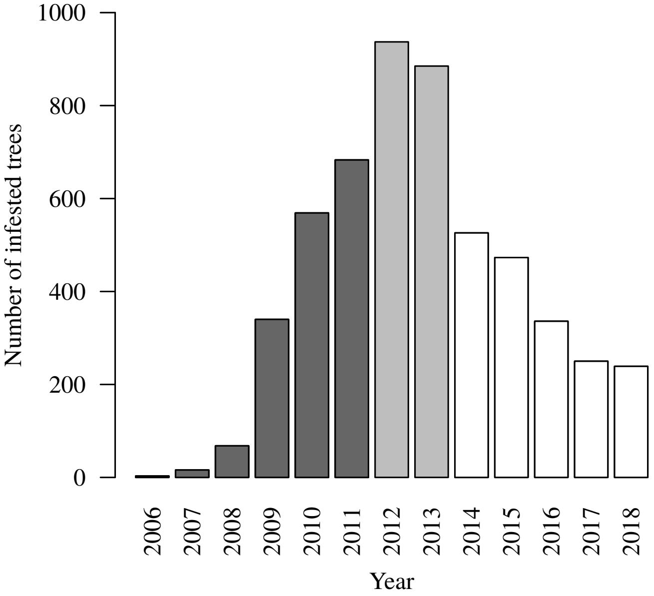

Figure 4.2 – Number of infested trees over time. The darker grey represents the outbreak onset. The grey represents the outbreak peak. The white represents the outbreak collapse.

5.5°C over the year represents MPB development rate and timing. The average daily wind speed in July and August represents the MPB long-distance dispersal dependence on wind (Safranyik & Carroll, 2006; Robertson *et al.*, 2007). To consider the fact that drought–especially in spring–reduces pines' ability to defend themselves and increase MPB attacks' success rate (Safranyik, 1978; Lusebrink *et al.*, 2016), we included the average daily relative humidity in spring. We also included the covariates pine cover, pine height, and pine age to consider the impact of host tree abundance and size on MPB infestation (Safranyik & Carroll, 2006).

The levels of beetle from a given year can be divided into uncontrolled infestation (I_u) and controlled infestation (I_c). Details of these variables are given in Table 4.1. The control was undertaken by managers from the Saskatchewan Forest service under a zero-tolerance policy. However, some infested trees inevitably remained undetected and these were identified as red-tops in the following year. Beetles from these uncontrolled infestations can disperse short distances within and between cells, and thus provide the beetle pressure for new infestations. The MPB presence two years prior to the observation is not included as MPB is generally univoltine (Bentz & Powell, 2014), so we assume that an infested tree can only be a source of beetle for the following year and not the years after that. We also included the distance to the park southern border, which was close to external infestations not managed by the Forest Service and potential sources of MPB. Finally, it is possible to calculate the total infestation by adding the uncontrolled and controlled infestations.

Lastly, we included the northerness and easterness of the slope as it can address bias in MPB dispersal due to the spatial configuration and wind direction.

We estimated weather variables with the BioSIM software (Régnière *et al.*, 2014) at the location of each grid cell centroid using topography data from the Canadian Digital Elevation Map downloaded from the Geogratis website (geogratis.cgdi.gc.ca).

The vegetation data was derived from Beaudoin *et al.* (2014). The authors computed, for each cell of a grid, the coverage of various species of pines and the leading species height and age from 2001 and 2011 MODIS imagery. To calculate the pine height and

Table 4.1 – Description of the variables used in the models.

Name	Description
$T_{\max}$	Highest maximum daily temperature during July and August
RH	Average daily relative humidity in spring
CT	Cold tolerance from Régnière & Bentz (2007)
Cover	Percentage of pine cover
Height	Pine height in meters
I_c	Previous-year controlled MPB infestation level in a 3-cell radius around each location I_c = number of infested cells with all trees controlled at the same location + 0.5 × number of infested cells with all trees controlled in radius 1 + 0.25 × number of infested cells with all trees controlled in radius 2 + 0.125 × number of infested cells with all trees controlled in radius 3 (Fig. 4.3)
I_u	Previous-year uncontrolled MPB infestation level in a 3-cell radius around each location I_u = number of infested cells with uncontrolled trees at the same location + 0.5 × number of infested cells with uncontrolled trees in radius 1 + 0.25 × number of infested cells with uncontrolled trees in radius 2 + 0.125 × number of infested cells with uncontrolled trees in radius 3 (Fig. 4.3)
Dist	Distance to the park southern border close to external infestations (Fig. 4.4)
N	Northerness: spatial property of a slope to face North
E	Easterness: spatial property of a slope to face East

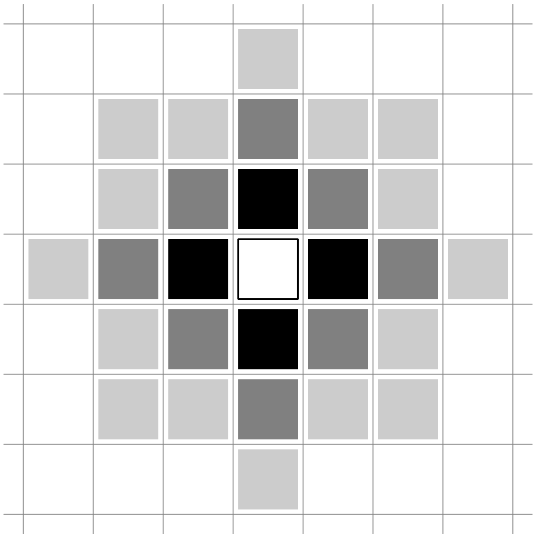

Figure 4.3 – Representation of the adjacent cells taken into account in the covariates (cf. Table 4.1). White: focus cell; dark grey: 4 adjacent cells (radius 1); medium grey: next 8 adjacent cells (radius 2); light grey: next 16 adjacent cells (radius 3). Based on Kunegel-Lion *et al.* (2019).

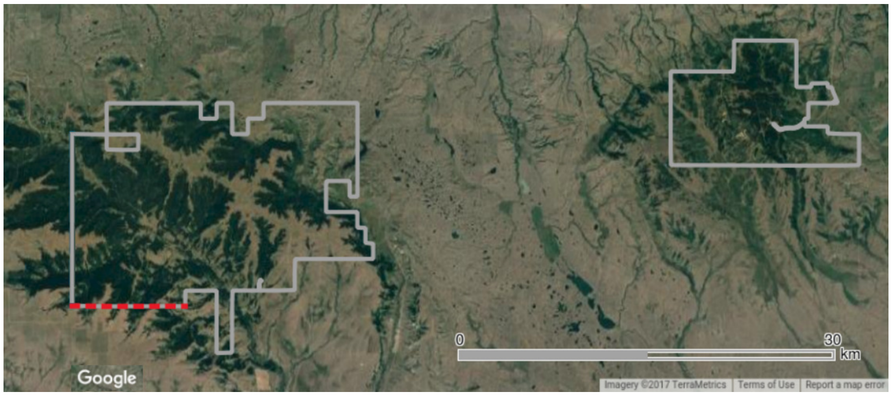

Figure 4.4 – Cypress Hills park boundaries in Saskatchewan (grey). The dashed red line represents the park border close to outside infestations in the South. Based on Kunegel-Lion *et al.* (2019).

age, we used the leading species values when the pine cover was more than 50%. We spatially interpolated the vegetation variables values at the location of our grid cell centroids using bicubic spline interpolation with the `interp` function of the `R` package `akima` (Akima & Gebhardt, 2016). Since the data were only available for 2001 and 2011, the vegetation variables were linearly interpolated over our study period.

4.2.2 Data analysis

In this study, we 1) test the hypothesis that MPB presence depends on a combination of weather, vegetation, topography, and beetle and host-related variables, 2) select relevant variables for each outbreak phase, and 3) show whether the selected variables have a different impact on MPB presence depending on the outbreak phase.

To test the hypothesis that MPB presence depends on a combination of weather, vegetation, topography, and beetle and host-related variables, we used a logistic regression where the probability of MPB presence $\pi(\underline{\beta})$ depend on parameters β_i and ecological and environmental covariates X_i, described in the previous section, as defined by

$$\pi(\underline{\beta}) = \frac{e^{\beta_0 + \sum \beta_i X_i}}{1 \; + \; e^{\beta_0 + \sum \beta_i X_i}}. \tag{4.1}$$

We trained the logistic regressions on each phase separately using the `train` function of the `R` package `caret` (R Core Team, 2018; Kuhn, 2018).

For each outbreak phase, we implemented a multiple working hypothesis approach. We used the exhaustive enumeration of subsets method (Sokal & Rohlf, 1995). This method compares all possible combinations of the covariates and selects the best models among the ones sharing the same number of covariates. We selected the best model overall and the best model per number of covariates using the Bayesian Information Criterion (BIC; Schwarz, 1978). With our goal of determining which factors are associated with MPB infestation at each outbreak phase, we chose the BIC over the Akaike Information Criterion (AIC; Akaike, 1974) for model comparison. This is because BIC is better indicator of the “true” model whereas AIC is more suited to determine which models should be used for predictions (Ghosh & Samanta, 2001; Elliott & Brook, 2007; Aho *et al.*, 2014). Additionally, our number of observations was very large (238 121 observations) compared to the parameter space (11 parameters) which also favours the BIC over the AIC. As with the AIC index, a low BIC means a good trade-off between the goodness of fit of the model and model complexity. Two models with a BIC difference less or equal to 2 are considered indistinguishable whereas a BIC difference of 8 or more provides strong evidence for the model with lower BIC (Burnham & Anderson, 2002; Potapov *et al.*, 2011).

To be able to differentiate the effect of each covariate, we removed highly correlated covariates ($|\rho| > 0.6$) from our analysis and we also removed combinations of covariates with a variance inflation factor (VIF) greater than 10 to handle the potential remaining multicollinearity (Kutner *et al.*, 2004). Therefore, we did not include the number of degree-days in the analysis as it was correlated with the maximum temperature ($\rho = 0.68$) over the entire time period. This can also be justified by the fact that the minimum number of degree-days above 5.5°C needed for MPB development (833; Safranyik *et al.*, 1975; Carroll *et al.*, 2006; Safranyik *et al.*, 2010) was always exceeded, with observed degree-days of 1054 and higher. Furthermore, high numbers of degree-days is not an issue with respect to inducing multivoltinism as this is rarely present in MPB (Bentz & Powell, 2014). We also did not include the wind speed, which was highly correlated with the distance to the park southern border ($\rho = -0.73$) over the entire time period. Indeed, both of these covariates can be connected to the impact on the infestation of the possible sources of MPB outside the park limits, which were not recorded. As might be expected, pine height and age were also correlated ($\rho = 0.95$). Therefore, we only kept pine height in our analyses. Note that the correlation coefficient for the controlled and uncontrolled infestations is 0.26, therefore the impact of these covariates on the infestation should be possible to differentiate.

To assess the performance of the selected models, we performed cross-validation with folds defined by year. For each outbreak phase, we tested the accuracy of the selected models on the test sets using the area under the receiver operating characteristic curve

(AUROC; Metz, 1978; Bradley, 1997) and the area under the precision-recall curve (AUPR; Raghavan *et al.*, 1989; Saito & Rehmsmeier, 2015).

A receiver operating characteristic (ROC) curve (Metz, 1978) depicts, for a range of probability thresholds, the true positive rate (or 1 - false negative rate, also referred to as sensitivity or recall) against the false positive rate (also referred to as 1 - specificity). A precision-recall curve (Raghavan *et al.*, 1989) depicts, for a range of probability thresholds, the proportion of true positives amongst the positive predictions (also referred to as precision or positive predictive value) against the true positive rate (sensitivity/recall). For the reader's convenience, more details on how to calculate these indices are given in Appendix G.

A high AUROC or AUPR ($0 \leq$ AUROC/AUPR ≤ 1) represents a good performance of a binary classifier in terms of correspondence between observed and predicted values. A null model has an AUROC of 0.5 and a AUPR value equals to the proportion of positive outcomes in the data. The precision-recall curve is more informative than the ROC curve for imbalanced data sets (Davis & Goadrich, 2006; Saito & Rehmsmeier, 2015) which is the case here as the rate of 0 to 1 in our three data sets is between 40:1 and 95:1.

To show whether the selected covariates have a different impact depending on the outbreak phase, we compared the order of importance of the standardized estimates β_i. Within a model, a large negative or positive β_i has, respectively, a large negative or positive impact on MPB presence whereas a small β_i has a small impact on MPB presence.

4.3 Results

For the outbreak onset, the best model used nine covariates: northerness, maximum temperature in July and August, relative humidity, cold tolerance, distance to the infested border, pine cover, pine height, and controlled and uncontrolled infestations (BIC = 8583.6; Table 4.2). However, the model without northerness or pine height gave a ΔBIC < 8, casting doubt on the importance of these indices on the MPB presence for the onset.

For the peak of the outbreak, the best model used 5 covariates: easterness, maximum temperature in July and August, pine height, and controlled and uncontrolled infestations (BIC = 5500.9; Table 4.3). However, the model with the covariate cold tolerance instead of easterness had ΔBIC < 8 which implies that cold tolerance and easterness could be interchanged.

The best model for the outbreak collapse used 5 covariates: maximum temperature in July and August, relative humidity in spring, distance to the infested border, and controlled and uncontrolled infestations (BIC = 7436.9; Table 4.4). However, the model without the maximum temperature gave a ΔBIC < 8, casting doubt on the importance of this variable on the MPB presence for the collapse.

For each phase, the selected models have high AUROC indicating a high level of predictive ability ($\text{AUROC}_{\text{onset}} = 0.874$, $\text{AUROC}_{\text{peak}} = 0.913$, and $\text{AUROC}_{\text{collapse}} = 0.936$; Tables 4.2–4.4). Compared to the null models, the AUPR values are consistent with the higher AUROC values ($\text{AUPR}_{\text{onset}} = 0.174$ with $\text{AUPR}_{\text{null}} = 0.010$, $\text{AUPR}_{\text{peak}} = 0.427$ with $\text{AUPR}_{\text{null}} = 0.024$, $\text{AUPR}_{\text{collapse}} = 0.306$ with $\text{AUPR}_{\text{null}} = 0.011$; Tables 4.2 to 4.4). The relatively high AUPR values show that the models predict well MPB presence without potentially wasting too much management resources on false alerts, which are incorrectly-predicted MPB presence.

For the outbreak onset, the order of the covariates by importance (absolute standardized estimates) is: maximum temperature in July and August, distance to the infested border, cold tolerance, uncontrolled infestations, relative humidity in spring, pine cover, controlled infestations, pine height, northerness (Fig. 4.5). The order and the selected covariates differ from the peak: uncontrolled infestations, pine height, easterness, maximum temperature in July and August, controlled infestations, and the collapse: uncontrolled infestations, controlled infestations, relative humidity in spring, distance to the infested border, maximum temperature in July and August (Fig. 4.5).

Table 4.2 – Comparison of the models' BIC, AUROC, and AUPR for the outbreak onset. All models are compared to the one with the lowest BIC using ΔBIC. For each number of variables, we show the best model in black and competing models with a difference of BIC ≤ 8 in grey. The model in bold is the one selected from the ΔBIC ≤ 2. "AUROC" stands for the area under the ROC curve, "AUPR" stands for the area under the precision-recall curve.

Size	Selected variables	VIF$_{\max}$	BIC	ΔBIC	AUROC	AUPR
0	*null*		12705.2	4121.6	0.500	0.010
1	I_u	1.0	10121.8	1538.3	0.650	0.154
2	I_u, I_c	1.0	9400.6	817.1	0.683	0.184
3	$T_{\max}$, I_u, I_c	1.1	8929.6	346.0	0.797	0.172
4	$T_{\max}$, Dist, I_u, I_c	1.2	8769.1	185.5	0.873	0.183
5	$T_{\max}$, CT, Dist, I_u, I_c	1.7	8661.0	77.4	0.867	0.182
6	$T_{\max}$, CT, Dist, Cover, I_u, I_c	1.7	8628.7	45.1	0.874	0.181
7	$T_{\max}$, RH, CT, Dist, Cover, I_u, I_c	3.0	8592.0	8.4	0.870	0.180
8	N, $T_{\max}$, RH, CT, Dist, Cover, I_u, I_c	3.0	8587.5	4.0	0.873	0.176
8	$T_{\max}$, RH, CT, Dist, Cover, Height, I_u, I_c	3.0	8589.9	6.3	0.871	0.178
9	**N, $T_{\max}$, RH, CT, Dist, Cover, Height, I_u, I_c**	**3.0**	**8583.6**	**0.0**	**0.874**	**0.174**
10	N, E, $T_{\max}$, RH, CT, Dist, Cover, Height, I_u, I_c	3.0	8594.8	11.2	0.874	0.173

Table 4.3 – Comparison of the models' BIC, AUROC, and AUPR for the outbreak peak. All models are compared to the one with the lowest BIC using ΔBIC. For each number of variables, we show the best model in black and competing models with a difference of BIC ≤ 8 in grey. The models in bold are the ones selected from the ΔBIC ≤ 2. "AUROC" stands for the area under the ROC curve, "AUPR" stands for the area under the precision-recall curve.

Size	Selected variables	VIF$_{\max}$	BIC	ΔBIC	AUROC	AUPR
0	*null*		8341.5	2842.1	0.500	0.024
1	I_u	1.0	5557.7	58.3	0.879	0.436
2	Height, I_u	1.0	5540.7	41.3	0.902	0.430
3	Height, I_u, I_c	1.4	5520.4	21.0	0.907	0.427
4	E, Height, I_u, I_c	1.4	5507.8	8.4	0.907	0.429
5	**E, $T_{\max}$, Height, I_u, I_c**	**1.4**	**5500.9**	**1.5**	**0.913**	**0.427**
5	$T_{\max}$, CT, Height, I_u, I_c	3.1	5505.1	5.7	0.904	0.416
6	**E, $T_{\max}$, CT, Height, I_u, I_c**	**3.1**	**5499.4**	**0.0**	**0.909**	**0.421**
7	E, $T_{\max}$, CT, Cover, Height, I_u, I_c	3.2	5505.9	6.6	0.909	0.419
8	E, $T_{\max}$, CT, Dist, Cover, Height, I_u, I_c	3.4	5514.7	15.3	0.908	0.419
9	N, E, $T_{\max}$, CT, Dist, Cover, Height, I_u, I_c	3.5	5524.3	24.9	0.906	0.422

Table 4.4 – Comparison of the models' BIC, AUROC, and AUPR for the outbreak collapse. All models are compared to the one with the lowest BIC using ΔBIC. For each number of variables, we show the best model in black and competing models with a difference of BIC ≤ 8 in grey. The model in bold is the one selected from the ΔBIC ≤ 2. "AUROC" stands for the area under the ROC curve, "AUPR" stands for the area under the precision-recall curve.

