Bulletin of Mathematical Biology manuscript No.
(will be inserted by the editor)

The signature of endemic populations in the spread of mountain pine beetle outbreaks

Dean Koch · Mark A. Lewis · Subhash Lele

Received: date / Accepted: date

Abstract The mountain pine beetle (MPB) is among the most destructive eruptive forest pests in North America. A recent increase in the frequency and severity of outbreaks, combined with an eastward range expansion towards untouched boreal pine forests, has spurred a great interest by government, industry and academia into the population ecology of this tree-killing bark beetle. Modern approaches to studying

Dean Koch is postodoctoral fellow in the Department of Mathematical and Statistical Sciences, University of Alberta (UofA), Edmonton, Canada T6G 2R3 (email: dkoch@ualberta.ca, ORCID: https://orcid.org/0000-0002-8849-859X); Subhash R. Lele is Professor of Mathematical and Statistical Sciences at UofA (email: slele@ualberta.ca); Mark A. Lewis (MAL) is Professor of Mathematical and Statistical Sciences and Biological Sciences at UofA (email: mark.lewis@ualberta.ca, ORCID: https://orcid.org/0000-0002-7155-7426). The authors thank the Lewis Research Group for providing expert advice and feedback, as well as Victor Shegelski and the Sperling Lab for providing flight mill data. MAL is also grateful for support through NSERC and the Canada Research Chair Program.

Dean Koch, Mark A. Lewis, and Subhash Lele

Department of Mathematical and Statistical Sciences, University of Alberta, 11324 89 Ave NW, Edmonton, AB, T6G 2J5, Canada.

E-mail: dkoch@ualberta.ca, mlewis@ualberta.ca, slele@ualberta.ca

the spread of the MPB often involve the analysis of large-scale, high resolution datasets on landscape level damage to pine forests. This creates a need for new modelling tools to handle the unique challenges associated with large sample sizes and spatial effects. In two companion papers (Koch *et al.*, 2020, JRSI; and Koch *et al.*, 2020, EEST), we explain how the computational challenges of dispersal and spatial autocorrelation can be addressed using separable kernels. In this paper, we use these ideas to capture nonstationary patterns in the dispersal flights of MPB. This facilitates a landscape-level inference of subtle properties of MPB attack behaviour based on aerial surveys of killed pine. Using this model, we estimate the size of the cryptic endemic MPB population, which formerly has been measurable only by means of costly and time-intensive ground surveys.

Keywords mountain pine beetle · endemic · nonstationary · redistribution · kernel · modelling

Declarations

Funding: This research was supported by a grant to MAL from the Natural Science and Engineering Research Council of Canada (Grant No. NET GP 434810-12) to the TRIA Network.

Conflicts of interest/Competing interests: TRIA Network was supported by contributions from Alberta Agriculture and Forestry, Foothills Research Institute, Manitoba Conservation and Water Stewardship, Natural Resources Canada-Canadian Forest Service, Northwest Territories Environment and Natural Resources, Ontario Ministry

of Natural Resources and Forestry, Saskatchewan Ministry of Environment, West Fraser and Weyerhaeuser.

Ethics approval: Not applicable

Consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and material: All data used in this paper are available in a zip archive as electronic supplementary material (Online resource 2).

Code availability: The data analysis results reported in this paper can be reconstructed using the R script files included in the electronic supplementary material (Online resource 3).

Authors' contributions: All authors contributed to the model conception, design, and application. The first draft of the manuscript was written by Dean Koch and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

1 Introduction

The mountain pine beetle (MPB) *Dendoctonus ponderosae* Hopkins (Coleoptera Curculionidae), is a tree-killing species of bark beetle native to pine forests of Western North America. Each year for a short period in summer, adult MPB seek to complete their life cycle by attacking a suitable living host pine. During attacks, MPB bore into the bark, introducing fungal pathogens in the process, and ultimately girdle the tree (Taylor et al. 2006). Death follows swiftly for a pine whose defence systems fail to eject these attackers. When an attack succeeds, MPB use the host to feed and reproduce, laying eggs in galleries excavated underneath its bark.

A successful attack often leads to a MPB outbreak, in which local populations rise dramatically and large numbers of healthy pine are attacked over a period of several years. These outbreaks are major disturbance agents of pine forests, so mathematical models to explain their origin and how they spread across the landscape are of great interest to forest ecologists.

With few exceptions, the adult MPB die after reproduction, and their progeny emerge as teneral adults the following summer to begin a new life cycle. This semelparity, and the approximately linear relationship between reproductive success and host death, are mathematically convenient properties when constructing models to track year-to-year changes in MPB populations. For example if a total of B beetles have attacked a stand containing H susceptible pines, killing a fraction $\phi(B)$ of them, then a rough estimate of the MPB population emerging in the next year is $\lambda\phi(B)H$, where $\lambda > 0$ represents average per-stem productivity.

Modellers sometimes exploit this relationship to project outbreak dynamics ahead to future years (Heavilin and Powell 2008), and to explore ecological factors in population growth (Aukema et al. 2008). However, two major complications in MPB dynamics become apparent when attempting to link the host mortality fraction $\phi(B)$ with the underlying beetle population (Nelson et al. 2008). First, any spatially explicit model for $\phi(B)$ must account for dispersal flights, which allow localized outbreaks to spread into nearby areas. Second, $\phi(B)$ must reflect the eruptive and nonlinear nature of MPB population growth. We review these aspects briefly below, before introducing a new mathematical approach to the modelling problem.