Size	Selected variables	VIF_{max}	BIC	ΔBIC	AUROC	AUPR
0	*null*		11071.0	3634.1	0.500	0.011
1	I_u	1.0	8182.2	745.3	0.864	0.264
2	I_u, I_c	1.0	7498.5	61.6	0.929	0.306
3	RH, I_u, I_c	1.0	7465.1	28.2	0.925	0.305
4	RH, Dist, I_u, I_c	1.1	7443.9	6.9	0.932	0.305
5	**T_{max}, RH, Dist, I_u, I_c**	**1.1**	**7436.9**	**0.0**	**0.936**	**0.306**
6	T_{max}, RH, Dist, Cover, I_u, I_c	1.1	7443.5	6.5	0.936	0.304
7	E, T_{max}, RH, Dist, Cover, I_u, I_c	1.1	7453.0	16.1	0.935	0.302
8	N, E, T_{max}, RH, Dist, Cover, I_u, I_c	1.1	7463.8	26.9	0.935	0.301
9	N, E, T_{max}, RH, Dist, Cover, Height, I_u, I_c	1.2	7474.9	37.9	0.934	0.301
10	N, E, T_{max}, RH, CT, Dist, Cover, Height, I_u, I_c	2.0	7486.2	49.2	0.935	0.302

Some covariates have a negative impact on MPB presence. A larger distance to the infested border or a higher maximum temperature in summer decreases the probability of infestation in a cell (Fig. 4.5). Other covariates have a positive impact on MPB presence. A larger nearby controlled or uncontrolled infestation, higher pine cover, larger pine height, northerness, or easterness increases the probability of infestation in a cell (Fig. 4.5). The relative humidity has both a positive impact of MPB presence at the onset and a negative impact at the collapse. However, this covariate has a relatively low impact at the onset (Fig. 4.5).

We can visually characterize the spatial patterns of infestations for each outbreak phase. During the onset, there are few large areas with high infestation risk and they are directly adjacent to the park infested border (Fig. 4.6). However, other smaller areas at risk are present in the rest of the park. During the peak, more large areas with high risk of infestation arise and they are located nearby previous infestations rather than adjacent to the park infested border (Fig. 4.6). Note that since the first areas with high infestation risk were close to this border, most areas at risk during the peak are still located in the same general region. During the collapse, the areas with high risk of infestation generally decrease in size but locations similar to the ones during the peak are at risk (Fig. 4.6).

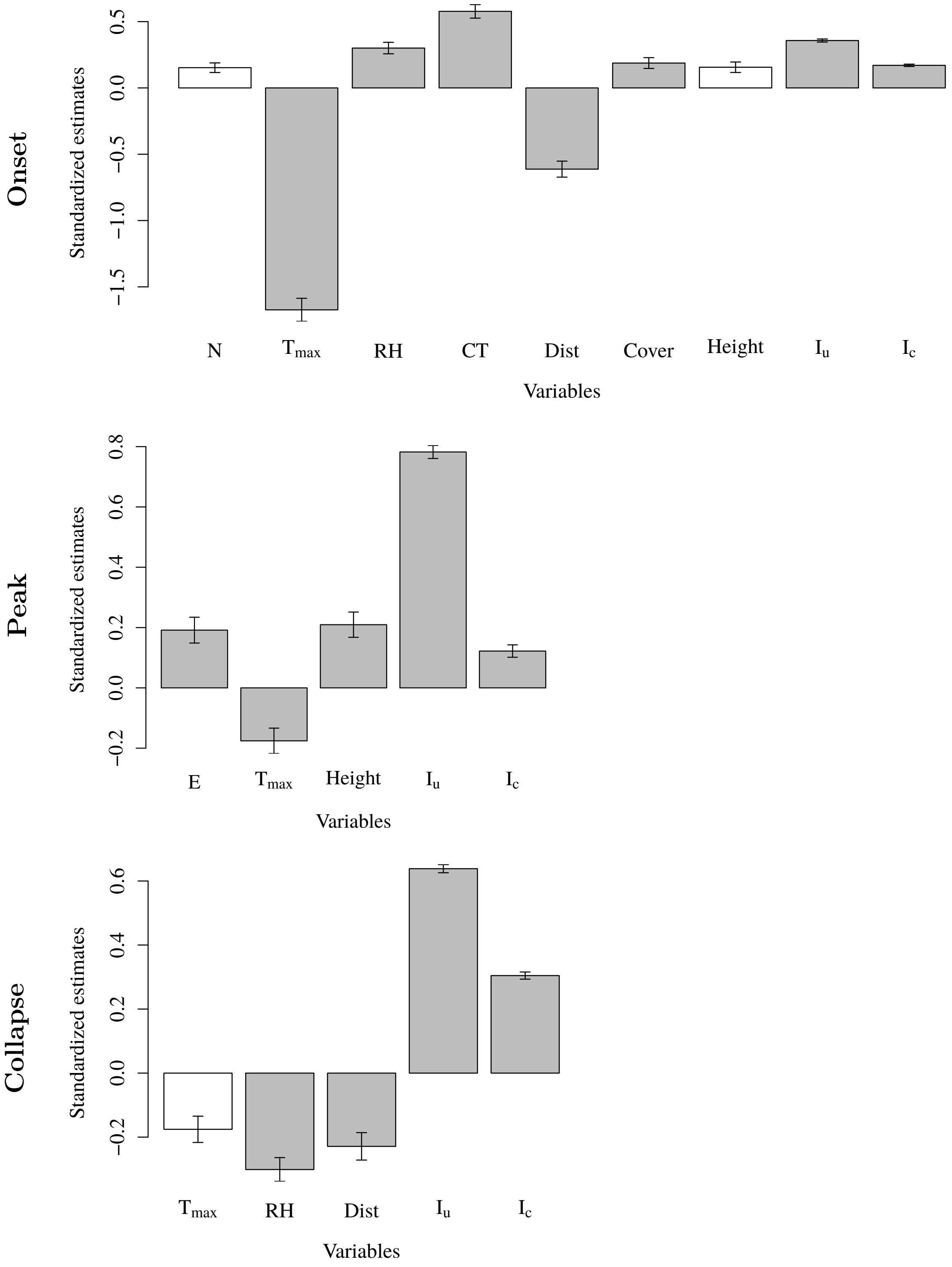

Figure 4.5 – Standardized estimates (± standard error) for each selected model by outbreak phases. Variables in white have weak evidence from ΔBIC (see Tables 4.2 to 4.4).

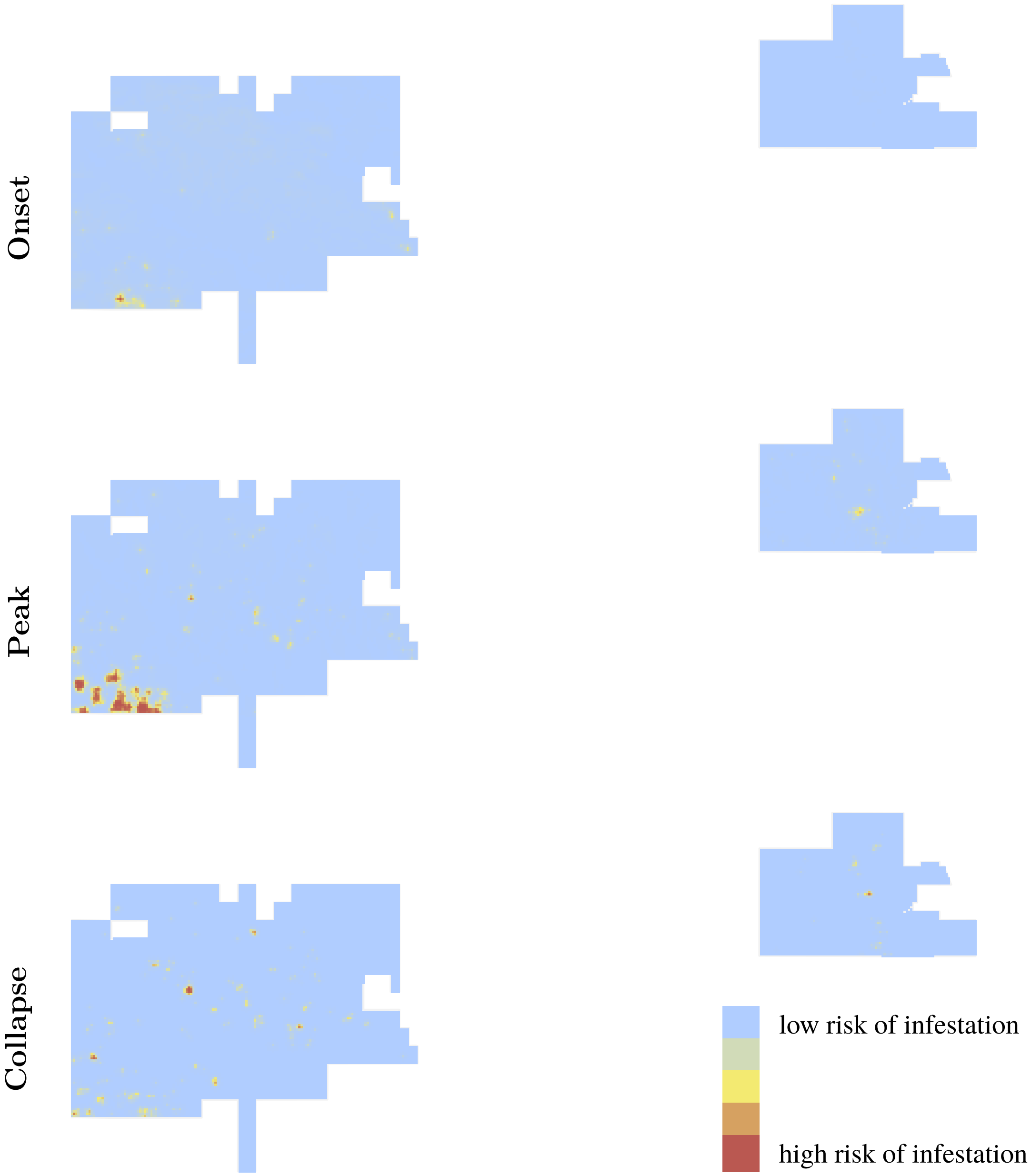

Figure 4.6 – Maps of the predicted infestation probabilities $\pi(\underline{\beta})$ using the parameters from the best model for each outbreak phase. The onset is represented by the year 2009, the peak by the year 2013, and the collapse by the year 2016. For each outbreak phase, the prediction patterns are similar among years. The risk of infestation ranges from low (blue) to high (red).

4.4 Discussion

Our analyses showed how the impact of environmental variables on MPB infestation change with the outbreak phase in a forest managed for MPB. The selected models showed high AUROC values and relatively high AUPR values compared to the null models suggesting, respectively, good predictive abilities overall and good predictions of MPB presence while avoiding false alerts. Therefore most variables driving MPB infestations are likely included in the models.

Parameters related to dispersal (distance to external sources, ability to emerge and fly restrained by high temperatures, source of beetle from the neighbouring uncontrolled trees) were shown to always have a major impact on MPB infestation, as expected. This is especially true at the outbreak onset. The source of beetles shifts from the infested border during the onset to nearby uncontrolled infested trees as beetles establish themselves in patches at the peak and collapse phases.

Parameters related to the host quality (pine height, pine defense ability restrained by low relative humidity) have a non-negligible impact on MPB infestation. During the peak, beetles show a preference for taller trees whereas at the collapse, when the total MPB population size decreases, beetles show a preference for trees with weaker defenses.

Parameters related to the habitat quality (direction of the slope, range of winter temperatures allowing beetle survival, host availability, infestation history) also have a non-negligible impact on MPB infestation. At the onset, beetles favoured areas with advantageous ranges of winter temperatures. During the peak, beetles were successful in locations oriented to the East. During the collapse, beetles favoured trees with a history of MPB infestation in the neighbourhood.

It is generally known that the transition from endemic to epidemic population size is usually due to 1) the relaxation of biotic and abiotic factors impacting beetle mortality such as weaker trees and warmer winter temperatures and/or 2) beetle flights from outside sources (Safranyik, 2004; Carroll *et al.*, 2006; Aukema *et al.*, 2008; Nelson *et al.*, 2018). In this study, MPB infestation is driven mostly by the beetles' dispersal ability during the onset. Here, the impact of the maximum summer temperature seems to reflect the negative effect of high temperatures on beetle emergence and flight (McCambridge, 1971) and not the positive effect of beetle development and spatial synchrony (Aukema *et al.*, 2008). Beetle development rate by itself could not be tested in this study because of the high correlation between the number of degree-days over the year and the maximum summer temperature. There is also evidence for higher winter survival due to cold tolerance but no evidence of weaker trees. So, in Cypress Hills, during the onset, population increase is mostly due to incoming flights from outside the park and warmer winter temperatures.

During the outbreak peak, our findings show that the main factors governing MPB presence are a high level of beetle pressure and history along with low maximum summer temperatures. The importance of beetle pressure and history has also been widely found in previous studies (Safranyik, 2004; Aukema *et al.*, 2008; Preisler *et al.*, 2012). Furthermore, MPB attacks are successful in areas with taller trees. Given the positive relationship between tree diameter and height, this result agrees with the fact that sufficient large and vigorous pines are necessary to sustain high epidemic population size (Safranyik, 2004). No evidence of low relative humidity in spring–indicating drought–was found during the peak which agrees with Safranyik (2004) but disagree with Creeden *et al.* (2014). This factor shows the importance of drought which, by impacting trees and their ability to produce toxic resin, facilitate MPB attacks success. In summary, at the outbreak peak, MPB infestation depends mainly on the availability of nearby beetle sources and, to a smaller extent, favourable weather conditions and habitat quality.

At the collapse, MPB infestation is negatively impacted by relative humidity in spring which could mean that individuals start targeting less vigorous trees. During this phase, pines encountered lower relative humidity than in the two previous phases with an overall decrease in pine cover. This loss of large and healthy trees is usually what is thought to drive an outbreak collapse back to endemic levels in addition to cold snaps in fall and early spring (Safranyik, 2004; Creeden *et al.*, 2014; Nelson *et al.*, 2018). In this study, there is no evidence of cold tolerance impacting MPB presence during the collapse. Therefore, the MPB outbreak collapse in Cypress Hills is linked to the decrease in susceptible pines.

As expected, the presence of uncontrolled trees in a neighbourhood has a positive impact on MPB presence by providing a source of beetles for the following year. Perhaps surprisingly, this positive impact on MPB presence was also seen for entirely controlled cells, albeit at a lower level. We interpret this as being correlative but not causal, and arising from the MPB showing a preference for certain environmental conditions. These conditions persist from year-to-year even after the MPB-infested trees are controlled. Good management reduces effectively the likeliness of infestation, since controlled infestations always have a lower impact than the presence of uncontrolled trees.

A limitation of this work comes from the fact that we are working with presence/absence in cells and not actual numbers of infested trees or beetles. Carroll *et al.* (2006) mentions that the number of infested trees is a good proxy for the beetle population size. Our use of presence/absence instead of number of infested trees does, however, allow us to deal with the issue that a small part of the data is expressed as infested zones and not actual tree locations.

Some ecological factors influencing MPB infestations, such as predators and competitors, were not available and therefore were not included in this analysis. These factors are not as readily and broadly available as weather variables, and thus are often not included in analyses. Other factors linked to host and stand characteristics were only available for a couple of years within our study period. Therefore, they were largely estimated. However, future work should focus on gathering such data and analyzing the impact of MPB predators and competitors, along with stand characteristics, on MPB location (see also Safranyik, 1999; Smith *et al.*, 2011; Krause *et al.*, 2018).

Finally, when determining the outbreak phases, we considered the outbreak status in

Cypress Hills as a whole instead of differentiating the status of each cell. For example, some cells could be newly infested during the outbreak peak or collapse. However, since the study area is small, it makes sense to see the outbreak as a whole as factors usually have larger yearly variations than within-year variations.

To conclude, the impact of weather, vegetation, and beetle or host-related factors on MPB infestations were shown to vary in a clear, ecologically interpretable manner during an outbreak. This gives managers guidance regarding which stands to focus on for an efficient control. For example, they could use the risk probability maps to inform survey locations (Kunegel-Lion *et al.*, 2019). These results also point out that the predictive ability of models using data from an incomplete outbreak to determine future infestations may be limited. Indeed, with such a change in the factor impacts from an outbreak phase to another, the predictions for a specific phase should be biased if model training is done with data from another phase. However, while the size of impact does change, the direction of impact of any given covariate seldom changes as a function of the outbreak phase so this may limit prediction error.

Chapter 5

Mountain pine beetle outbreak duration and pine mortality depend on control effort

5.1 Introduction

Bark beetle outbreaks are a major threat for western North American pine forests, impacting tens of thousands of square kilometers (Romme *et al.*, 1986; Fettig & Hilszczański, 2015). Beetle epidemic populations caused a decline of pine forests and changes in forest structure and composition, eventually leading to significant economic losses (Diskin *et al.*, 2011; Walton, 2013). Detection of infested trees is typically via aerial surveys, ground-based surveys, or a combination of the two (Fettig & Hilszczański, 2015). Managers employ various control tactics to reduce levels of tree mortality and these can be divided divided into indirect–or preventive–and direct methods. The potential of direct control methods to be effective depends on the ability to detect bark beetle infestations. In turn, the level of direct control impacts outbreak duration and the extent of tree mortality. In this paper, we focus on direct control.

Several bark beetle species are capable of causing significant levels of tree mortality.

Among them are the European spruce beetle, *Ips typographus*, infesting various species of spruce in central Europe, the spruce beetle, *Dendroctonus rufipennis*, affecting forests of spruce trees in North America, the southern pine beetle, *Dendroctonus frontalis*, particularly active in pine forests in Central America and southeastern North America, and the mountain pine beetle, *Dendroctonus ponderosae*, which is causing significant losses in western North America. These species are responsible for extensive ecological and economic damages triggering intensive control in many geographical locations.