1.1 Dispersal flights

In modelling the evolution of an outbreak over multiple years it is often convenient to track the beetle population in discrete time, where B_t is the value of B in year t. Such models cannot easily relate B_{t+1} and $\lambda\phi(B_t)H$ without incorporating dispersal. Flights of the MPB allow it to escape depleted stands, spark outbreaks in neighbouring areas, and expand its range (de la Giroday et al. 2012). By modelling B_t as the outcome of a spatially explicit dispersal event, we are better equipped to capture these interesting and important ecological phenomena, and achieve a higher precision in fitting $\phi(B)$ to data.

A variety of MPB dispersal models can be found in the literature (*e.g.* Goodsman et al. 2016; Preisler et al. 2012; Aukema et al. 2008; Heavilin and Powell 2008), but most make two simplifying assumptions out of mathematical convenience: that

movements occur in all directions with equal probability (isotropy); and that patterns of redistribution do not vary with spatial location (stationarity).

The main novelty in the methods presented here is that we drop both of these assumptions. Our dispersal model has the flexibility to capture directed and location-dependent (anisotropic and nonstationary) events. It is meant as a phenomenological alternative to dynamical systems based approaches to the same problem (*e.g.* Garlick et al. 2011; Powell and Bentz 2014; Powell et al. 2018), but with a simpler mathematical representation that borrows computationally efficient methods from spatial statistics. Our mathematical approach is based on ideas presented in two companion papers; on covariance structure (Koch et al. 2020a), and redistribution kernels (Koch et al. 2020b).

1.2 Colonization curves

The colonization curve, function $\phi(B)$, should be highly nonlinear to accommodate the distinct behaviours exhibited in different phases of MPB populations (Berryman 1978). During the incipient-epidemic phase, attacks occur at densities low enough to be defended by hosts, so cooperative efforts in overcoming these defences leads to a positive density dependence (Allee effect) in $\phi(B)$ (Boone et al. 2011). However, as the number of attacking individuals rises, the MPB enters epidemic and post-epidemic phases, in which the density dependence turns negative as a result of scramble competition (Woodell and Peters 1992).

Empirical data on $\phi(B)$ therefore reveals an S-shaped, or sigmoid relationship (Raffa and Berryman 1983). This form is reminiscent of the familiar type-III func-

tional responses for parasitism behaviour (Holling 1959), and indeed many aspects of MPB population dynamics are well described by this parasitoid-prey systems theory (Goodsman et al. 2016). In Section 2, we show how these functions can be adapted to model MPB population growth, generalizing the models of Heavilin and Powell (2008) and Koch et al. (2020b).

1.3 Aerial Overview Surveys

The extent to which a model can be complexified is limited by the amount and quality of data available for fitting and validation. Thus while the ideas outlined above lead to sophisticated population models, they also demand an unusually large spatial dataset for parameter inference. For this reason, we demonstrate our methods in an analysis of data from the Aerial Overview Survey (AOS) in British Columbia (BC).

The AOS maintains an annual record of the spatial patterns of insect damage to forests in BC. Operators fly in fixed-wing aircraft over most of the forested land in the province each summer, logging the locations of damage (and the presumed cause) as polygon and point data on maps, which are then digitized and published online.

These data are sometimes dismissed as too imprecise for detailed population modelling, since the process of visual observation and manual delineation of damaged areas is prone to human error (*e.g.* Kautz 2014; Wulder et al. 2006). Nevertheless, because the AOS covers such an impressively large extent and timeline of forest damage patterns in BC, a considerable body of landscape-level MPB research draws from the AOS and its predecessor, the Forest Insect and Disease Survey (*e.g.* Aukema

et al. 2006; Robertson et al. 2009; Chen and Walton 2011; Reyes et al. 2012; Sambaraju et al. 2012; Chen et al. 2015a,b).

1.4 Paper Outline

We show in Section 2 how a generalization of the Heavilin and Powell (2008) model allows us to relate data from the AOS with ground surveys of MPB activity. We use these ideas in Section 3 to demonstrate the remarkable amount of information that can be extracted – with the right modelling tools – from the AOS alone.

In particular our model accurately estimates the size of the cryptic, low-density endemic MPB population using only spatial data on outbreaks. This is remarkable given that pine mortality caused by the endemic phase happens at levels far below the operational detection threshold of the AOS (Cooke and Carroll 2017). Studies of endemic MPB more typically rely on intensive ground surveys of attacked pine (*e.g.* Boone et al. 2011; Bleiker et al. 2014). Our model estimates the rate of endemic attacks using only AOS data on outbreak-level pine mortality.

This is important because, in comparison to more reliable ground survey methods, aerial survey programs such as the AOS are a far less expensive and time-consuming means of monitoring MPB activity over large geographical areas. Note that similar datasets are available for the neighbouring province of Alberta (AB), in which a highly consequential MPB range expansion is currently underway.

Section 2.1 introduces the model by reviewing a popular mathematical representation for the colonization curve $\phi(B)$, before introducing several refinements in Sections 2.2-2.4. Our representation of dispersal flight is then introduced Section

2.5, and an error model suitable for the AOS dataset is proposed in Section 2.6. We demonstrate the model in Section 3 by fitting to data on outbreaks of the MPB in BC during the years 2006-2008.

2 Methods

Our case study covers a pine-rich region of roughly 10,000 km^2, centered over the Merritt Timber Supply Area (TSA) of Southern BC (Figure 1). We divided this into a 1 hectare (ha) resolution grid (*sensu* Aukema et al. 2006) to form a 1000 x 1000 lattice of cells, with matching layers provided by the province (http://www.hectaresbc.org) on wildfire, cutblocks, and topography.

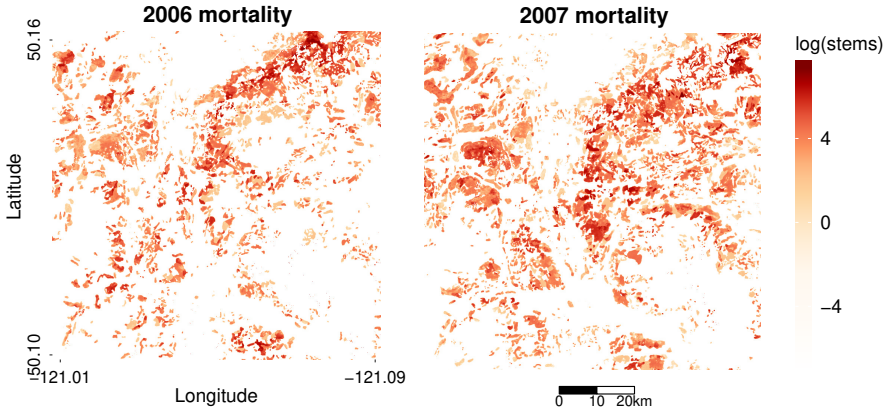

Fig. 1: Hosts killed by MPB ($\phi_{i,t}H_{i,t}$, in stems/ha) in the summers of 2006-2007. AOS data on damage severity were rasterized to approximate susceptible host mortality ($\phi_{i,t}$). Host density $H_{i,t}$ was derived from pine volume estimates in Beaudoin et al. (2014), as described in Appendix 1.1

171 As we are interested in how dispersal patterns are related to outbreak development,
172 we analysed the attack years 2006-2008, in which a large number of pine-leading stands
173 would see transitions from endemic to epidemic behaviour (the incipient-epidemic
174 population phase). This period captures the peak of an epidemic in the Merritt TSA
175 (in terms of basal area damaged) at a time when around one out of four cells in the
176 area exhibited crown-fade due to MPB activity.

177 Our analysis tracks four state variables, indexed by year (t) and location (i): Only
178 two of them are measured in practice: pine mortality ($\phi_{i,t}$) and host density ($H_{i,t}$,
179 in stems/ha) (Appendices 1.1–1.2); The others, MPB density *pre*-dispersal $\tilde{B}_{i,t}$ and
180 *post*-dispersal $B_{i,t}$ (in females/ha), are latent variables, inferred by the model but
181 never directly observed (Table 1).

location i	vectorized	definition	units	type
$H_{i,t}$	$\boldsymbol{H}_t$	pre-attack susceptible pine density	stems/ha	observed
$\phi_{i,t}$	$\boldsymbol{\phi}_t$	proportion of $H_{i,t}$ killed by MPB	unitless	
$\tilde{B}_{i,t}$	$\tilde{\boldsymbol{B}}_t$	emerging MPB density (pre-dispersal)	females/ha	latent
$B_{i,t}$	$\boldsymbol{B}_t$	MPB attack density (post-dispersal)		

Table 1: Notation for state variables in the MPB attack dynamics model. Indexing is by year t and location i, and boldface denotes the vector of all n locations, *e.g.* $\boldsymbol{\phi}_t = (\phi_{1,t}, \phi_{2,t}, \ldots, \phi_{n,t})'$.

182 In the remainder of Section 2 we construct a model connecting these four state
183 variables. We start in Section 2.1 by extending the red-top model of Heavilin and

Powell (2008), interpreting its attack parameters in a new light. We then link this attack model to novel submodels describing four important components of MPB population dynamics (Nelson et al. 2008): stand susceptibility, endemic populations, reproduction, and dispersal (Sections 2.2-2.5, respectively). Finally, in Section 2.6 we describe the data and statistical methodology used for fitting the full model, before presenting our results and connecting them to empirical findings from the MPB literature in Sections 3-4.

2.1 Attack dynamics (ϕ)

Our equation for pine mortality $\phi_{i,t}(B_{i,t})$ generalizes the red-top model of Heavilin and Powell (2008) to better match the types of colonization curves fitted in Cooke and Carroll (2017). The red-top model is best introduced by focusing at first on a particular location and year; so for notational convenience we omit the subscripts i and t until they are needed again in Section 2.4. Thus, for i and t fixed, we relate the attack density B (females/ha) to pine mortality ϕ by:

$$\text{proportion of } H \text{ killed} = \phi(B) = \frac{B^\kappa}{a^\kappa + B^\kappa} \quad \text{where } a > 0,\ \kappa > 0. \tag{1}$$

Parameter a is the half-saturation value, or attack density (in females/ha) at which 50% mortality occurs, and κ is a shape parameter controlling the density dependence. The special case $\kappa = 2$ recovers the red top model of Heavilin and Powell (2008) (after multiplying both sides by H). Other κ values reflect alternative regimes of density dependence. Larger κ and/or a values coincide with a stronger defensive response by

pines. When $\kappa \leq 1$, the Allee effect vanishes, reflecting compromised defences, as might occur, for example, during a drought.