To obtain the necessary information on beetle locations and implement control methods, managers carry out aerial and ground-based surveys. Aerial surveys often are efficient for detecting previously or, in rare cases, currently infested trees (Fettig & Hilszczański, 2015). On the other hand, ground -based surveys are the primary methods used to identify currently infested trees. Using aerial surveys of previously infested trees to inform ground-based surveys allows managers to direct detection resources to susceptible locations by using the propensity of beetles to engage in short-distance dispersal from the previously infested trees to attack new trees (Safranyik & Carroll, 2006).

Once an outbreak is detected, preventive measures can be implemented in high-risk locations and direct control can be implemented by removing live broods from the environment. Preventive methods focus on making the environment less susceptible to bark beetles. For example, thinning reduces the number of susceptible trees but also induces changes in micro-climate which impact the beetle's life cycle (Fettig & Hilszczański, 2015). By way of contrast, direct control aims to actively reduce beetle population and slow beetle spread, for example, by felling and burning infested trees. Several studies have compared different direct control methods (*e.g.* Nelson *et al.*, 2006; Seidl *et al.*, 2008; Strom & Clarke, 2011). The methods can have different control efficiency due either to a different detection rate or to a different removal success. With the assumption that the number of infested trees is a good proxy for beetle population size (Safranyik, 1988; Carroll *et al.*, 2006), using the number of infested trees removed from the environment allows for a fair comparison of methods.

In this paper, we focus on mountain pine beetle outbreaks in Cypress Hills, Saskatchewan.

Mountain pine beetles have a one-year life cycle (Safranyik & Carroll, 2006). During the summer, adults emerge and attack new pines by drilling galleries under the bark. There, they mate and females lay eggs. The new generation overwinters as larvae before resuming their development to adult stage in the summer. Adults usually do not survive the winter. During outbreaks, mountain pine beetle individuals typically present a mass attack behaviour where they coordinate their effort using pheromones to attack large and healthy pines in sufficient number to overwhelm their defences (Safranyik & Carroll, 2006). The mountain pine beetle's main host in western North America is the lodgepole pine (*Pinus contorta*) although it is a threat for almost every pine species (Safranyik & Carroll, 2006).

The efficacy of direct control of mountain pine beetle outbreaks is disputed: while some studies report significant reduction in the levels of tree mortality due to direct control, others notice little to no effect compared to uncontrolled areas (Six *et al.*, 2014). For example, Wulder *et al.* (2009) reported that control activities slowed the rate of increase of infested trees compared to no control. However, Trzcinski & Reid (2008) reported that beetle populations continued to increase at the same rate in treated and untreated areas and that between 45% and 79% of infested trees were left undetected in the treated area. It is generally accepted that extremely high detection and control rates are required to eradicate mountain pine beetle outbreaks (Six *et al.*, 2014).

In order to obtain quantitative information on the shape of the relationship between control rates and mountain pine beetle outbreaks, Carroll *et al.* (2006) built a simple and aspatial theoretical framework describing the proportion of infested trees that must be treated to maintain a constant beetle population size P as a function of the yearly rate of increase of infested trees R:

$$P = 1 - 1/R. \tag{5.1}$$

Combined with the following geometric growth population model of beetle population N at time t

$$N_t = N_0\big(R(1 - P_a)\big)^t \tag{5.2}$$

where P_a is the actual proportion of infested trees treated, they were able to show, theoretically, how many years of sustained control effort is necessary to achieve outbreak suppression depending on the rate of increase and the control effort. Indeed, when the actual proportion of infested trees treated P_a reaches the proportion of infested trees that must be treated P, Eq. 5.2 becomes $N_t = N_0$. While this gives useful insights, it does not fully account for spatial and temporal heterogeneities found in natural environments. In turn, these translate into heterogeneities for infestations. For example, variations in temperatures, host densities, or host vigour impact beetle life cycles (Safranyik & Carroll, 2006) and thus, the rate of increase of infested trees. In addition, natural outbreaks are typically not isolated in space, making them susceptible to rebound even under control. Thus, there is the need for evaluation of control in spatially and temporally variable environments by coupling theoretical models to detailed data regarding environmental conditions and infestation levels.

In this study, we analyze quantitatively the impact of control effort on bark beetle and pine populations under the influence of a spatially and temporally changing environment with beetle intake from outside sources. The objectives are: 1) to build a semi-empirical model of infestation, 2) to simulate outbreaks under different control rates, and 3) to explore the impact of control effort on beetle and pine population sizes. To do so, we use exhaustive mountain pine beetle infestation data from the Cypress Hills area and simulate outbreaks using observed environmental variable values.

5.2 Material and methods

5.2.1 Study organism, area, and data

The data used to calculate the pine population growth rate, parametrize the predictive model, and initiate the simulation model comes from the mountain pine beetle infestation in the Saskatchewan portion of Cypress Hills interprovincial park. This park is located at the border between the provinces of Saskatchewan and Alberta. The Saskatchewan portion of the park is divided in two sections that are 20-km apart. Their combined extent covers 184 km^2.

The Saskatchewan Forest Service is responsible for managing mountain pine beetles in the Saskatchewan portion of the park. To do so, they survey the park every year to locate infested trees and control them (Saskatchewan Ministry of Environment, 2016). First, a complete aerial survey of the park extent is performed in order to detect red-top trees. Red-top trees are the dying trees infested by mountain pine beetle in the previous year and they typically do not carry live broods anymore. Second, circular surveys with a radius of 50 meters are conducted around each red-top tree to find trees currently infested by live brood. Finally, the infested trees that are detected are then controlled, principally using a fall and burn tactic to ensure that beetles are killed. However, in some areas, red-top trees are clustered together. To thoroughly survey these areas, polygons are delineated around each red-top trees cluster. The extent of each polygon is then entirely checked for infested trees using line surveys. All detected infested trees are controlled. Polygon locations and shapes typically change from one year to the other. However, they are consistently located in the same highly infested and therefore highly surveyed areas. The Forest Service has been following this procedure since the mountain pine beetle infestation was detected in 2006 up to the current collapse in 2018. In summary, we obtained the following data for each year: locations of red-top trees, locations and shapes of polygons, number of infested trees for each circular survey, locations of infested trees within the polygons.

To get an estimate of all of the infested trees location for every year, we used the following method. For the controlled infested trees found in circular survey plot (called type C1), we used the location of the plot's centroid. For the controlled infested trees found in line surveys (*i.e.* within polygons, called type C2), we used the exact location of the infested trees. For the uncontrolled infested trees that were missed at year t, became red-top trees at year $t+1$ and were not part of a polygon at year $t+1$ (called type U1), we used the exact location of the red-top trees at year $t+1$. For the uncontrolled infested trees that were missed at year t, became red-top trees at year $t+1$ and were part of a polygon at year $t+1$ (called type U2), we first assumed that the proportion of uncontrolled infested trees within these highly surveyed areas is 0.11 (Kunegel-Lion *et al.*, 2019). Therefore, the proportion of controlled infested trees is $1 - 0.11 = 0.89$. Using these proportions, the number of controlled trees is equal to 0.89 × the number of infested trees. So, the number of infested trees is the number of controlled trees / 0.89. Therefore, the number of uncontrolled infested trees is

$$\begin{aligned}\#\text{ uncontrolled infested trees} &= 0.11 \times \#\text{ infested trees} \\ &= 0.11 \times \frac{\#\text{ controlled infested trees}}{0.89} \end{aligned} \tag{5.3}$$

Second, using Eq. 5.3, we modelled, for each $t+1$ polygon, the number of infested trees that would have been missed and fell within the polygon areas at year $t+1$ as a Poisson random variable with mean equal to the number of uncontrolled infested trees at t that fall into a $t+1$ polygon. Third, we randomly distribute in space the U2 infested trees in each polygon area. In total, we had 2672 trees of type C1, 740 trees of type C2, 1819 trees of type U1, and 93 trees of type U2 (Fig. 5.1). We were able to estimate the location of all infested trees every year with a precision of a few meters for 48% of the infested trees (types C2 and U1), 50 meters for 50% of the infested trees (type C1), and several hundred meters depending on the polygons size for the remaining 2% (type U2).

The ecological and environmental covariates and the infestation response value were discretely distributed in space and time. We superimposed a grid of 722 cells of size 500×500 meters over the park extent, counted how many infested trees fell in each cell, and obtained ecological and environmental variable values for each cell over the study period from 2005 to 2018. To reflect the impact of high temperatures on mountain pine beetle emergence and dispersal (Safranyik & Carroll, 2006), we included the maximum temperature over the spring and summer. Because mountain pine beetle larvae are very sensitive to sudden changes in minimum temperatures in the fall, winter, and spring (Safranyik, 2004), we included the cold tolerance metric from Régnière & Bentz (2007). A small fraction of mountain pine beetles engage in long-distance dispersal events and

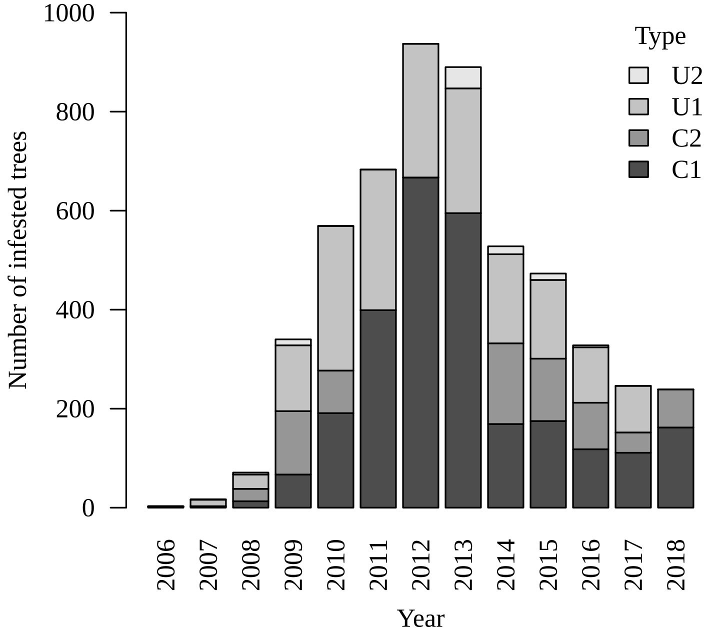

Figure 5.1 – Number of infested trees per type over the years in Cypress Hills, SK. Type C1 represents the controlled infested trees found in circular surveys whereas type C2 represents the controlled infested trees found in line surveys. Types U1 and U2 represent the uncontrolled infested trees at year t that become red-top trees, respectively, outside or inside of a year-$t+1$ polygon. For the year 2018, the numbers of U1 and U2 infested trees are unknown at the time of the study as they can only be estimated with 2019 data.

disperse further than a few hundred meters by getting caught in the wind (Safranyik & Carroll, 2006). Therefore, we included the wind speed during spring and summer which corresponds to the dispersal season. Water stress reduces the pines' ability to resist mountain pine beetle attacks (Safranyik, 1978; Creeden *et al.*, 2014). Therefore, we included the relative humidity during the spring. Weather variables were estimated for each grid cell and each year using the BioSIM software (Régnière *et al.*, 2014). In addition to weather variables, we included elevation, slope, northerness and easterness to address bias in mountain pine beetle dispersal due to spatial configuration and wind direction. Topography variables were calculated from the Canadian Digital Elevation Map downloaded from the Geogratis website (geogratis.cgdi.gc.ca). We also included in our list of covariates pine height and number of pines per cell. From Beaudoin *et al.* (2014), we estimated, for each cell in 2001 and 2011, the leading species height, the pine cover and the tree volume. Then, we obtained pine height by using the leading species height when the pine cover is greater than 50%. We spatially interpolated the pine height values at the location of each cell using bicubic spline interpolation provided by the function `interp` of the `R` package `akima` (Akima & Gebhardt, 2016). We obtained the pine volume by multiplying the pine cover by the tree volume. To obtain vegetation variable values for every year, we linearly interpolated the values over the time period for each cell. We estimated the number of pine trees in each cell from the pine volume using the process and equation described in Goodsman *et al.* (2016). The expected number of pines greater than 10 cm at breast height $E(S)$ depends on the pine volume per hectare V following the equation

$$E(S) = AV \exp(-\delta V), \tag{5.4}$$

where A and δ are free parameters. Pines with a diameter at breast height smaller than 10 cm are seldom the target of mountain pine beetle attacks and therefore were not included in the pine count (Safranyik & Carroll, 2006). Based on Monserud *et al.* (2006), Cypress Hills has a site index of 15 to 18 meters at 50 year breast height age for lodgepole pine which corresponds to a medium to good site in Alberta Sustainable

Resource Development (1985). Therefore, we fit A and δ to data from the yield tables for medium and good sites provided in Alberta Sustainable Resource Development (1985) using a nonlinear regression and obtained the values $A = 18.18$ (± 0.23 SE) and $\delta = 5.4 \times 10^{-3}$ ($\pm\, 0.4 \times 10^{-3}$ SE).

5.2.2 Model

The simulation model used a baseline pine population growth rate γ and predictions of the number of infested trees I_x^{t+1} depending on ecological and environmental variables $\underline{X}_x^t$ and number of pines H_x^t at year t for the same location. The model symbols are defined in Table 5.1.

Pine population growth rate

The pine population growth rate in the absence of mountain pine beetles provides a baseline to which we can compare the observed pine mortality. Using the number of pines H_x^t in 2001 and in 2011 only for cells where no beetles were present between 2001 and 2011, we calculated the pine yearly growth rate using the equation:

$$\gamma \;=\; \left(\frac{\sum\limits_x H_x^{2011}}{\sum\limits_x H_x^{2001}} \right)^{1/(2011-2001)} \tag{5.5}$$

In addition, we estimated variability in the pine population growth rate by calculating the standard deviation of the yearly pine population growth rate per cell.

Predicting the number of infested pines

In this section, we describe our a semi-empirical infestation model used to predict the number of infested trees from ecological and environmental variables.

An infested tree at year t produces new infested trees at year $t + 1$ according to $\text{Poisson}\big(R_x^t\big)$ where R_x^t is the geometric rate of increase of infested trees at year t and location x. We assume that the rate of increase of infested trees in a cell depends on the number of pines available H_x^t as well as other ecological and environmental variables

Table 5.1 – Description of the symbols used in the model.

Symbol	Description
R_x^t	Geometric rate of increase of infested trees from year t to $t+1$
$\underline{X}_x^t$	Vector of environmental variable measurements at year t
H_x^t	Number of pine trees in a cell x at year t
$\underline{\beta}$	Vector of parameters associated with variables $\underline{X}_x^t$ and H_x^t
M_x^t	Number of uncontrolled infested trees at year t and location x
$N(x)$	Von Neumann neighbourhood of location x
w	Weight associated with the number of uncontrolled infested trees in each cell of the neighbourhood $N(x)$
r_0	Number of infested trees from other sources
I_x^{t+1}	Number of infested trees at year $t+1$ and location x
B_x^t	Number of uncontrolled infested trees added to M_x^t to describe boundary conditions at year t and location x

$\underline{X}_x^t$ described in the previous section. We express the rate of increase of infested trees R_x^t as

$$R_x^t \;=\; e^{\beta_0 \,+\, \beta_1 H_x^t \,+\, \sum\limits_{i\geq 2} \beta_i X_{x,i}^t} \tag{5.6}$$

At a location x, the sources of new infestations at $t+1$ are:

1. M_x^t: number of uncontrolled infested trees at t in the same location x,
2. $\sum\limits_{\xi \in N(x)} M_\xi^t$: number of uncontrolled infested trees at t in the Von Neumann neighbourhood of x,
3. B_x^t: number of uncontrolled infested trees at t in the cells outside the park limits and adjacent to x, this variable allows us to include boundary conditions in the model,
4. r_0: number of infested trees arising from other sources, such as long-distance dispersal events originating from inside or outside the park limits and transitions from endemic to epidemic population level, and representing a background infestation level.

The number of infested trees I_x^{t+1} at year $t+1$ and location x is a sum of R_x^t for each source of infestation, and thus, a Poisson-distributed random variable itself. Therefore, we can write

$$I_x^{t+1} \;\sim\; \text{Poisson}\Bigg(\left(\frac{M_x^t + w\big(B_x^t + \sum_{\xi \in N(x)} M_\xi^t\big)}{1+4w} + r_0 \right) R_x^t \Bigg) \tag{5.7}$$

where w is the weight associated with each location adjacent to x. The term $1+4w$ arises due to the weight of 1 attributed to the number of uncontrolled infested trees at the same location and the weight of w attributed to the number of uncontrolled infested trees in each of the 4 cells of the Von Neumann neighbourhood.

Eq. 5.7 can be rewritten as

$$I_x^{t+1} \;\sim\; \text{Poisson}\Bigg(\Big(M_x^t + w\big(B_x^t + \sum_{\xi \in N(x)} M_\xi^t\big) + \widetilde{r_0} \Big) \, \widetilde{R}_x^t \Bigg). \tag{5.8}$$

where

$$\widetilde{r_0} \;=\; r_0(1+4w) \tag{5.9}$$

and

$$\widetilde{R}_x^t \; = \; \frac{R_x^t}{1+4w}. \tag{5.10}$$

Therefore, the expected value of I_x^{t+1} is

$$\begin{aligned} E\big(I_x^{t+1}\big) &= \big(M_x^t + w\big(B_x^t + \sum_{\xi \in N(x)} M_\xi^t\big) + \widetilde{r_0}\big)\; \widetilde{R}_x^t \\ &= e^{\ln\left(M_x^t + w\left(B_x^t + \sum\limits_{\xi \in N(x)} M_\xi^t\right) + \widetilde{r_0}\right)} \; e^{\beta_0' + \beta_1 H_x^t + \sum\limits_{i \geq 2} \beta_i X_{x,i}^t} \\ &= e^{\ln\left(M_x^t + w\left(B_x^t + \sum\limits_{\xi \in N(x)} M_\xi^t\right) + \widetilde{r_0}\right) \; + \; \beta_0' \; + \; \beta_1 H_x^t \; + \; \sum\limits_{i \geq 2} \beta_i X_{x,i}^t}, \end{aligned} \tag{5.11}$$

where $\beta_0' = \beta_0 - \ln(1+4w)$.