Parameter estimation becomes simpler if Equation (1) can be made linear in its parameters. Observing that the odds-ratio of pine mortality $\phi/(1-\phi)$ is $(B/a)^\kappa$, we can take logarithms to get a linear equation on the logit-log scale:

$$\text{logit}(\phi) = -\kappa \log(a) + \kappa \log(B), \tag{2}$$

where $\log(B)$ is the logarithm of attack density, and $\log(a)$ the density (on the log scale) at which one half of susceptible hosts are expected to be colonized.

This also happens to be the mathematical form of the colonization curve fitted in Cooke and Carroll (2017) to the data reported in Boone et al. (2011) on attacked pines in our study area. Their analysis estimated $\hat{\kappa} = 1.66$ for the 2 years leading up to 2006. In years prior, a much lower value (0.56) was estimated, suggesting that environmental stressors on pine may have relaxed the Allee effect and bolstered endemic populations to spark the large-scale outbreaks of 2006-2008 (Figure 1).

Once started, outbreaks are not easily stopped. Irruptions in MPB populations are accompanied by behavioural changes in which host-preference switches from stressed to healthy pine (Carroll et al. 2006). This allows population growth to continue even after pine vigour recovers from a period of stress. Above a certain density threshold, B_T, the MPB have sufficient numbers to cooperatively attack a healthy pine (a mass attack), releasing them from the ordinary pressures of the Allee effect and marking the beginning of the incipient-epidemic phase.

The nature of this density dependence is reflected by the (equivalent) equations (1) and (2). The case $\kappa > 1$, corresponding to attacks on healthy pine, is illustrated

graphically in Figure 2. We discuss two features of this curve below, the incipient-epidemic transition point (B_T) and mass attack number (m_A). We report on their estimates in Section 3.

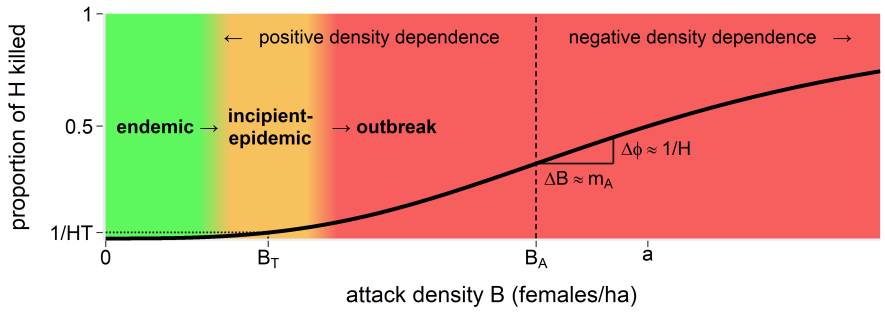

Fig. 2: Host mortality as a function (Equation 1) of MPB attack density B for $\kappa = 3$. Above the inflection point (B_A) is a regime of negative density dependence. When $B \approx 0$, the endemic population is too small to mass-attack healthy pine. When B rises to the incipient-epidemic transition point B_T, mass attacks become feasible and the MPB are released from the endemic phase. At moderate densities, each attacked pine accounts for $\approx m_A$ beetles. At higher densities, intraspecific competition leads to diminishing returns and negative density dependence

2.1.1 The incipient-epidemic transition (B_T)

Empirical data from our study area suggest that a density of $\hat{B}_T \approx$ 300-600 females/ha is sufficient to initiate the incipient-epidemic phase in an area of $T = 15.3$ ha (Carroll

et al. 2006). B_T is of course scale dependent; Amman (1984) estimated a quite different transition point at the $T = 40.5$ ha scale.

However, given any scale of interest, and given the values of κ and a, B_T can be estimated via equation (1). This is done by setting $\phi(B_T) = 1/HT$ (*i.e.* assuming a single host death in the specified area, T) and inverting (1) to get:

$$B_T = \frac{a}{\sqrt[\kappa]{HT - 1}}. \tag{3}$$

2.1.2 The mass attack number (m_A)

The number of MPB aggregating during attacks is carefully moderated by pheromones. This allows the beetle to optimize its reproductive success in mass attacks, by attacking in numbers high enough to overcome tree defenses, but low enough to avoid crowd competition (Taylor et al. 2006). For example, Raffa and Berryman (1983) reported an optimum of around 61 attackers/m^2, implying that a tree with 5.5 m^2 of bark available for attack (typical of the pine-leading stands in our study area) would have an optimum around 340 females/stem.

Under ideal conditions for MPB attack, this optimal density will presumably match the average attack density per attacked tree, which we call the *mass attack number* m_A (in females/stem). This average is approximated by the slope of $\phi(B)$ near its inflection point $B_A = a\sqrt[\kappa]{(\kappa - 1)/(\kappa + 1)}$ (where $\phi''(B_A) = 0$), since, at this intermediate density, $\phi(B)$ is nearly linear, and increases with B at rate $\phi'(B_A) \approx 1/(m_A H)$. From (1) we can therefore compute the approximation:

$$m_A \approx 1/(\phi'(B_A)H) = \frac{4B_A\kappa}{H(\kappa^2 - 1)}. \tag{4}$$

2.2 Stand susceptibility (a)

Although (3) and (4) both depend nonlinearly on κ, both equations scale linearly with the half-saturation value a. Stands that are highly susceptible to MPB attack have lower values (requiring fewer attacking beetles to initiate an outbreak), and vice versa. We can therefore interpret a as a simple measure of susceptibility to attack.

One can expect a to vary with environmental factors, such as weather, and stand characteristics, such as pine density. These factors have a complex and nuanced relationship with susceptibility (Preisler et al. 2012), and a clear biology-based model for this relationship is lacking. To avoid overcomplicating our model, we simply take the best linear approximation on the logit-log scale, writing $\boldsymbol{x}\boldsymbol{\beta} = -\kappa \log(a)$ for a set of unknown regression parameters $\boldsymbol{\beta} = (\beta_1, \ldots \beta_{n_\beta})$ and covariates $\boldsymbol{x} = (x_1, \ldots x_{n_\beta})$.

Thus in equation (2) for the mortality log-odds, $\text{logit}(\phi)$, we swap out the intercept term with a linear predictor. Similar regression models, such as in Aukema et al. (2008) and Preisler et al. (2012), have been useful for identifying environmental factors that have a significant ($\beta_k \neq 0$) effect on outbreak occurence. For our purposes $\boldsymbol{\beta}$ simply serves as a (location-wise) correction of a through which to estimate MPB population sizes, so we do not focus on the β_k or their effect sizes in our analysis. However, interested readers will find the full set of linear regression covariates listed in Appendix 1.2.

2.3 Endemic populations (ϵ)

The (aspatial) red-top model of Heavilin and Powell (2008) has no endemic equilibrium: low density populations are viewed as unstable, tending to extinction, and occuring only by means of immigrations from a reservoir of distant outbreaks appearing stochastically across the landscape. However, empirical data (*e.g.* Boone et al. 2011; Bleiker et al. 2014) suggest that resident endemic populations are widespread and persistent. These low-density populations subsist on a small number of defensively compromised pines and an assemblage of secondary bark beetle species that assist in the colonization of weakened trees.

We introduce a stable endemic equilibrium into the red-top model (Equation 1) by adding a small positive term $\epsilon > 0$ (in females/ha) to the post-dispersal MPB population (B) in the red-top model at all sites/years prior to attack. Specifically, with the addition of endemic beetles the attack function $\phi(B)$ becomes:

$$\text{proportion of } H \text{ killed} = \phi(B+\epsilon) = \frac{(B+\epsilon)^\kappa}{a^\kappa + (B+\epsilon)^\kappa} \text{ where } \epsilon > 0, \tag{5}$$

which, on the log-scale, produces $\text{logit}(\phi) = -\kappa \log(a) + \kappa \log(B+\epsilon)$.

This constant introduces a spatially uniform background level of MPB. Should an in-flight from a neighbouring outbreak occur, its density is added to the endemic cohort ϵ, and the combined population attacks pines according to Equation (1). The effect of ϵ is therefore to boost the effective size of spreading populations, increasing the likelihood that an incipient-epidemic transition will succeed in sparking a local outbreak.

In the absence of immigrating MPB, the endemic population is too small to attack healthy pines, so it instead seeks out defensively weakened trees. Because this pool of suitable hosts is ephemeral and extremely small compared to H, these endemic MPB incur a much higher flight-establishment mortality cost than do outbreaking populations: Taylor et al. (2006) estimates the generation mortality of endemic MPB at 97.5%. Assuming most of this loss can be attributed to the search flight, the rate of attack on defensively weakened hosts under this model would be $(1 - 0.975)\epsilon$ females/ha (or slightly above), with the healthy pine population variable H unaffected.

However, if an endemic population joins with a cohort of immigrating outbreak-level MPB, suitable hosts suddenly become abundant, and the flight-establishment losses should drop accordingly. The generation mortality in populations capable of mass-attacks is thought to lie in the range 80-98.6% (Taylor et al. 2006; Amman 1984). We assume that these losses mostly occur as a result of tree defences and crowd-competition. Unlike search flight losses, the latter are subsumed into $\phi(B)$ under the model (1). Therefore, we estimate the total number of attacking beetles at a given site as the sum of ϵ and any MPB (local or immigrant) originating from mass-attacked trees.

2.4 Reproduction (λ)

Reproduction connects subsequent years, so we must now make the dependence of our model variables on time and location explicit. In the red-top model, reproduction is summarized by $\tilde{B}_{i,t} = \lambda_{t-1}\phi_{i,t-1}H_{i,t-1}$. This expresses that $\tilde{B}_{i,t}$, the density of

(non-endemic) mature MPB emerging in year t at location i, is proportional to the number of mass-attacked stems in year $t-1$.

The productivity parameter λ_{t-1} specifies the average number of female MPB brood to emerge from each attacked tree in the year following an attack. This counts only those individuals that hatch, survive to maturity, and engage in search flights for new hosts. Note that by aggregating demographic information to the level of the tree (and year), we forego some precision. However this formulation simplifies the model considerably, summarizing in a single constant the many MPB within-tree growth and development processes that cannot be observed in aerial surveys (Berryman 1974).