The boundary conditions were such that cells outside the park limits have a number of infested trees equal to 0 except for cells just outside the park southern border which has a number of infested trees greater or equal to 0. This allowed us to take into account the fact that there were no mountain pine beetle flights from outside the park limits except outside the park southern border (Rory McIntosh, *pers. comm.*). We made the reasonable assumption that the infestation outside the park limits lasted from 2006 to 2014 (estimated from Brian Poniatowski, *pers. comm.*). We optimized the boundary conditions outside the infested border using the function `optim` of the `R` package `stats`. For each cell x adjacent to the infested border, we added to $\sum\limits_{\xi \in N(x)} M_\xi^t$ a value b_x^t estimated with the optimization algorithm. Thus,

$$B_x^t = \begin{cases} b_x^t & \text{if } x \text{ is adjacent to the infested border,} \\ 0 & \text{otherwise.} \end{cases} \tag{5.12}$$

The optimization algorithm used the Nelder-Mead method with initial parameter values $b_x^{2005 \text{ to } 2010} = \{0, 0, ..., 0\}$. The function maximized the coefficient of determination r^2 between the observed and predicted number of infested trees once Eq. 5.11 was fitted to data.

To fit Eq. 5.11 to data, we divided the data set in folds by holding out a different year for each fold. We only kept the folds where the year holdout was between 2009 and 2018 as the years 2006 to 2008 had a very small number of cells with $I_x^t > 0$. Therefore, we had 10 folds. With the function `gnm` of the `R` package `gnm`, we fit, on the training set of each fold, a Poisson nonlinear model using Eq. 5.11. We evaluated each of the 10 fits on the test set of the corresponding fold using the coefficient of determination r^2 between observed and predicted number of infestation. Then, we fit the model on the entire data set and referred to it as the final model.

Simulations

We simulated, between 2006 and 2018, outbreaks subject to various control rates. We used a time unit of 1 year and a spatial unit of a 500×500 meters cell. Table 5.2 shows the state variables used in the simulations. At each time step and for each location, the number of infested trees is the sum of the number of controlled and uncontrolled trees:

$$I_x^t = M_x^t + C_x^t. \tag{5.13}$$

For each control effort e, we ran 1000 simulations. The initial values corresponded to the park observations in 2006. The total number of infested trees in 2006 was $\sum_x I_x^{2006} = 3$. Starting with $t = 2006$, we repeated the following process over the years for each location:

1. We drew a baseline pine population growth rate g_x^t in a normal distribution with mean γ and the standard deviation calculated in the previous section.
2. We calculated the number of pines susceptible to beetle attacks using $S_x^{t+1} = g_x^t H_x^t$.
3. We predicted the number of infested trees I_x^{t+1} using the model described in Eq. 5.11 and parameters fitted in the previous section.
4. We drew the number of controlled infested trees C_x^{t+1} in Binomial(e, I_x^{t+1}).
5. We calculated the number of uncontrolled infested trees using $M_x^{t+1} = I_x^{t+1} - C_x^{t+1}$.
6. We calculated the number of pines surviving the year using

Table 5.2 – State variables used in the simulation model.

Variable	Description
S_x^t	Number of pines susceptible to beetle attacks at year t and location x
H_x^t	Number of pines surviving after beetle attacks at year t and location x
I_x^t	Number of infested pines at year t and location x
M_x^t	Number of uncontrolled infested pines at year t and location x
C_x^t	Number of controlled infested pines at year t and location x

$$S_x^{t+1} \;=\; H_x^{t+1} \;-\; I_x^{t+1}.$$

To compare the simulations to the observations in Cypress Hills, we performed the simulations with the observed control effort e^t_{obs} each year. The observed control effort e^t_{obs} was calculated using the observed values of C_x^t and I_x^t following the equation:

$$e^t_{\text{obs}} = \frac{\sum\limits_x C_x^t}{\sum\limits_x I_x^t}. \tag{5.14}$$

5.3 Results

5.3.1 Pine population growth rate

In the absence of mountain pine beetle, the pine geometric growth rate estimated using Eq. 5.5 was 0.95, indicating a diminishing population. The standard deviation of the pine population growth rate per cell was 0.17.

5.3.2 Predicting the number of infested trees

To estimate the values of b_x^t, we used the optimization algorithm described in the previous section. We obtained the maximum coefficient of determination $r^2 = 0.72$ for the values described in Table 5.3.

To evaluate the predictive model of the number of infested trees, we compared the observed and predicted values on the 10 folds of the data set using the coefficient of determination r^2. The average r^2 was 0.63 with a minimum of 0.21 and a maximum of 0.86. Visually, the predicted infested areas correspond to the observations (see Fig. 5.2 for a typical example). We obtained the parameter values $\widetilde{r_0} = 0.045$ (± 0.002 SE) and $w = 0.089$ (± 0.004 SE). Therefore, using, Eq. 5.9, r_0 is equal to 0.033. Using Eq. 5.10, we obtain the values of R_x^t from $\widetilde{R}_x^t$. Forested areas have a mean R_x^t equal to 2.1 (see Fig. 5.3 for a typical example). The areas with no pine cover as estimated from aerial imagery indicate R_x^t values greater than 0 to be able to handle the possibility of isolated

Table 5.3 – Parameter values estimated for b_x^t.

Year	Cell									
	1	2	3	4	5	6	7	8	9	10
2006	1	4	2	0	0	0	0	2	8	0
2007	3	0	25	37	41	14	0	0	0	0
2008	45	10	128	65	63	167	72	18	0	11
2009	67	13	80	78	72	135	95	19	7	0
2010	22	26	37	59	80	84	37	0	0	34
2011	11	0	108	34	525	261	128	1	58	0
2012	53	0	88	42	202	191	2	16	0	3
2013	0	0	58	30	164	210	107	11	0	0
2014	32	7	0	0	123	105	25	0	0	0

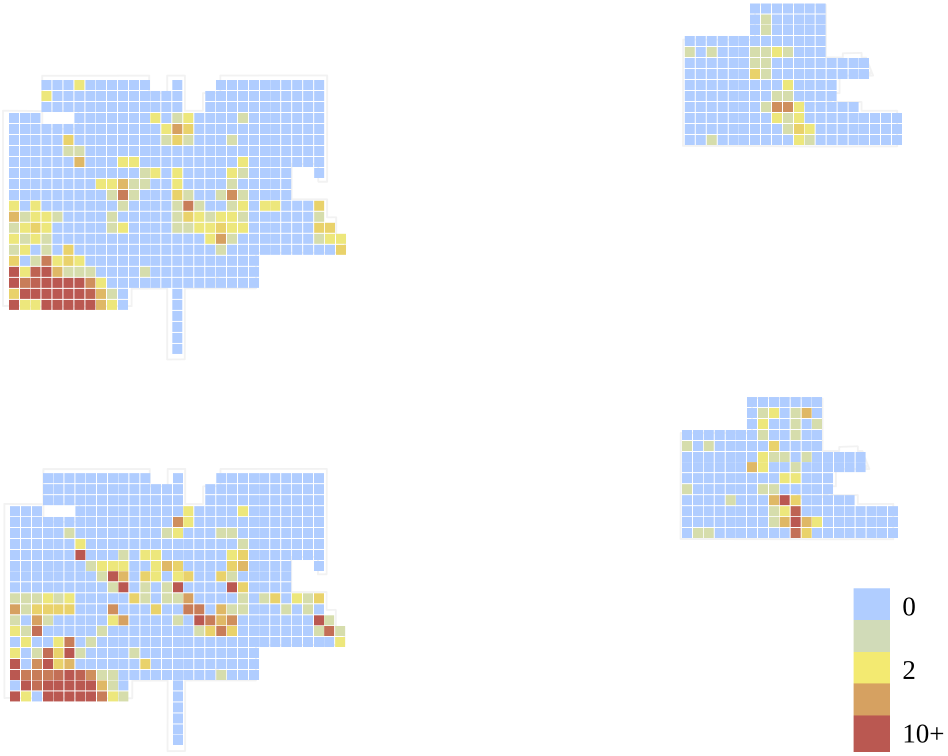

Figure 5.2 – Predicted (top) vs observed (bottom) number of infested trees I_x^{2013} per 500×500-meters cell in 2013 on a log scale.

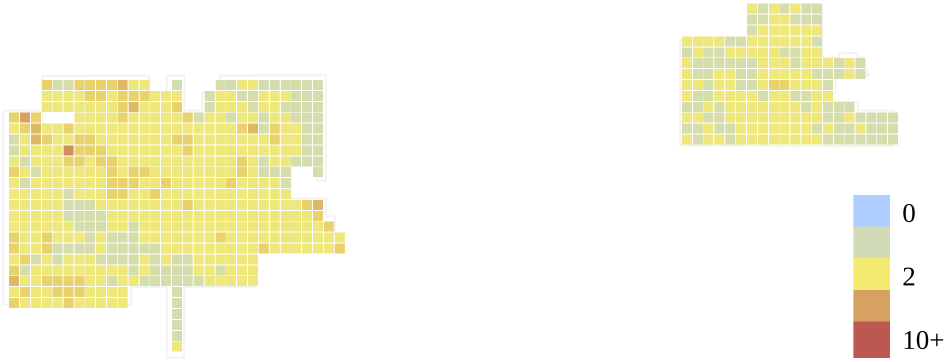

Figure 5.3 – Yearly rate of increase R_x^{2012} per 500×500-meters cell on a log scale estimated using a non-linear regression to obtain the number of infested trees I_x^{2013} for the year 2013.

pines in the simulation model. The sum of the predicted infested trees over the domain each year matches the sum of the observed infested trees (Fig. 5.4).

5.3.3 Simulations

Using Eq. 5.14, we calculated the observed control effort in Cypress Hills and obtained a median of 62% of controlled infested trees over the years (Fig. 5.5). Using these control proportions each year in our simulations, we obtained the total number of infested trees, the number of uncontrolled infested trees and the number of pines that match the observed or estimated values in the park (Fig. 5.6). The model described with good precision the outbreak dynamics although the simulated number of infested trees is lower at the peak than the observations.

As control effort decreases, outbreaks were less likely to be eradicated and the number of infested trees increased exponentially over time (Fig. 5.7). Even a small control effort allowed a substantial decrease in the number of infested trees compared to no control. For example, implementing 40% control reduces the number of infested trees after 10 years under 5% of the number of infested trees under a 1% control (Fig. 5.7 and Table 5.4). Likewise, respectively 50% and 65% control are necessary to reduce the number of infested trees after 10 years to 1% and 0.1%. Control effort above 55-60% actively reduced the total beetle population over time. When under 99% control, the tree population decreased by 46% over 13 years whereas it decreased by 61% when under 1% control (Fig. 5.8). Outbreak suppression refers to the act of driving epidemic populations below the eruptive threshold. Therefore, when one or less mass-attacked tree remains, managers achieve outbreak suppression although a small number of beetles might persist in the domain. The number of years of sustained control effort necessary to achieve suppression of half of the outbreak simulations went from over 13 years for a proportion of controlled infested trees up to 80% to 12 years at 81% control and 9 years as managers approach complete control of the infestations (99%) (Fig. 5.9). Achieving suppression of mountain pine beetle in 95% of the outbreak simulations in less than 13 years required a proportion of controlled infested trees to be at least 89%.

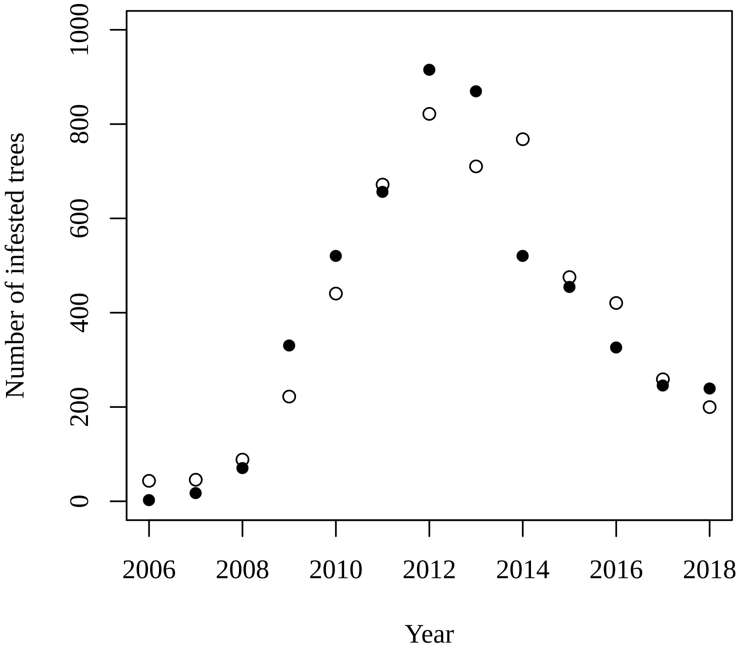

Figure 5.4 – Predicted (white) vs observed (black) number of infested trees per year in the entire domain $\sum_{x} I_x^t$.

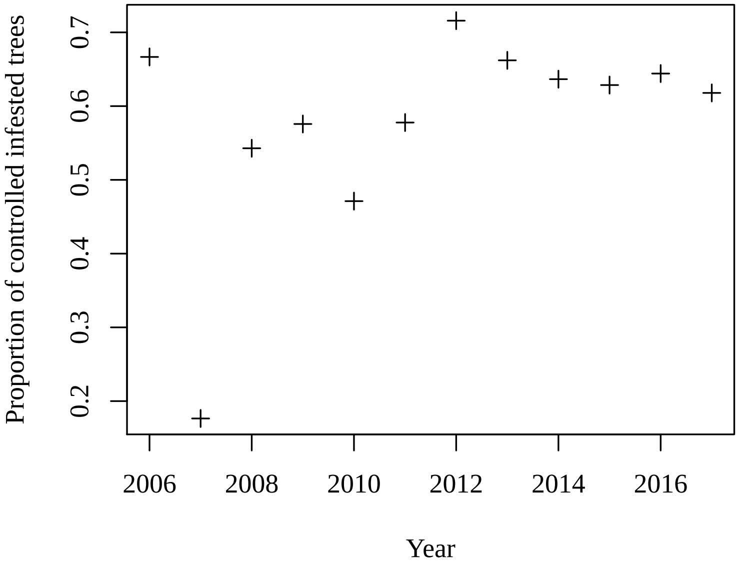

Figure 5.5 – Observed control efficiency in Cypress Hills, Saskatchewan, Canada.

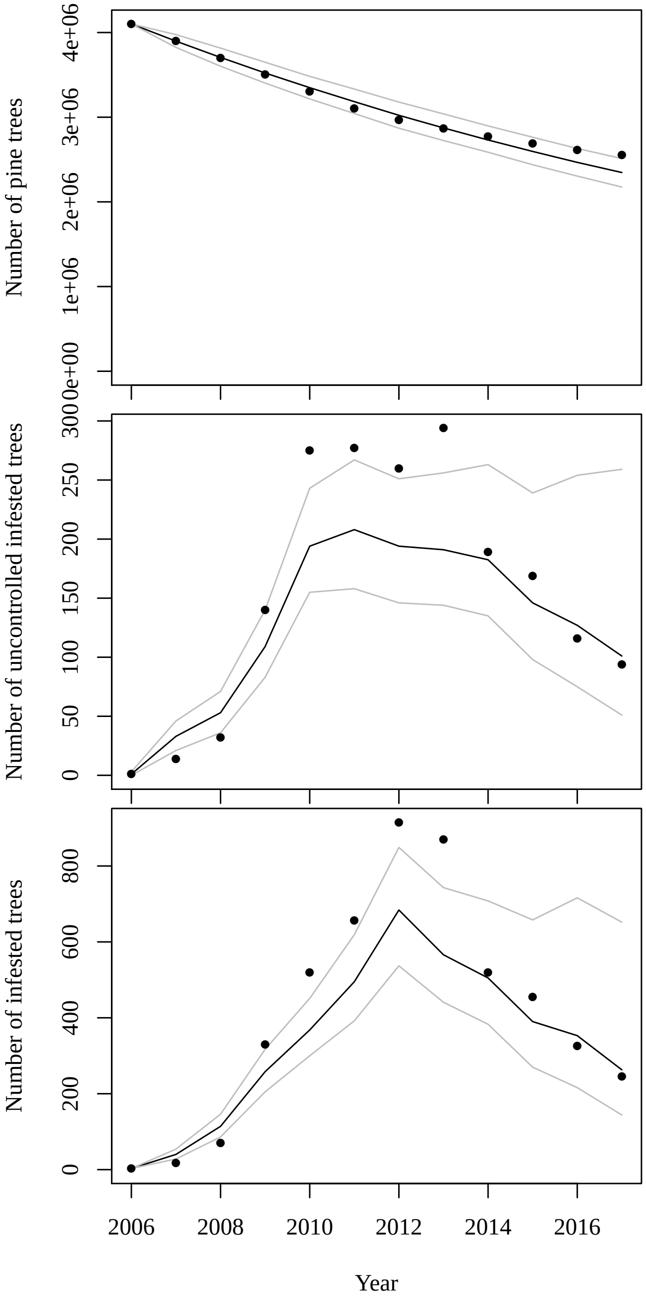

Figure 5.6 – Simulated number of pine trees $\sum_x H_x^t$ (top), number of uncontrolled infested trees, $\sum_x M_x^t$ (middle), and number of infested trees $\sum_x I_x^t$ (bottom) over the years for a control effort matching the observed control. The black and grey lines are, respectively, the median and 95% confidence interval over 1000 simulations. The points represent the observations.

Table 5.4 – Sustained control effort necessary to obtain close to 10, 5, 1, or 0.1% of infested trees after 10 years compared to 1% control.

Control effort	Number of infested trees	Percentage of infested trees
1%	184 492	100.00%
30%	18 521	10.04%
40%	6 482	3.51%
50%	1 959	1.06%
65%	245	0.13%

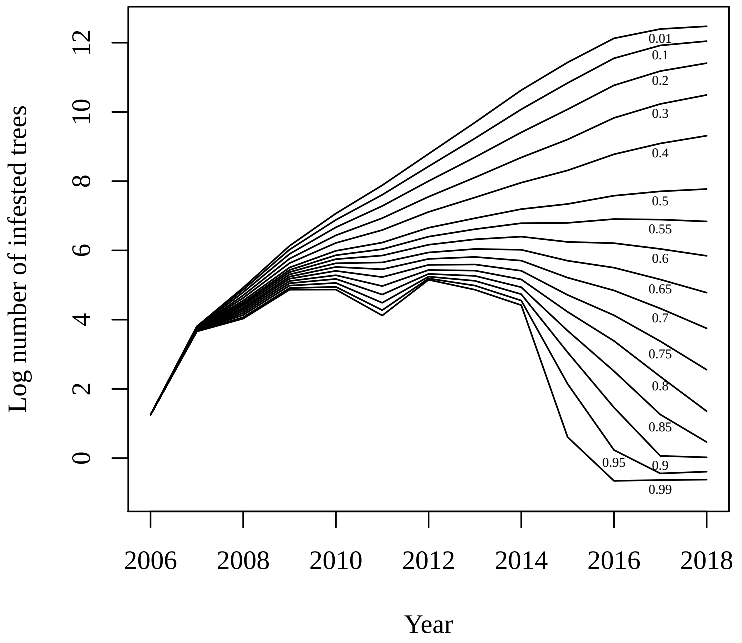

Figure 5.7 – Total number of infested trees in the domain $\sum_x I_x^t$ for control rates between 0.01 and 0.99 on a log scale. Each line represents the average over 1000 simulations for a control rate.