Under this model, productivity λ_t is not identifiable from data on $\phi_{i,t}$ and $H_{i,t}$ without knowledge of $a_{i,t}$. So we instead fixed the value of $\lambda_t = \lambda$ in all years to a plug-in estimate of $\lambda = (2/3)(250) = 166.7$ (females/stem) suggested by empirical productivity data for epidemic phase MPB (Cole, W. and Amman 1969, Fig. 9), and assuming a 1:2 male-female sex ratio (Reid 1962). This productivity value is consistent with a 90% generation mortality rate, calculated using the brood production regression in (Safranyik 1988, eq. 14) on the mean diameters (Carroll et al. 2006) and heights (Safranyik and Linton 1991) of pine in our study area.

Although a time (and space) dependent λ would be more realistic, it would complicate the model considerably. We do however allow all other process model parameters to vary with time (*e.g.* ϵ_t, κ_t, $\boldsymbol{\beta}_t$, and the parameters of D_t), estimating them separately for each year in our analysis. Variations in productivity are therefore reflected in changing stand susceptibility $a_{i,t}$, which varies both spatially and temporally through $\boldsymbol{\beta}_t$ and the local covariates $\boldsymbol{x}_{i,t}$ (Table 2).

submodel	vector	contents	definition	units
attack	$\boldsymbol{\theta}_{\phi_t}$	κ_t	density dependence shape value	unitless
		λ	beetle production per attacked host	females/stem
		ϵ_t	emerging endemic MPB population level	females/ha
		$a_{i,t}$	half-saturation / susceptibility value	females/ha
		$\boldsymbol{\beta}_t$	linear regression coefficients for $a_{i,t}$	-
dispersal	$\boldsymbol{\theta}_{D_t}$	$\boldsymbol{\Delta}_{k,t}$	pWMY kernel: angle, shape and range	-
error	$\boldsymbol{\theta}_{V_t}$	σ_t^2	marginal variance	unitless
		$\boldsymbol{\rho}_t$	Gaussian autocorrelation range (x and y)	km

Table 2: Parameters of the generalized red-top model ($\boldsymbol{\theta}_t$), organized into categories of attack ($\boldsymbol{\theta}_{\phi_t}$), dispersal ($\boldsymbol{\theta}_{D_t}$), and error ($\boldsymbol{\theta}_{V_t}$). All except for λ are fitted to data separately by year (t). For dispersal, a 5-parameter product-WMY (pWMY) kernel (Appendix 2) is assigned to each of $m = 625$ data blocks, indexed by $k = 1, \ldots m$. A vector of 44 regression coefficients ($\boldsymbol{\beta}_t$) defines stand susceptibility through the linear model $\kappa_t \log(a_{i,t}) = \boldsymbol{x}_{i,t}\boldsymbol{\beta}_t$ for local covariates $\boldsymbol{x}_{i,t}$ (Appendix 1.3), where i indexes location.

2.5 Dispersal ($\tilde{B} \rightarrow B$)

Dispersal can be represented in population models using redistribution kernels (Neubert et al. 1995). These are functions, D_t, specifying a probability distribution for location following movement events. If the emerging MPB population $\tilde{B}_{t,i}$ is observed at n spatial locations, D_t specifies an $n \times n$ matrix ($\boldsymbol{D}_t$) whose i, j^{th} entry $[D_t]_{ij}$ is the expected proportion of the population $\tilde{B}_{t,j}$ that will move to cell i in the course of dispersal (Appendix 2). Thus, after adding the endemic MPB, the expected attack density is $\mathbb{E}\left(B_{i,t}\right) = \epsilon_t + \sum_j \left([D_t]_{ij}\tilde{B}_{t-1,j}\right)$. The equivalent matrix-vector equation

$\mathbb{E}(\boldsymbol{B}_t) = \epsilon_t \boldsymbol{I} + \boldsymbol{D}_t \tilde{\boldsymbol{B}}_t$ allows us to drop the cumbersome location indices (i), so we will use this simpified notation whenever possible.

The i^{th} entry of $\mathbb{E}(\boldsymbol{B}_t)$ estimates the density of attackers, sometimes called beetle pressure, at stand location i in year t. Beetle pressure is a common feature of MPB outbreak risk models (*e.g.* Wulder et al. 2006; Preisler et al. 2012), where it expresses proximity to infestations by a weighted sum of severity values or presence/absence indicators in a neighbourhood of the target stand. The weights in this calculation are provided by the kernel function D_t.

The choice of D_t therefore reflects assumptions about how MPB redistribute in search of new hosts. Ad-hoc assignments of weights to $[D_t]_{ij}$ often suffices in simpler predictive models (*e.g.* Kärvemo et al. 2014; Kunegel-Lion et al. 2019) but, whenever possible, it is desireable to choose a kernel derived from models of the physical flight process (Nelson et al. 2008).

Our flight model approximates the Whittle-Matérn-Yasuda (WMY) kernel family (Yasuda 1975), which describes diffusive movements through complex habitat (Koch et al. 2020b). Included in this family are a number of distinct isotropic kernels that have been advocated in previous studies of similar datasets (*e.g.* Turchin and Thoeny 1993; Heavilin and Powell 2008; Goodsman et al. 2016). Figure 3 (middle) is one example, arising from diffusion with constant settling.