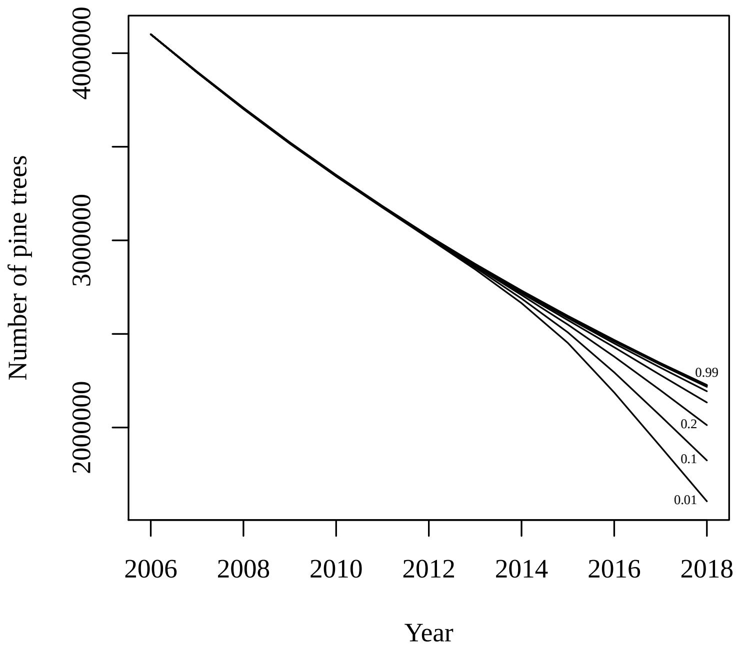

Figure 5.8 – Total number of pine trees in the domain $\sum_{x} H_x^t$ for control rates between 0.01 and 0.99. Each line represents the average over 1000 simulations for a control rate.

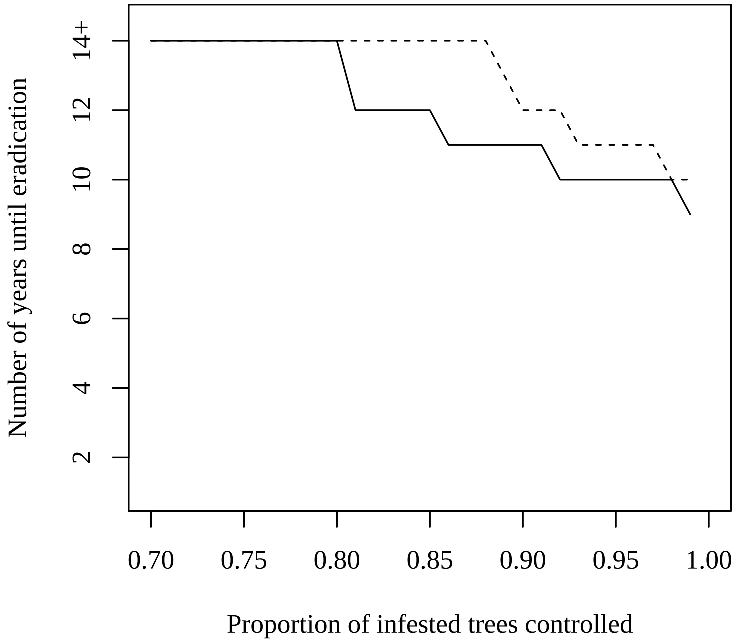

Figure 5.9 – Number of years of sustained control effort until outbreak suppression depending on the control rate. The solid line represent the number of years necessary to achieve outbreak suppression in half of the 1000 simulations and the dashed line in 95% of the simulations.

5.4 Discussion

In this study, we showed that using even a little control on mountain pine beetle infested trees decreased significantly the beetle population compared to a no control scenario after controlling for environmental factors. Indeed, maintaining 40% control over 10 years reduces the beetle population to under 5% of what it would be for 1% control. However, moderate control (55-60%) is required to reduce the beetle population over time despite the outbreak being initially supported by outside flights. In addition, high control (above 89%) can lead to true outbreak suppression under 13 years.

The observed control efficiency in Cypress Hills was relatively constant over time and centred at about 62% of controlled infested trees with the exception of the year 2007. This result agrees with the previous assessment of Cypress Hills control efficiency where the proportion of controlled trees were calculated to be 56% and 71%, respectively in 2011 and 2012 (Kunegel-Lion *et al.*, 2019). Note that, in 2007, control efficiency was surprisingly low. This efficiency could be explained by a surge of beetle flights from the infestations outside the park limits happening that year. Indeed, the current control strategy is not able to detect long-distance dispersal events efficiently since it focuses on the close neighbourhood of previous-year infested trees.

The predictive model accuracy when comparing number of infested trees per cell is average. However, when taking into account the overall spatial distribution of the number of infested trees in the domain, the predictions of the highly or little infested areas correspond to the observations to a large extent. This correspondence between observed and predicted infested areas is reflected by the fact that the sum of the predicted number of infested trees in the domain over the years overlap well with the observed number of infested trees.

Using a semi-empirical predictive model is useful to get estimates of relevant parameters. For example, using our predictive model parameters, we could back-calculate the yearly rate of increase in Cypress Hills and found that it has an average of 2.1 in forested areas. Likewise, we found that the background infestation level is 0.03. This means that for each 500×500-meters cell, containing hundreds to thousands of pines in average, we expect an average of 0.03 pine to get infested from long-distance dispersal events or transitions from endemic to epidemic population. These parameters are not observed easily on the field.

In Cypress Hills over the study period, the average rate of increase of infested trees was 2.1 with three trees initially infested. This means that one infested tree produces 2.1 infested trees the following year in average. With these values, an outbreak is eradicated in average before 13 years for a control rate of at least 81%. This result is of the same order of magnitude as the results of Carroll *et al.* (2006) although their initial values are 10 000 infested trees. Using the model described by Carroll *et al.* (2006) and parametrized for a study area at the border of the provinces of British Columbia and Alberta, Canada, Coggins *et al.* (2011) estimated that 11 years of continuous control effort was necessary to achieve outbreak suppression with a control rate of 70%.

When an infested area is not controlled, outbreak duration can be very short (Fettig *et al.*, 2014). However, we do not observe this behaviour in our study. Our time period–13 years–is rather short. For a longer period of time under a small control rate, the number of pines available for beetle attacks might become limiting and thus, lead to outbreak collapse due to a lack of hosts (Safranyik, 2004).

Infested areas are typically not isolated in space. For example, intensive control or no control of *I. typographus* leads to similar tree mortality, mostly because of beetle flights from outside the intensively managed area (Grodzki *et al.*, 2006). We included this connectivity in our model by considering the beetle pressure from outside the park southern border and a background infestation level. Because of it, under any control effort, the outbreak first peaked and only started to decline once the infestations just outside the park limits had died out.

One limitation of this study is that we were constrained to the time period and location of the current Cypress Hills outbreak in order to use ecological and environmental factors as input for our model. Because of that, we were not able to have a precise estimate of the number of years of sustained control effort required to achieve outbreak suppression for control rate lower than 81%. However, we were able to determine that control levels above 55-60% are effectively decreasing beetle population sizes over time and so should eventually lead to suppression.

In conclusion, this study allowed to bridge the gap between theory and field observations. Direct control impact beetle outbreak duration and pine mortality even at a low level. Moderate control can eradicate mountain pine beetle infestations in the long term. However, a significant control rate is needed to achieve a quick suppression. Future work could examine the spatial patterns of infested and non-infested trees for outbreaks under different control effort in order to get further insights.

Chapter 6

General conclusion

Forest insect outbreaks are currently one of the main sources of disturbance in North American forests (Dale *et al.*, 2001). For example, mountain pine beetle outbreaks are affecting extremely large areas of pine forests in Canada since the 1900s (Taylor *et al.*, 2006; Walton, 2013). To avoid damages on biodiversity and loss of ecosystem services, pest management is necessary (Dale *et al.*, 2001; Chan-McLeod, 2006). In this thesis, I assessed the current MPB management strategy in Cypress Hills, Saskatchewan, and considered ways of improvement while gathering insights on pest control and population dynamics.

6.1 Mountain pine beetle control

The main conclusion arising from this thesis is that MPB control in Cypress Hills is efficient although it could be improved. In chapter 2, I show that managers in Cypress Hills are able to obtain a good control efficiency compared to a random strategy by detecting and controlling 62% of the infested trees in the park. In chapter 5, however, I show that this level of control is just enough to decrease the MPB population in the long term. Finding ways to increase further the efficiency would allow a more certain outcome. However, it is not necessary to reach 100% control in order to achieve suppression. Indeed, in populations presenting an Allee effect such as the MPB, reducing population size under their Allee threshold should suffice to eliminate them (Liebhold & Bascompte, 2003).

Control efficiency could be improved in three different ways. First, I show in chapter 2 that using survey radius larger than 50 meters would help increase detection–and thus control efficiency–in a cost-effective manner. Indeed, if more beetles are detected and controlled a given year, the future cost of controlling their offspring is avoided. Therefore, it is more cost-efficient to invest now rather than in the future.

Second, in the application section of chapter 3, I show that increasing the budget would help increase detection and control given a management strategy searching around red-top trees and the MPB spatial distribution, although this effect is less important at low density of infested trees. Therefore, some management goals are more feasible than others. According to the simulations, controlling more than 20% of the infested trees is an unattainable goal. However, removing all individuals above 0.5 infested trees per hectare is achievable. One possible explanation is the Allee effect: there might not be sufficient remaining beetles to produce offspring. Therefore, managers need to define carefully their objectives within the realm of possibilities. Studies like the one in chapter 3 can help define what are the management possibilities.

Finally, the main detection strategy used for MPB involves looking in the vicinity of previously infested trees. In chapter 2, I show that including environmental factors such as weather and stand characteristics into a predictive model in addition to beetle pressure and history gives good predictions of infested locations one year ahead. It is known that weather and stand characteristics affect MPB development and dispersal (Safranyik, 1978; Shore & Safranyik, 1992). These factors help calculate the susceptibility of a stand to be attacked by MPB. In addition, MPB typically disperse within a stand and thus, red-top trees are a good proxy for the detection of new attacks. Gathering both types of information into one model helps define areas at risk. It has been studied at large scales (Aukema *et al.*, 2008; Preisler *et al.*, 2012). However, maps of small-scale predictions, such as the ones presented in this thesis, provide a direct tool that managers can use to inform their surveys.

6.2 Mountain pine beetle population dynamics

As mentioned previously, weather, topography, vegetation, and beetle pressure and history influence the location of MPB-infested trees. In chapter 4, I show that their impact depend on the outbreak phase. The initiation of an outbreak requires an increase in MPB population size coming from either external sources or endemic populations. In 2006, Cypress Hills most likely received incoming flights from southern populations in the United States. Therefore, the most important factor was the distance to the park southern border. During the peak, beetles choose locations with large trees close to their emergence site. The outbreak collapse is driven by a decrease in susceptible and vigorous hosts. This information can be used to determine how areas at risk would vary in time. However, MPB are very sensitive to changes in microclimate and the presence of competing/facilitating species during specific population phases. Such data were not available over the study area extent and collecting them would help further distinct the outbreak phases.

In chapter 5, I simulate what would happen to the pine and beetle populations without management. I show that, although MPB outbreaks negatively affect pine populations, it would still be declining in the absence of MPB in Cypress Hills. This result shows that Cypress Hills is experiencing a change in the forest community that might not just be due to MPB outbreaks although MPB accelerate this change.

In chapter 3, I show that the interaction between human and mountain pine beetle populations can be modelled using functional responses. This is a novel way to look at pest management in the context of population dynamics. For instance, MPB control follows a type II functional response in Cypress Hills. Adding such component to a population model would allow precise projections of MPB populations given that MPB outbreaks are now mostly actively controlled. To complete the model even further regarding the link with humans communities, one could add a component describing the economic costs of MPB depending on the population density (Yokomizo *et al.*, 2009).

6.3 Limitations and applicability

Cypress Hills is a small isolated study area. Although, it may limit the scope of the research, it also gives some advantages. For example, the total cost of surveying the entire park extent is relatively low and thus feasible. Therefore, researchers can have access to extensive and extremely detailed surveys to parametrize their models. Furthermore, the park isolation combined with the fact that long-distance events are rare in MPB allows to study an outbreak with relatively low interference from other populations. Note that incoming flights still occur but in such low frequency and magnitude that their effect is minor. Finally, the outbreak studied is still in progress. Therefore, some important information regarding the outbreak collapse could be missing from the data set. Still, the number of MPB-infested trees in Cypress Hills is currently low and has been consistently decreasing for the past five years. In addition, information about potential competing and facilitating species was only partially available in the study area. Focusing on these aspects would greatly increase our understanding of the system. Future work could use such highly-detailed data set to corroborate and further simulation and empirical studies (*e.g.* Logan *et al.*, 1998; Aukema *et al.*, 2006) by including information about inter-species interactions and through the use of spatial analyzes.

The MPB genetic structure in Cypress Hills is close to the ones in southeastern British Columbia populations and most different from the ones in northern British Columbia populations (Samarasekera *et al.*, 2012). In the study of a MPB fungal symbiont, gene flow occurs between populations in Montana, southern British Columbia and Cypress Hills (Tsui *et al.*, 2012). Therefore, it is reasonable to assume that MPB in Cypress Hills have similar characteristics with southern British Columbia and Montana populations, and the results presented in this thesis would be similar from the ones obtained at these locations and in an environment similar to Cypress Hills.

In chapter 3, I present the application of the functional response framework to human-pest systems. This framework is transferable to human-resources systems with little modifications. Indeed, the components are survey area, survey and control costs, budget, pest intensity, number of pest controlled. Respectively, they could be transformed into foraging area, search and handling costs, budget, resource intensity, number of resource items harvested. Using such functional response in a population model of the resource would help get precise projections of future stocks. This is already partially implemented in fisheries models which a function similar to a type I functional response.

6.4 Model implications for pest management

Statistical models such as regressions are particularly useful to determine which environmental and ecological factors are explaining pest locations in different contexts (*e.g.* Gumpertz *et al.*, 2000; Aukema *et al.*, 2008). We often know from experimental and observational studies how these factors, such as weather and stand characteristics, affect the insect life-cycle. However, combining pest pressure information with ecological and environmental factors to obtain predictions of risk and associated uncertainty requires modelling tools (Koch *et al.*, 2009).

Local-scale risk maps created with such an approach can inform pest management. Using models to evaluate the risk associated with specific stands have the following advantages. First, the accuracy of the model can be evaluated quantitatively using a validation method such as cross-validation. Second, the model can produce a quantitative map of the infestation risk. Such maps allow managers to plan efficiently their future actions. Short-, medium-, and long-term risk predictions can be developed. However, uncertainty and variability increase for long-term estimations (Boyce *et al.*, 2006; Doak *et al.*, 2008).

Mechanistic models and simulations are valuable tools to use in the context of pest management. For example, with such tools, the efficiency of management strategies can be assessed and compared to one another extensively (*e.g.* Strohm *et al.*, 2016; Nelson *et al.*, 2018). Indeed, using simulations instead of real-life experiments allows to implement management scenarios that would be otherwise either too expensive to carry out, too hazardous in terms of ecological and environmental consequences, too extended in time or space to be feasible, or too caught up in interferences to separate the effects of specific factors. However, a simulated environment might lack realism in terms of mechanisms, parameter values, time and spatial scales, etc. (*e.g.* Keane *et al.*, 2002). Therefore, simulations on their own would not be useful. They need to be informed in details by experiments and observations.

Furthermore, mechanistic models can describe the population dynamics of a controlled pest more accurately if they include a realistic management component. Such component can have different forms, each of them affecting population growth and dispersal in different ways. For example, simply reducing the growth parameter would imply that pest density does not affect the efficiency of the management strategy which seems unrealistic. Therefore, using a well-defined framework such as the functional responses can help consider the implications of different management forms on population dynamics.

In conclusion, models are used more and more frequently in the context of pest management to assess detection and control strategies, and consider ways of improvement. Mountain pine beetle detection is improved by including, in addition to beetle pressure and history, weather, stand characteristics, and outbreak phase information. Its current control level leads to a decline in population but only an extremely high control effort can lead to a quick suppression of the beetle population. Reducing beetle populations to low densities help suppress outbreaks. These results are important for researchers and managers to consider as pest damages in North American forests are predicted to increase as a consequence of climate change (Bentz *et al.*, 2010).

Bibliography

Aho, K., Derryberry, D. & Peterson, T. 2014. Model selection for ecologists: The worldviews of AIC and BIC. *Ecology* **95**: 631–636.

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**: 716–723.

Akima, H. & Gebhardt, A. 2016. *akima: Interpolation of irregularly and regularly spaced data*. R package version 0.6-2.

Alberta Sustainable Resource Development 1985. Alberta phase 3 forest inventory. *Technical report*, Alberta Sustainable Resource Development.

Alberta Sustainable Resource Development 2007. Mountain pine beetle management strategy.

Allee, W.C. 1931. Animal aggregations. In: *A study in general sociology*. The University of Chicago Press, Chicago.

Arditi, R. & Ginzburg, L.R. 1989. Coupling in predator-prey dynamics: Ratio-dependence. *Journal of theoretical biology* **139**: 311–326.

Aukema, B.H., Carroll, A.L., Zheng, Y., Zhu, J., Raffa, K.F., Dan Moore, R. *et al.* 2008. Movement of outbreak populations of mountain pine beetle: influences of spatiotemporal patterns and climate. *Ecography* **31**: 348–358.

Aukema, B.H., Carroll, A.L., Zhu, J., Raffa, K.F., Sickley, T.A. & Taylor, S.W. 2006. Landscape level analysis of mountain pine beetle in British Columbia, Canada: Spatiotemporal development and spatial synchrony within the present outbreak. *Ecography* **29**: 427–441.