We calculate the $[D_t]_{ij}$ values using pWMY kernels, which in addition to closely approximating the WMY, easily incorporate anisotropic (directed) movement patterns (Figure 3, right) as might be expected from the effect of local winds (Ainslie and Jackson 2011) and patchy habitat (Powell et al. 2018).

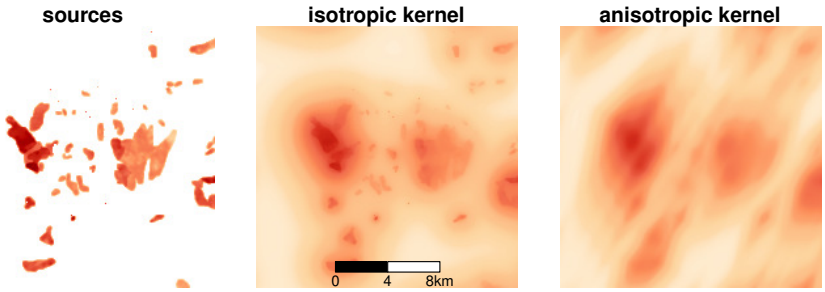

Fig. 3: MPB density pre-dispersal (left) and post-dispersal (middle and right) for two models of MPB flight patterns: an isotropic Bessel kernel (middle) with parameters from Goodsman et al. (2016), and an anisotropic pWMY kernel (right) parametrized to resemble the Bessel, but with the addition of a northeast-facing directionality

Writing $\boldsymbol{\delta}_{ij} = \left(x_i - x_j, y_i - y_j\right)'$ for the vector difference between the x-y coordinates at locations i and j, the equation of the anisotropic pWMY kernel is:

$$D_t\left(\boldsymbol{\delta}_{ij}; \alpha, \rho_x, \nu_x, \rho_y, \nu_y\right) = cD_W\left(d_{ij}^x; \rho_x, \nu_x\right) D_W\left(d_{ij}^y; \rho_y, \nu_y\right),$$
$$\text{where } \left(d_{ij}^x, d_{ij}^y\right)' = \boldsymbol{R}_\alpha \boldsymbol{\delta}_{ij}, \text{ and } D_W\left(d; \rho, \nu\right) = \left(d/\rho\right)^\nu K_\nu\left(d/\rho\right), \quad (6)$$

where K_ν denotes the ν^{th} order modified Bessel function of the second kind, $\boldsymbol{R}_\alpha$ is the standard 2D rotation matrix for angle α, and c is the kernel normalization constant. The parameters of this kernel are explained in detail in Appendix 2.

Importantly, the pWMY can be computed far more quickly than the WMY. Computational simplicity allows different dispersal patterns to be quickly fitted at different sites within a dataset. In our study area, this revealed a complex pattern of directionality (nonstationarity) that varies depending on the position of the source population.

Nonstationarity in dispersal patterns over a large geographic area is unsurprising in light of work by Powell and Bentz (2014), whose differential-equation based movement model connects environmental cues to direction and motility in MPB flights. Recognizing the importance of this nonstationarity, but lacking high-resolution data on its cues, we opted for a novel phenomenological model that combines multiple stationary (pWMY) kernels to form a nonstationary one.

We fitted each pWMY kernel to a relatively small square geographical area (a block) before combining them by computing a weighted average of their fitted values, with weights inversely related to distance from the block centroid (Figure 4). The effective contribution of each kernel to beetle pressure $\mathbb{E}(\boldsymbol{B}_t)$ is therefore restricted to a neighbourhood (dashed outer line in Figure 4, right) of the block over which it was fitted (solid line).

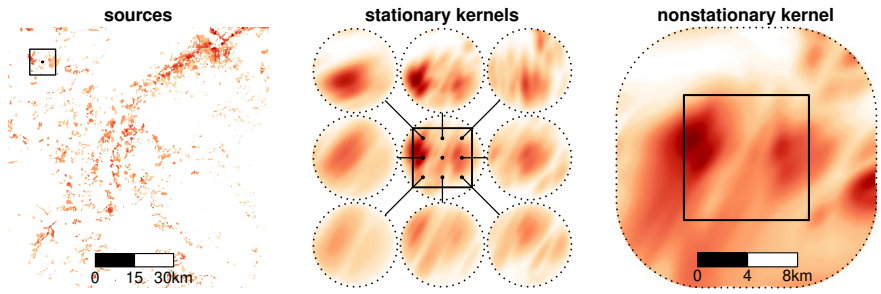

Fig. 4: A nonstationary flight pattern estimation scheme: stationary kernels are separately fitted to small overlapping blocks of data (at left, a block and its centroid). Expected beetle pressure (detail, at right) is computed as the distance-weighted average of nearby kernel predictions. The middle panel shows the nearest 9 block centroids and their kernel predictions before averaging

The resulting nonstationary dispersal model is itself a redistribution kernel, so we refer to it as D_t (with associated matrix $\boldsymbol{D}_t$). Its explicit mathematical form is derived in Appendix 2.

The virtue of this approach is that it captures complex (nonstationary) dispersal patterns by means of simpler stationary kernels, whose parameters can be fitted rapidly by well-established techniques over small neighbourhoods within which a stationarity assumption is reasonable. Moreover, there is no requirement for detailed environmental data, such as the stand density values used by Powell and Bentz (2014). Movement patterns are instead estimated directly from the available attack data.