Avgar, T., Kuefler, D. & Fryxell, J.M. 2011. Linking rates of diffusion and consumption in relation to resources. *The American Naturalist* **178**: 182–190.

Baddeley, A., Rubak, E. & Turner, R. 2015. *Spatial point patterns: Methodology and applications with R*. Chapman and Hall/CRC Press, London.

Beaudoin, A., Bernier, P., Guindon, L., Villemaire, P., Guo, X., Stinson, G. *et al.* 2014. Mapping attributes of Canada's forests at moderate resolution through kNN and MODIS imagery. *Canadian Journal of Forest Research* **44**: 521–532.

Bentz, B.J., Duncan, J.P. & Powell, J.A. 2016. Elevational shifts in thermal suitability for mountain pine beetle population growth in a changing climate. *Forestry* **89**: 271–

283.

Bentz, B.J. & Powell, J.A. 2014. Mountain pine beetle seasonal timing and constraints to bivoltinism: A comment on Mitton and Ferrenberg, “Mountain pine beetle develops an unprecedented summer generation in response to climate warming”. *The American Naturalist* **184**: 787–796.

Bentz, B.J., Régnière, J., Fettig, C.J., Hansen, E.M., Hayes, J.L., Hicke, J.A. *et al.* 2010. Climate change and bark beetles of the western United States and Canada: Direct and indirect effects. *BioScience* **60**: 602–613.

Berec, L., Kean, J.M., Epanchin-Niell, R., Liebhold, A.M. & Haight, R.G. 2015. Designing efficient surveys: Spatial arrangement of sample points for detection of invasive species. *Biological Invasions* **17**: 445–459.

Billings, R.F., Gara, R.I. & Hrutfiord, B.F. 1976. Influence of ponderosa pine resin volatiles on the response of *Dendroctonus ponderosae* to synthetic trans-verbenol. *Environmental Entomology* **5**: 171–179.

Bleiker, K.P. & Six, D.L. 2007. Dietary benefits of fungal associates to an eruptive herbivore: Potential implications of multiple associates on host population dynamics. *Environmental Entomology* **36**: 1384–1396.

Boissard, P., Martin, V. & Moisan, S. 2008. A cognitive vision approach to early pest detection in greenhouse crops. *Computers and Electronics in Agriculture* **62**: 81–93.

Bordon, J.H. 1982. Aggregation pheromones. In: J.B. Mitton & K.B. Sturgeon, eds., *Bark Beetles in North American Conifers A System for the Study of Evolutionary Biology*, pp. 74–139. University of Texas Press, Austin, TX.

Boyce, M.S., Haridas, C.V., Lee, C.T. & the NCEAS Stochastic Demography Working Group 2006. Demography in an increasingly variable world. *Trends in Ecology & Evolution* **21**: 141–148.

Bradley, A.P. 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition* **30**: 1145–1159.

Burnham, K.P. & Anderson, D.R. 2002. *Model selection and multimodel inference: A practical information-theoretic approach.* Springer, New York, NY, 2nd edn..

Carroll, A.L. & Safranyik, L. 2004. The bionomics of the mountain pine beetle in lodgepole pine forests: Establishing a context. *Information Report BC-X-399*, Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, British Columbia, Canada.

Carroll, A.L., Shore, T.L. & Safranyik, L. 2006. Direct control: Theory and practice. In: L. Safranyik & B. Wilson, eds., *The mountain pine beetle: a synthesis of biology, management, and impacts on lodgepole pine.*, pp. 155–172. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, British Columbia.

Carroll, A.L., Taylor, S.W., Régnière, J. & Safranyik, L. 2004. Effects of climate change on range expansion by the mountain pine beetle in British Columbia. *Infor-*

mation Report BC-X-399, Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Kelowna, British Columbia, Canada.

Cerezke, H.F. 1989. Mountain pine beetle aggregation semiochemical use in Alberta and Saskatchewan, 1983-1987. In: *Proceedings from the Symposium on the management of lodgepole pine to minimize losses to the mountain pine beetle, Kalispell, Montana July 12–14, 1988*, General Technical Report INT, pp. 108–113. USDA Forest Service, Intermountain Research Station, Ogden, UT.

Chan-McLeod, A.C.A. 2006. A review and synthesis of the effects of unsalvaged mountain-pine-beetle-attacked stands on wildlife and implications for forest management. *Journal of Ecosystems and Management* **7**: 119–132.

Chen, H. & Jackson, P.L. 2017. Climatic conditions for emergence and flight of mountain pine beetle: Implications for long-distance dispersal. *Canadian Journal of Forest Research* **47**: 974–984.

Chen, H. & Walton, A. 2011. Mountain pine beetle dispersal: Spatiotemporal patterns and role in the spread and expansion of the present outbreak. *Ecosphere* **2**: 1–17.

Coggins, S.B., Coops, N.C., Wulder, M.A., Bater, C.W. & Ortlepp, S.M. 2011. Comparing the impacts of mitigation and non-mitigation on mountain pine beetle populations. *Journal of Environmental Management* **92**: 112–120.

Cole, W.E. 1981. Some risks and causes of mortality in mountain pine beetle populations: A long-term analysis. *Researches on Population Ecology* **23**: 116–144.

Cooke, B.J. & Carroll, A.L. 2017. Predicting the risk of mountain pine beetle spread to eastern pine forests: Considering uncertainty in uncertain times. *Forest Ecology and Management* **396**: 11–25.

Cosner, C., DeAngelis, D.L., Ault, J.S. & Olson, D.B. 1999. Effects of spatial grouping on the functional response of predators. *Theoretical Population Biology* **56**: 65–75.

Creeden, E.P., Hicke, J.A. & Buotte, P.C. 2014. Climate, weather, and recent mountain pine beetle outbreaks in the western United States. *Forest Ecology and Management* **312**: 239–251.

Cullingham, C.I., Cooke, J.E.K., Dang, S., Davis, C.S., Cooke, B.J. & Coltman, D.W. 2011. Mountain pine beetle host-range expansion threatens the boreal forest. *Molecular Ecology* **20**: 2157–2171.

Cullingham, C.I., James, P.M.A., Cooke, J.E.K. & Coltman, D.W. 2012. Characterizing the physical and genetic structure of the lodgepole pine x jack pine hybrid zone: Mosaic structure and differential introgression. *Evolutionary Applications* **5**: 879–891.

Dale, V.H., Joyce, L.A., McNulty, S., Neilson, R.P., Ayres, M.P., Flannigan, M.D. *et al.* 2001. Climate change and forest disturbances. *BioScience* **51**: 723–734.

Davis, J. & Goadrich, M. 2006. The relationship between precision-recall and ROC curves. In: *Proceedings of the 23rd international conference on Machine learning - ICML '06*, pp. 233–240. ACM Press, Pittsburgh, Pennsylvania.

de la Giroday, H.M.C., Carroll, A.L. & Aukema, B.H. 2012. Breach of the northern Rocky Mountain geoclimatic barrier: Initiation of range expansion by the mountain pine beetle. *Journal of Biogeography* **39**: 1112–1123.

Dhar, A., Parrott, L. & Heckbert, S. 2016. Consequences of mountain pine beetle outbreak on forest ecosystem services in western Canada. *Canadian Journal of Forest Research* **46**: 987–999.

Dick, J.T.A., Alexander, M.E., Jeschke, J.M., Ricciardi, A., MacIsaac, H.J., Robinson, T.B. *et al.* 2013. Advancing impact prediction and hypothesis testing in invasion ecology using a comparative functional response approach. *Biological Invasions* **16**: 735–753.

Diskin, M., Rocca, M.E., Nelson, K.N., Aoki, C.F. & Romme, W. 2011. Forest developmental trajectories in mountain pine beetle disturbed forests of Rocky Mountain National Park, Colorado. *Canadian Journal of Forest Research* **41**: 782–792.

Doak, D.F., Estes, J.A., Halpern, B.S., Jacob, U., Lindberg, D.R., Lovvorn, J. *et al.* 2008. Understanding and predicting ecological dynamics: Are major surprises inevitable. *Ecology* **89**: 952–961.

Elliott, L.P. & Brook, B.W. 2007. Revisiting Chamberlin: Multiple working hypotheses for the 21st century. *BioScience* **57**: 608–614.

Epanchin-Niell, R.S. & Hastings, A. 2010. Controlling established invaders: Integrating economics and spread dynamics to determine optimal management. *Ecology Letters* **13**: 528–541.

Erbilgin, N. 2019. Phytochemicals as mediators for host range expansion of a native invasive forest insect herbivore. *New Phytologist* **221**: 1268–1278.

Erbilgin, N., Cale, J.A., Hussain, A., Ishangulyyeva, G., Klutsch, J.G., Najar, A. *et al.* 2017a. Weathering the storm: How lodgepole pine trees survive mountain pine beetle outbreaks. *Oecologia* **184**: 469–478.

Erbilgin, N., Cale, J.A., Lusebrink, I., Najar, A., Klutsch, J.G., Sherwood, P. *et al.* 2017b. Water-deficit and fungal infection can differentially affect the production of different classes of defense compounds in two host pines of mountain pine beetle. *Tree Physiology* **37**: 338–350.

Fettig, C.J., Gibson, K.E., Munson, A.S. & Negrón, J.F. 2014. Cultural practices for prevention and mitigation of mountain pine beetle infestations. *Forest Science* **60**: 450–463.

Fettig, C.J. & Hilszczański, J. 2015. Chapter 14 - Management strategies for bark beetles in conifer forests. In: F.E. Vega & R.W. Hofstetter, eds., *Bark Beetles*, pp. 555–584. Academic Press, San Diego.

Fettig, C.J., Klepzig, K.D., Billings, R.F., Munson, A.S., Nebeker, T.E., Negrón, J.F. *et al.* 2007. The effectiveness of vegetation management practices for prevention and control of bark beetle infestations in coniferous forests of the western and southern United States. *Forest Ecology and Management* **238**: 24–53.

Finke, D.L. & Denno, R.F. 2002. Intraguild predation diminished in cmplex-structured vegetation: Implications for prey suppression. *Ecology* **83**: 643–652.

Ghosh, J. & Samanta, T. 2001. Model selection - An overview. *Current Science* **80**: 1135–1144.

Goodsman, D.W., Koch, D., Whitehouse, C., Evenden, M.L., Cooke, B.J. & Lewis, M.A. 2016. Aggregation and a strong Allee effect in a cooperative outbreak insect. *Ecological Applications* **26**: 2623–2636.

Goodsman, D.W. & Lewis, M.A. 2016. The minimum founding population in dispersing organisms subject to strong Allee effects. *Methods in Ecology and Evolution* **7**: 1100–1109.

Gray, C.A., Runyon, J.B., Jenkins, M.J. & Giunta, A.D. 2015. Mountain pine beetles use volatile cues to locate host limber pine and avoid non-host great basin bristlecone pine. *PLOS ONE* **10**: 1–13.

Grodzki, W., Jakuš, R., Lajzová, E., Sitková, Z., Maczka, T. & Škvarenina, J. 2006. Effects of intensive versus no management strategies during an outbreak of the bark beetle *Ips typographus* (L.) (Col.: Curculionidae, Scolytinae) in the Tatra Mts. in Poland and Slovakia. *Annals of Forest Science* **63**: 55–61.

Gumpertz, L., Wu & Pye, M. 2000. Logistic regression for southern pine beetle outbreaks with spatial and temporal autocorrelation. *Forest Science* **46**: 95–107.

Gustafson, E.J. 1998. Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems* **1**: 143–156.

Hassell, M.P. 1978. *The dynamics of arthropod predator-prey systems.* Princeton University Press.

He, Y., Zhao, J., Zheng, Y., Desneux, N. & Wu, K. 2012. Lethal effect of imidacloprid on the coccinellid predator *Serangium japonicum* and sublethal effects on predator voracity and on functional response to the whitefly *Bemisia tabaci*. *Ecotoxicology* **21**: 1291–1300.

Heavilin, J. & Powell, J. 2008. A novel method of fitting spatio-temporal models to data, with applications to the dynamics of mountain pine beetles. *Natural Resource Modeling* **21**: 489–524.

Hochberg, M.E. & Holt, R.D. 1999. The uniformity and density of pest exploitation as guides to success in biological control. In: B.A. Hawkins & H.V. Cornell, eds., *Theoretical approaches to biology control*, pp. 71–88. Cambridge University Press, Cambridge.

Holling, C.S. 1959a. The components of predation as revealed by a study of small-mammal predation of the european pine sawfly. *The Canadian Entomologist* **91**: 293–320.

Holling, C.S. 1959b. Some characteristics of simple types of predation and parasitism. *The Canadian Entomologist* **91**: 385–398.

Hopkins, B. & Skellam, J.G. 1954. A new method for determining the type of distribution of plant individuals. *Annals of Botany* **18**: 213–227.

Hossie, T.J. & Murray, D.L. 2016. Spatial arrangement of prey affects the shape of ratio-dependent functional response in strongly antagonistic predators. *Ecology* **97**: 834–841.

Hynum, B.G. & Berryman, A.A. 1980. *Dendroctonus ponderosae* (Coleoptera: Scolytidae): Pre-aggregation landing and gallery initiation on lodgepole pine. *The Canadian Entomologist* **112**: 185–191.

Ives, A.R., Schooler, S.S., Jagar, V.J., Knuteson, S.E., Grbic, M. & Settle, W.H. 1999. Variability and parasitoid foraging efficiency: A case study of pea aphids and *Aphidius ervi*. *The American Naturalist* **154**: 652–673.

Jackson, P.L., Straussfogel, D., Lindgren, B.S., Mitchell, S. & Murphy, B. 2008. Radar observation and aerial capture of mountain pine beetle, *Dendroctonus ponderosae* Hopk. (Coleoptera: Scolytidae) in flight above the forest canopy. *Canadian Journal of Forest Research* **38**: 2313–2327.

Janssen, A. & Sabelis, M.W. 1992. Phytoseiid life-histories, local predator-prey dynamics, and strategies for control of tetranychid mites. *Experimental & Applied Acarology* **14**: 233–250.

Jeschke, J.M., Kopp, M. & Tollrian, R. 2002. Predator functional responses: Discriminating between handling and digesting prey. *Ecological Monographs* **72**: 95–112.

Jeschke, J.M., Kopp, M. & Tollrian, R. 2004. Consumer-food systems: Why type I functional responses are exclusive to filter feeders. *Biological Reviews* **79**: 337–349.

Keane, R.E., Parsons, R.A. & Hessburg, P.F. 2002. Estimating historical range and variation of landscape patch dynamics: limitations of the simulation approach. *Ecological Modelling* **151**: 29–49.

Kermack, W.O. & McKendrick, A. G. 1927. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A* **115**: 700–721.

Koch, F.H., Yemshanov, D., McKenney, D.W. & Smith, W.D. 2009. Evaluating critical uncertainty thresholds in a spatial model of forest pest invasion risk. *Risk Analysis* **29**: 1227–1241.

Krause, A.M., Townsend, P.A., Lee, Y. & Raffa, K.F. 2018. Predators and competitors of the mountain pine beetle *Dendroctonus ponderosae* (Coleoptera: Curculionidae) in stands of changing forest composition associated with elevation. *Agricultural and Forest Entomology* **20**: 402–413.

Kuhn, M. 2018. *caret: Classification and regression training.*

Kunegel-Lion, M., McIntosh, R.L. & Lewis, M.A. 2019. Management assessment of mountain pine beetle infestation in Cypress Hills, SK. *Canadian Journal of Forest Research* **49**: 154–163.

Kutner, M.H., Kutner, M.H., Nachtsheim, C. & Neter, J. 2004. *Student solutions manual for use with applied linear regression models.* McGraw-Hill/Irwin.

Levin, S.A. 1992. The problem of pattern and scale in ecology: The Robert H. MacArthur award lecture. *Ecology* **73**: 1943–1967.

Lewis, M.A., Nelson, W. & Xu, C. 2010. A structured threshold model for mountain pine beetle outbreak. *Bulletin of Mathematical Biology* **72**: 565–589.

Liebhold, A. & Bascompte, J. 2003. The Allee effect, stochastic dynamics and the eradication of alien species. *Ecology Letters* **6**: 133–140.

Little, E.L.J. 1971. Atlas of United States trees. *Tech. Rep. Vol 1, conifers and important hardwoods*, U.S. Department of Agriculture Miscellaneous Publication 1146.

Liu, B., Teng, Z. & Chen, L. 2006. Analysis of a predator–prey model with Holling II functional response concerning impulsive control strategy. *Journal of Computational and Applied Mathematics* **193**: 347–362.

Lockwood, J.L., Hoopes, M.F. & Marchetti, M.P. 2013. *Invasion ecology.* John Wiley & Sons.

Logan, J. & Powell, J. 2001. Ghost forests, global warming, and the mountain pine beetle (Coleoptera : Scolytidae). *American Entomologist* **47**: 160–173.

Logan, J.A., White, P., Bentz, B.J. & Powell, J.A. 1998. Model analysis of spatial patterns in mountain pine beetle outbreaks. *Theoretical Population Biology* **53**: 236–255.

Long, J.A., Robertson, C., Nathoo, F.S. & Nelson, T.A. 2012. A Bayesian space–time model for discrete spread processes on a lattice. *Spatial and Spatio-temporal Epidemiology* **3**: 151–162.

Lotan, J., Brown, J. & Neuenschwander, L. 1985. Role of fire in lodgepole pine forests. *Lodgepole Pine : The Species and Its Management Symposium Proceedings* pp. 133–152.

Lusebrink, I., Erbilgin, N. & Evenden, M.L. 2016. The effect of water limitation on volatile emission, tree defense response, and brood success of *Dendroctonus ponderosae* in two pine hosts, lodgepole, and jack pine. *Frontiers in Ecology and Evolution* **4**: 1–13.

Marmion, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. 2009. The performance of state-of-the-art modelling techniques depends on geographical distribution of species. *Ecological Modelling* **220**: 3512–3520.

Matthews, B., Netherer, S., Katzensteiner, K., Pennerstorfer, J., Blackwell, E., Henschke, P. *et al.* 2018. Transpiration deficits increase host susceptibility to bark beetle attack: Experimental observations and practical outcomes for *Ips typographus* hazard assessment. *Agricultural and Forest Meteorology* **263**: 69–89.

Maxwell, B.D., Backus, V., Hohmann, M.G., Irvine, K.M., Lawrence, P., Lehnhoff, E.A. *et al.* 2012. Comparison of transect-based standard and adaptive sampling methods for invasive plant species. *Invasive Plant Science and Management* **5**: 178–193.

McCambridge, W.F. 1971. Temperature limits of flight of the mountain pine beetle, *Dendroctonus ponderosae*. *Annals of the Entomological Society of America* **64**: 534–

535.

McCreadie, J.W. & Adler, P.H. 2019. Spatial distributions of common and rare species in a widespread group of stream insects. *Freshwater Science* **38**: 53–63.

Messier, F. 1994. Ungulate population models with predation: A case study with the North American moose. *Ecology* **75**: 478–488.

Metz, C.E. 1978. Basic principles of ROC analysis. *Seminars in Nuclear Medicine* **8**: 283–298.

Mills, N.J. & Getz, W.M. 1996. Modelling the biological control of insect pests: A review of host-parasitoid models. *Ecological Modelling* **92**: 121–143.

Mitton, J.B. & Ferrenberg, S. 2014. Field studies demonstrate bivoltinism in the mountain pine beetle: A reply to Bentz and Powell. *The American Naturalist* **184**: 797–801.

Mitton, J.B. & Ferrenberg, S.M. 2012. Mountain pine beetle develops an unprecedented summer generation in response to climate warming. *The American Naturalist* **179**: E163–E171.

Monserud, R.A., Huang, S. & Yang, Y. 2006. Predicting lodgepole pine site index from climatic parameters in Alberta. *The Forestry Chronicle* **82**: 562–571.

Murdoch, W.W. & Oaten, A. 1975. Predation and population stability. *Advances in Ecological Research* **9**: 1–131.

Nachman, G. 2006. A functional response model of a predator population foraging in a patchy habitat. *Journal of Animal Ecology* **75**: 948–958.

Negrón, J.F. & Fettig, C.J. 2014. Mountain pine beetle, a major disturbance agent in US western coniferous forests: A synthesis of the state of knowledge. *Forest Science* **60**: 409–413.

Nelson, M.F., Murphy, J.T., Bone, C. & Altaweel, M. 2018. Cyclic epidemics, population crashes, and irregular eruptions in simulated populations of the mountain pine beetle, *Dendroctonus ponderosae*. *Ecological Complexity* **36**: 218–229.

Nelson, T., Boots, B., White, K.J. & Smith, A.C. 2006. The impact of treatment on mountain pine beetle infestation rates. *Journal of Ecosystems and Management* **7**: 20–36.

Ono, H. 2004. The mountain pine beetle: Scope of the problem and key issues in Alberta. *Information Report BC-X-399*, Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, British Columbia, Canada.

Potapov, A., Muirhead, J., Yan, N., Lele, S. & Lewis, M. 2011. Models of lake invasibility by *Bythotrephes longimanus*, a non-indigenous zooplankton. *Biological Invasions* **13**: 2459–2476.

Powell, J.A., Garlick, M.J., Bentz, B.J. & Friedenberg, N. 2018. Differential dispersal and the Allee effect create power-law behaviour: Distribution of spot infestations during mountain pine beetle outbreaks. *Journal of Animal Ecology* **87**: 73–86.

Preisler, H.K., Hicke, J.A., Ager, A.A. & Hayes, J.L. 2012. Climate and weather influences on spatial temporal patterns of mountain pine beetle populations in Washington and Oregon. *Ecology* **93**: 2421–2434.

Pye, J.M., Holmes, T.P., Prestemon, J.P. & Wear, D.N. 2011. Economic impacts of the southern pine beetle. In: R.N. Coulson & K.D. Klepzig, eds., *Southern Pine Beetle II*, *General Technical Report*, vol. SRS-140, pp. 213–222. U.S. Department of Agriculture Forest Service, Southern Research Station, Asheville, NC.

R Core Team 2016. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

R Core Team 2018. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Raffa, K.F., Aukema, B.H., Bentz, B.J., Carroll, A.L., Hicke, J.A., Turner, M.G. *et al.* 2008. Cross-scale drivers of natural disturbances prone to anthropogenic amplification: The dynamics of bark beetle eruptions. *BioScience* **58**: 501–517.

Raffa, K.F. & Berryman, A.A. 1982. Gustatory cues in the orientation of *Dendroctonus ponderosae* (Coleoptera: Scolytidae) to host trees. *The Canadian Entomologist* **114**: 97–104.

Raffa, K.F. & Berryman, A.A. 1983. The role of host plant resistance in the colonization behavior and ecology of bark beetles (Coleoptera: Scolytidae). *Ecological Monographs* **53**: 27–49.

Raghavan, V., Bollmann, P. & Jung, G.S. 1989. A critical investigation of recall and precision as measures of retrieval system performance. *ACM Trans. Inf. Syst.* **7**: 205–229.

Ridgeway, G. 2015. *gbm: Generalized boosted regression models.*

Riel, W.G., Fall, A., Shore, T.L. & Safranyik, L. 2004. A spatio-temporal simulation of mountain pine beetle impacts on the landscape. *Information Report BC-X-399*, Pacific Forestry Centre, Canadian Forest Service, Kelowna, British Columbia, Canada.

Rincon, D.F., Cañas, L.A. & Hoy, C.W. 2017. Modeling changes in predator functional response to prey across spatial scales. *Theoretical Ecology* **10**: 403–415.

Robertson, C., Nelson, T.A. & Boots, B. 2007. Mountain pine beetle dispersal: The spatial-temporal interaction of infestations. *Forest Science* **53**: 395–405.

Rodrigues, L.A.D., Mistro, D.C., Cara, E.R., Petrovskaya, N. & Petrovskii, S. 2015. Patchy invasion of stage-structured alien species with short-distance and long-distance dispersal. *Bulletin of Mathematical Biology* **77**: 1583–1619.

Romme, W.H., Knight, D.H. & Yavitt, J.B. 1986. Mountain pine beetle outbreaks in the Rocky Mountains: Regulators of primary productivity? *The American Naturalist* **127**: 484–494.

Rosenberger, D.W., Venette, R.C., Maddox, M.P. & Aukema, B.H. 2017. Colonization behaviors of mountain pine beetle on novel hosts: Implications for range expansion into northeastern North America. *PLOS ONE* **12**: 1–26.

Régnière, J. & Bentz, B. 2007. Modeling cold tolerance in the mountain pine beetle, *Dendroctonus ponderosae*. *Journal of Insect Physiology* **53**: 559–572.

Régnière, J., Saint-Amant, R. & Béchard, A. 2014. BioSIM 10 – User's manual. *Tech. Rep.*, Natural Resources Canada, Canadian Forest Service, Laurentian Forest Center, Québec (QC).

Saab, V.A., Latif, Q.S., Rowland, M.M., Johnson, T.N., Chalfoun, A.D., Buskirk, S.W. *et al.* 2014. Ecological consequences of mountain pine beetle outbreaks for wildlife in western North American forests. *Forest Science* **60**: 539–559.

Safranyik, L. 1978. Effects of climate and weather on mountain pine beetle populations. In: A.A. Berryman, G. Amman & R. Stark, eds., *Proceedings of Symposium on Theory and Practice of Mountain Pine Beetle Management in Lodgepole Pine Forests, April 25-27, 1978, Washington State University, Pullman, Washington*, pp. 77–84. College of Forestry, Wildlife and Range Sciences, University of Idaho, Moscow, Idaho.

Safranyik, L. 1988. Estimating attack and brood totals and densities of the mountain pine beetle in individual lodgepole pine trees. *The Canadian Entomologist* **120**: 323–331.

Safranyik, L., ed. 1999. *Effects of induced competitive interactions with secondary bark beetle species on the establishment and survival of mountain pine beetle broods in lodgepole pine*. No. BC-X-384 in Information report. Canadian Forest Service, Pacific Forestry Centre, Victoria, B.C.

Safranyik, L. 2004. Mountain pine beetle epidemiology in lodgepole pine. In: T. Shore, J. Brooks & J. Stones, eds., *Mountain Pine Beetle Symposium: Challenge and Solutions. October 30-31, 2003, Kelowna, British Columbia.*, Information Report BC-X-399, pp. 33 – 40. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, B.C.

Safranyik, L., Carroll, A., Régnière, J., Langor, D., Riel, W., Shore, T. *et al.* 2010. Potential for range expansion of mountain pine beetle into the boreal forest of North America. *The Canadian Entomologist* **142**: 415–442.

Safranyik, L. & Carroll, A.L. 2006. The biology and epidemiology of the mountain pine beetle in lodgepole pine forests. In: L. Safranyik & B. Wilson, eds., *The mountain pine beetle: a synthesis of biology, management and impacts on lodgepole pine*, pp. 3–66. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, Canada.

Safranyik, L., Shrimpton, D.M. & Whitney, H.S. 1975. An interpretation of the interaction between lodgepole pine, the mountain pine beetle, and its associated blue stain fungi in western Canada. In: *Management of Lodgepole Pine Ecosystems Sympo sium Proceedings*, pp. 406 – 428. Washington State University Coop. Extension Service, Pullman, WA, Pullman, Washington, USA.

Safranyik, L., Silversides, R., McMullen, L.H. & Linton, D.A. 1989. An empirical approach to modeling the local dispersal of the mountain pine beetle (*Dendroctonus ponderosae* Hopk.) (Col., Scolytidae) in relation to sources of attraction, wind direction and speed. *Journal of Applied Entomology* **108**: 498–511.

Saito, T. & Rehmsmeier, M. 2015. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE* **10**: 1–21.

Samarasekera, G.D.N.G., Bartell, N.V., Lindgren, B.S., Cooke, J.E.K., Davis, C.S., James, P.M.A. *et al.* 2012. Spatial genetic structure of the mountain pine beetle (*Dendroctonus ponderosae*) outbreak in western Canada: Historical patterns and contemporary dispersal. *Molecular Ecology* **21**: 2931–2948.

Sambaraju, K.R., Carroll, A.L., Zhu, J., Stahl, K., Moore, R.D. & Aukema, B.H. 2012. Climate change could alter the distribution of mountain pine beetle outbreaks in western Canada. *Ecography* **35**: 211–223.

Saphores, J.D.M. & Shogren, J.F. 2005. Managing exotic pests under uncertainty: optimal control actions and bioeconomic investigations. *Ecological Economics* **52**: 327–339.

Saskatchewan Ministry of Environment 2016. Mountain pine beetle operational procedures document.

Schwarz, G. 1978. Estimating the dimension of a model. *The Annals of Statistics* **6**: 461–464.

Seidl, R., Rammer, W., Jäger, D. & Lexer, M.J. 2008. Impact of bark beetle (*Ips typographus* L.) disturbance on timber production and carbon sequestration in different management strategies under climate change. *Forest Ecology and Management* **256**: 209–220.

Shore, T.L., Brooks, J.E. & Stone, J.E., eds. 2004. *Mountain pine beetle symposium: Challenges and solutions. October 30-31, 2003, Kelowna, British Columbia.* Information Report BC-X-399. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, BC Canada.

Shore, T.L. & Safranyik, L. 1992. *Susceptibility and risk rating systems for the mountain pine beetle in lodgepole pine stands*, vol. 336.

Shore, T.L., Safranyik, L. & Whitehead, R.J. 2006. Principles and concepts of management. In: L. Safranyik & B. Wilson, eds., *The Mountain Pine Beetle: A Synthesis of Biology, Management, and Impacts on Lodgepole Pine*, pp. 117–121. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, Canada.

Sidder, A.M., Kumar, S., Laituri, M. & Sibold, J.S. 2016. Using spatiotemporal correlative niche models for evaluating the effects of climate change on mountain pine beetle. *Ecosphere* **7**: 1–22.

Six, D.L., Biber, E. & Long, E. 2014. Management for mountain pine beetle outbreak suppression: Does relevant science support current policy? *Forests* **5**: 103–133.

Smith, G.D., Carroll, A.L. & Lindgren, B.S. 2011. Facilitation in bark beetles: Endemic mountain pine beetle gets a helping hand. *Agricultural and Forest Entomology* **13**: 37–43.

Sokal, R. & Rohlf, F. 1995. *Biometry*. A series of books in biology. W. H. Freeman.

StatSoft Inc. 2013. Introduction to boosting trees for regression and classification. In: *Electronic Statistics Textbook*. StatSoft, Tulsa, OK.

Stephens, P.A., Sutherland, W.J. & Freckleton, R.P. 1999. What is the Allee effect? *Oikos* **87**: 185–190.

Strohm, S., Reid, M.L. & Tyson, R.C. 2016. Impacts of management on mountain pine beetle spread and damage: A process-rich model. *Ecological Modelling* **337**: 241–252.

Strom, B.L. & Clarke, S.R. 2011. Use of semiochemicals for southern pine beetle infestation management and resource protection. In: *Southern Pine Beetle II*, no. SRS-140 in General Technical Report (GTR), pp. 381–397. U.S. Department of Agriculture Forest Service, Southern Research Station, Asheville, NC.

Sutherland, W.J. 2001. Sustainable exploitation: A review of principles and methods. *Wildlife Biology* **7**: 131–140.

Taylor, S.W., Carroll, A.L., Alfaro, R.I. & Safranyik, L. 2006. Forest, climate and mountain pine beetle outbreak dynamics in western Canada. In: L. Safranyik & B. Wilson, eds., *The mountain pine beetle: a synthesis of biology, management and impacts on lodgepole pine*, pp. 67–94. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, Canada.

Therrien, J., Mason, C.J., Cale, J.A., Adams, A., Aukema, B.H., Currie, C.R. *et al.* 2015. Bacteria influence mountain pine beetle brood development through interactions with symbiotic and antagonistic fungi: Implications for climate-driven host range expansion. *Oecologia* **179**: 467–485.

Thompson, S.K. 1990. Adaptive cluster sampling. *Journal of the American Statistical Association* **85**: 1050–1059.

Trzcinski, M.K. & Reid, M.L. 2008. Effect of management on the spatial spread of mountain pine beetle (*Dendroctonus ponderosae*) in Banff National Park. *Forest Ecology and Management* **256**: 1418–1426.

Tsui, C.K.M., Roe, A.D., El-Kassaby, Y.A., Rice, A.V., Alamouti, S.M., Sperling, F.a.H. *et al.* 2012. Population structure and migration pattern of a conifer pathogen, *Grosmannia clavigera*, as influenced by its symbiont, the mountain pine beetle. *Molecular Ecology* **21**: 71–86.

Turchin, P. 2013. *Complex population dynamics: A theoretical/empirical synthesis*. Monographs in Population Biology. Princeton University Press, Princeton.

Van Den Berg, H., Ankasah, D., Muhammad, A., Rusli, R., Widayanto, H.A., Wirasto, H.B. *et al.* 1997. Evaluating the role of predation in population fluctuations of the soybean aphid *Aphis glycines* in farmer's fields in Indonesia. *Journal of Applied Ecology*

34: 971–984.

Venier, L. & Holmes, S. 2010. A review of the interaction between forest birds and eastern spruce budworm. *Environmental Reviews* **18**: 191–207.

Walton, A. 2013. Provincial-level projection of the current mountain pine beetle outbreak: Update of the infestation projection based on the Provincial Aerial Overview Surveys of Forest Health conducted from 1999 through 2012 and the BCMPB model (year 10). *Tech. Rep.*, BC Ministry of Forests, Lands and Natural Resources Operations, Victoria, BC.

Wermelinger, B. 2004. Ecology and management of the spruce bark beetle *Ips typographus*—A review of recent research. *Forest Ecology and Management* **202**: 67–82.

Wulder, M.A., Ortlepp, S.M., White, J.C., Coops, N.C. & Coggins, S.B. 2009. Monitoring the impacts of mountain pine beetle mitigation. *Forest Ecology and Management* **258**: 1181–1187.

Yokomizo, H., Possingham, H.P., Thomas, M.B. & Buckley, Y.M. 2009. Managing the impact of invasive species: The value of knowing the density–impact curve. *Ecological Applications* **19**: 376–386.

Youden, W.J. 1950. Index for rating diagnostic tests. *Cancer* **3**: 32–35.

Youssef, A.M., Pourghasemi, H.R., Pourtaghi, Z.S. & Al-Katheeri, M.M. 2016. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. *Landslides* **13**: 839–856.

Appendices

Appendix A: Varying the probability exponent

To vary the amount of noise that we introduced in the random sampling of locations from the model probabilities, we raised the model probabilities to an exponent ranging from 0 to 5. We then sampled the locations without replacement using the new probabilities as weight. The exponent 0 gives the same weight to all locations and, therefore, would give results equivalent to the random strategy. In opposition, a high exponent value increases the differences between low and high probabilities and eventually leads to a deterministic situation where the same locations with the highest probabilities are always chosen.

When we fixed the net survey area and varied the exponent, the predictions control efficiency varies from values similar to the random search at exponent 0 to values similar to the local search at high exponent (Fig. A.1). When the fixed survey area is equivalent to the one used in the current strategy (2 200 000 m^2), we can see that the local control efficiency is always higher than the predictions control efficiency no matter the exponent value. However, for a net survey area of 5 000 000 m^2, the prediction control efficiency is larger than the local control efficiency for an exponent value from about 1-1.5 to 5.

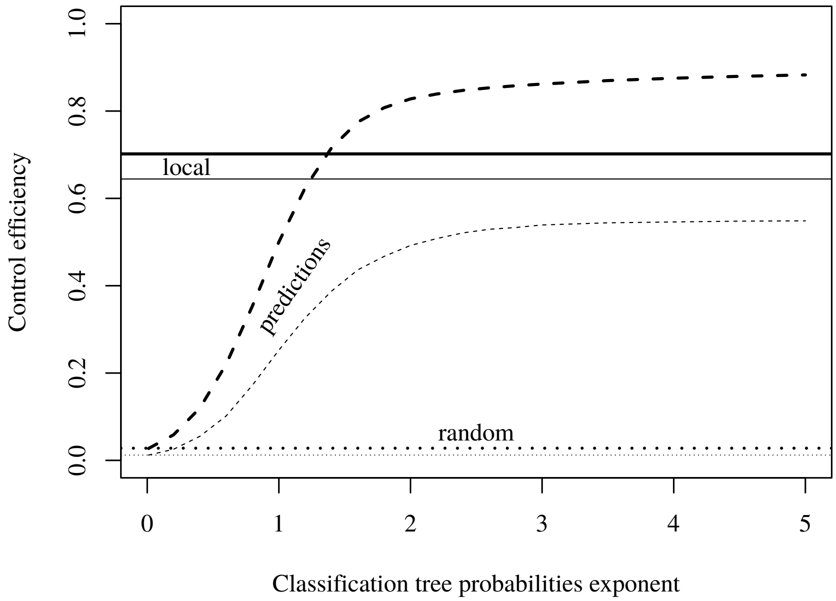

Figure A.1 – Control efficiency in relation to the classification tree probabilities exponent. Increasing the classification tree probabilities exponent gives more weight to locations with high predicted risks of infestation. Solid lines represent the local search around red-top trees for 2011. Dashed lines represent the search at locations chosen from model predictions for 2011. Dotted lines represent the search around random locations for 2011. Thin lines correspond to a survey area equivalent to the current Forest Service strategy (50 meter-radius circular plot; 2 200 000 m^2). Thick lines correspond to a survey area of 6 000 000 m^2 which correspond to the circular plot radius 90 m for the local search. The data for 2012, not presented here, display similar patterns.

Appendix B: Varying the cost of a missed green infestation

We varied the cost of a missed green infestation θ from 0 to 2000 and investigated its impact on the optimal survey area and the minimum cost per controlled tree depending on the detection strategy.

The optimal net survey area increases with θ for both the local and predictions strategies, although the optimal area is consistently larger using the predictions strategy (Fig. B.1a). However, the minimum total cost per controlled tree associated with the optimal survey area is lower for the predictions strategy than the local strategy for $\theta \geqslant 500$ (Fig. B.1b).

This means that the more expensive a green infestation, *i.e.* the more new infestations produced by one infested tree, the better in term of costs it is to increase the management effort now rather than controlling the additional new infestations in the future.

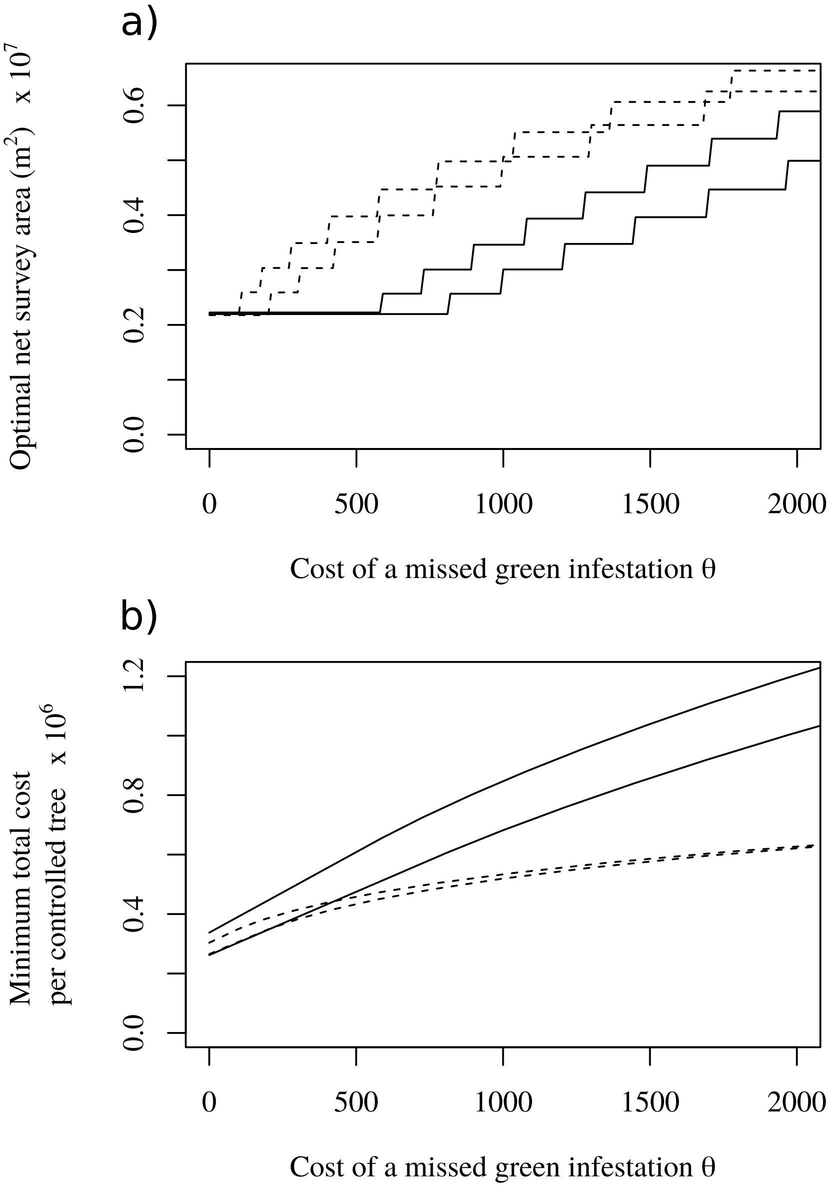

Figure B.1 – Optimal net survey area (a) and minimum total cost per controlled tree (b) in relation to the cost of missing a green infestation θ. Solid lines represent the values for the local search whereas dashed lines represent the values for the model predictions strategy for each 2011 and 2012.

Appendix C: Functional response curves for $\delta = 0$, $\gamma = 0$, $A = 64$ and $S = 1$

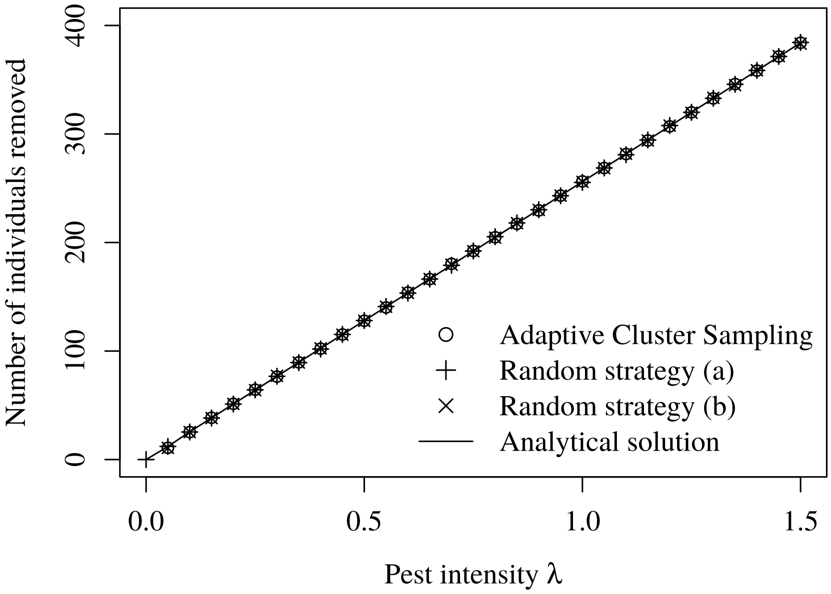

Figure C.1 – Functional response curves for fixed values of $\delta = 0$, $\gamma = 0$, $A = 64$ and $S = 1$ for the random strategy (a) and $A = 1$ and $S = 64$ for the random strategy (b) and the Adaptive Cluster Sampling strategy, and for a Poisson pest spatial distribution. Circles represent the means of 2000 simulations of the Adaptive Cluster Sampling strategy, pluses of the random strategy (a), and crosses of the random strategy (b). Solid lines represent the values for the analytical solution of random strategy (a) and (b).

Appendix D: Impact of parameters δ, γ, and A

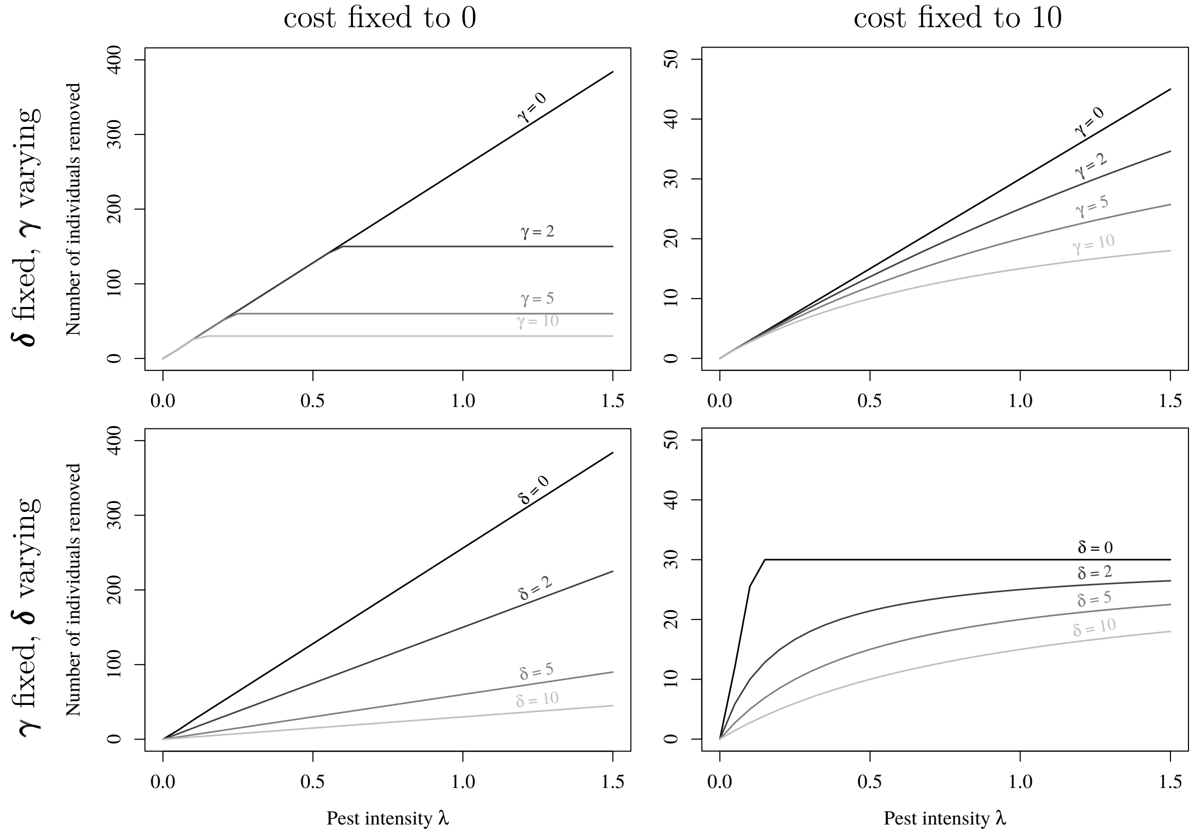

Figure D.1 – Functional response curves showing the impact of one of the cost parameters (the survey cost δ or the removal cost γ) when the other is fixed for the random strategies (a) and (b) on a Poisson pest spatial distribution. The shades of grey represent the values for the varying cost parameter from black (cost $= 0$) to light grey (cost $= 10$). The search area is set to 64 and the budget to 300.

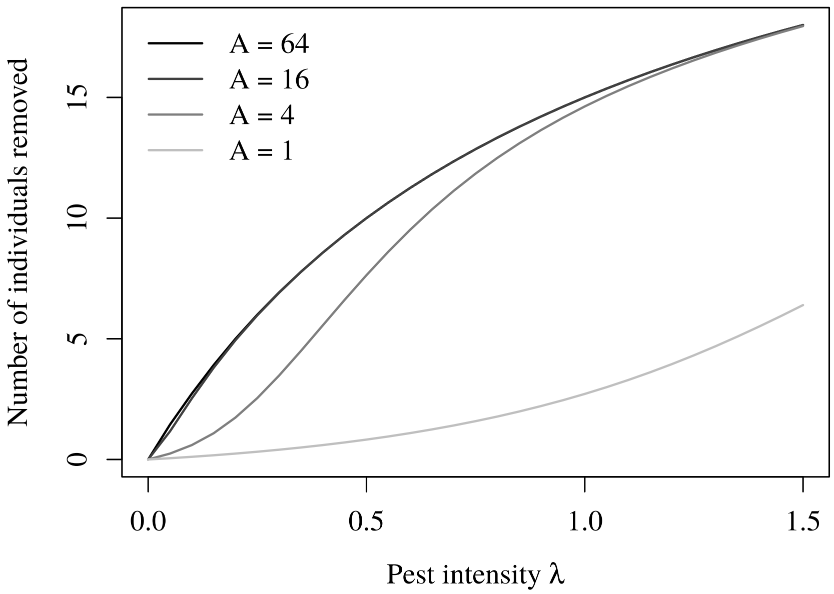

Figure D.2 – Functional response curves for fixed values of survey cost $\delta = 10$, removal cost $\gamma = 10$ and for the random strategies (a) and (b) on a Poisson pest spatial distribution. The shades of grey represent the values of the search area A: light grey $A = 1$ to black $A = 64$.

Appendix E: Comparison with Holling's disk equation

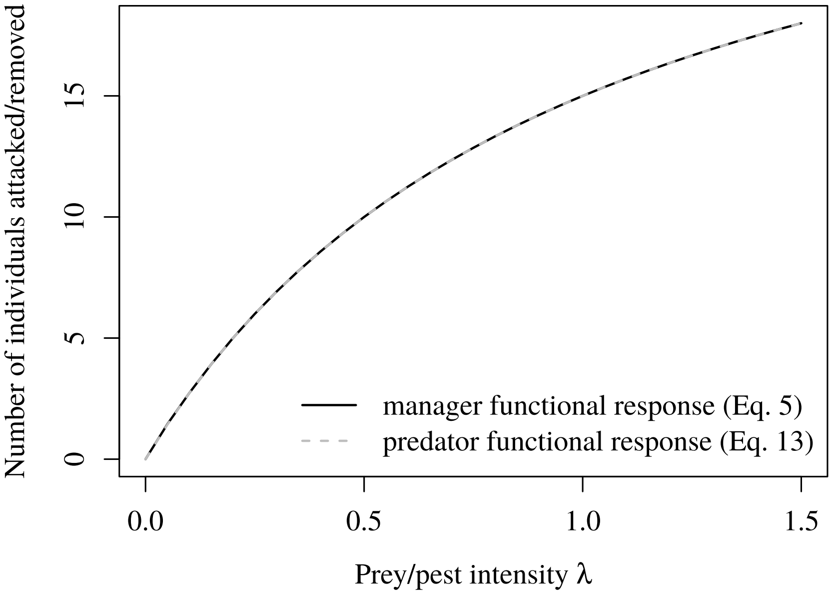

Figure E.1 – Functional response curves for fixed values of survey cost $\delta = 10$, removal cost $\gamma = 10$ and for a random strategy on a Poisson pest spatial distribution. The search area is set to 64 and the budget to 300. The black solid line represents the result of Eq. 3.5 and the grey dotted line represents the result of Eq. 3.13.

Appendix F: Functional response of the mountain pine beetle management in Cypress Hills, Saskatchewan

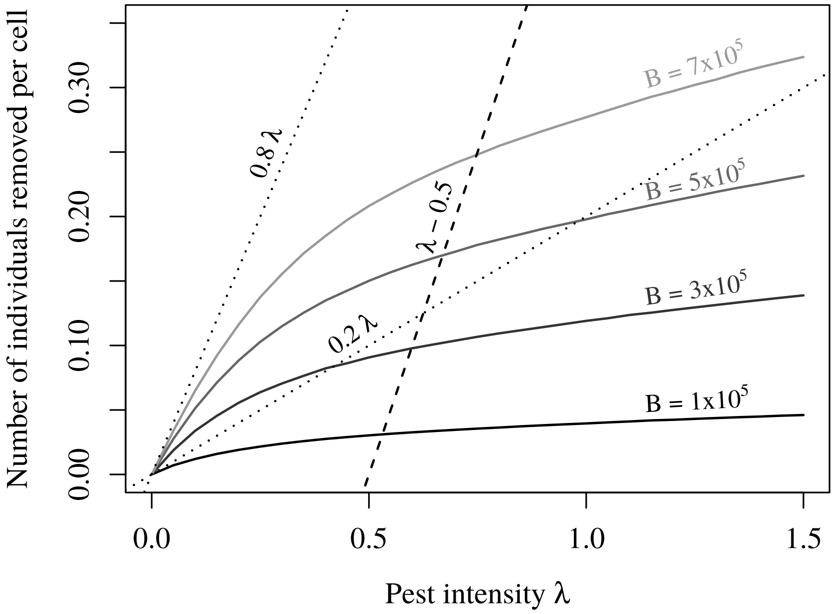

Figure F.1 – Functional response of the mountain pine beetle management in Cypress Hills, Saskatchewan using the adaptive cluster sampling process. Solid lines represent the mean of 1000 simulations of the management process for different budget values: $1{\times}10^5$, $3{\times}10^5$, $5{\times}10^5$, and $7{\times}10^5$. Dotted lines represent the management goal 1): removing 80% and 20% of all the individuals. The dashed line represent the management goal 2): removing all individuals above the threshold 0.5 individual per cell

Appendix G: True and false, positive and negative rates

A true positive (TP) is an observed presence that is also a predicted presence whereas a false negative (FN) is an observed presence that is also a predicted absence. A true negative (TN) is an observed absence that is also a predicted absence whereas a false positive (FP) is an observed absence that is also a predicted presence (Table G.1). A classifier gives as results the probability of having a presence for each observation; therefore, we determine the number of predicted presence or absence using a probability threshold. For example, setting the threshold to 0.5 would mean that if the probability of having MPB presence is higher than 0.5, we would consider it as a predicted presence and if it were below 0.5, we would consider it an absence. A different probability threshold would then give different number of true positives, false positives, true negatives and false negatives.

A receiver operating characteristic (ROC) curve (Metz, 1978) depicts, for a range of probability thresholds, the true positive rate (or 1 - false negative rate, also referred to as sensitivity or recall) against the false positive rate (also referred to as 1 - specificity). The true positive rate (TPR) is calculated following the equation: $\text{TPR} = \frac{\text{TP}}{\text{TP+FN}}$. The false positive rate (FPR) is calculated following the equation: $\text{FPR} = \frac{\text{FP}}{\text{FP+TN}}$.

A precision-recall curve (Raghavan *et al.*, 1989) depicts, for a range of probability thresholds, the proportion of true positives among the positive predictions (also referred to as precision or positive predictive value) against the true positive rate (sensitivity/recall). The precision is calculated following the equation: $\text{precision} = \frac{\text{TP}}{\text{TP+FP}}$.

Table G.1 – Confusion matrix displaying the number of observations by observed and predicted outcome.

		Observed	
		Absence	Presence
Predicted	Absence	TN	FN
	Presence	FP	TP