Our construction of D_t used a total of 625 pWMY kernels in a 25×25 grid arrangement of blocks, each of size 10×10 km. Since each pWMY kernel captures only the local flight patterns within its respective block, we chose a distance-weighting function (Appendix 5.2) that assigns zero weight beyond the centroid-to-corner distance within a block (7.1 km). This scheme tracks movements up to 14.2 km, a reasonable upper bound on self-powered dispersal given laboratory studies suggesting fewer than 10% of MPB are capable of flight beyond this distance (Shegelski et al. 2019).

To avoid overparameterizing an already complicated model – and lacking data on wind patterns – we assumed that atmospherically-driven flight events (such as those documented by Jackson et al. 2008) were rare enough to ignore. Furthermore, although both block size and the number of blocks can be viewed as tuning parameters for the dispersal model, we assigned them ad-hoc values in this case to (roughly) coincide with the aforementioned self-powered dispersal limitations of MPB, rather than attempting to optimize them via model selection.

Some edge effects are unavoidable with this modelling strategy. For example, $[D_t]_{ij}$ values for a location coinciding with a block centroid will be determined almost entirely by the data within that single block, whereas for a location halfway between block centroids, the $[D_t]_{ij}$ values are influenced by data from two (or more) overlapping blocks – a much larger geographical extent. We believe, however, that this type of inconsistency pales in comparison to the roughcast assumption of stationary and isotropic dispersal patterns.

2.6 Model-fitting

2.6.1 Data

Pine density $\boldsymbol{H}_t$ was estimated using the model output of Beaudoin et al. (2014) for the year 2001, after adjusting for losses due to wildfire, logging, and pest damage incurred during the intervening years (Appendix 1.1). For simplicity we did not attempt to model regeneration, but rather assume that changes in density due to growth were small enough to ignore over the period 2001-2008.

Pine mortality data are drawn from the AOS of the Merritt TSA (Figure 1) for the attack years 2006-2008. These were rasterized by standard methods (Appendix 1.2) to produce a 1000×1000 grid of sample locations at a 1 ha resolution, matching the geometry of the pine density dataset. To avoid edge effects in dispersal calculations, we excluded a $\approx$10km buffer at the edge of this grid from the response data, forming the (logit-transformed) vector $\boldsymbol{\phi_t}$ from the subgrid of dimensions 893×893 centered on this region (a within-year sample size of 797,449 points).

2.6.2 Statistical model

A redistribution kernel is a probabilistic model – it connects MPB damage patterns to the *expected* density of attackers arriving next year at each location $\mathbb{E}(\boldsymbol{B}_t)$. Variations of $\boldsymbol{B}_t$ about this mean should therefore be modeled as error. Investigations into ecological dispersal by Preston (1948) and Limpert et al. (2001) inform us these errors are likely to be lognormally distributed. Assuming, $\left(\mathbb{E}(B_{i,t}) - B_{i,t}\right) \overset{iid}{\sim}$ lognormal$(0, \tilde{\sigma}_t^2)$, we can summarize Sections 2.1-2.5 in the equation:

$$\underbrace{\text{logit}\left(\boldsymbol{\phi}_t\right)}_{\text{pine mortality log-odds}} = \underbrace{\boldsymbol{X}_t\boldsymbol{\beta}_t}_{\text{susceptibility}} + \kappa_t \log(\underbrace{\epsilon_t \boldsymbol{I} + \lambda \boldsymbol{D}_t\left(\boldsymbol{\phi}_{t-1} \odot \boldsymbol{H}_{t-1}\right)}_{\text{beetle pressure}}) + \underbrace{\boldsymbol{Z}_t}_{\text{error}}, \quad (7)$$

where $\boldsymbol{X}_t = (\boldsymbol{x}'_{1,t}, \ldots \boldsymbol{x}'_{n,t})'$ is the (covariate) data matrix for year t, and $\boldsymbol{Z}_t$ is the vector of process errors arising from $\boldsymbol{B}_t$. The logit and log functions are applied elementwise, and the symbol $\odot$ denotes elementwise multiplication. This slight abuse of notation allows us to suppress the location indices i and write the complete model (7) in terms of length-n vector operations.

Under the lognormal assumption, $\boldsymbol{Z}_t$ is mean-zero multivariate normal (MVN), with a variance $\kappa_t \tilde{\sigma}_t^2$ that scales with the strength of the density dependence in $\phi(B)$. We assume that measurement error introduces an additional mean-zero MVN random vector appearing additively on the logit scale of (7). Since these errors are presumably independent of $\boldsymbol{B}_t$, their effect (by standard MVN theory) is to simply increase the variance of $\boldsymbol{Z}_t$. Thus, ignoring any autocorrelation (for now), we could write $\boldsymbol{Z}_t \sim \text{MVN}\left(\boldsymbol{0}, \sigma_t^2 \boldsymbol{I}\right)$, where σ_t^2 is the sum of the variances from process and measurement error.

For simplicity we ignored temporal autocorrelation by treating each year of data in the analysis as independent, as is commonly done in large-scale MPB outbreak analyses (*e.g.* Heavilin and Powell 2008; Goodsman et al. 2016). While this is not ideal, it avoids the difficulties associated with aligning consecutive years of raster data containing a large number of slight positional errors (Wulder et al. 2009), while simplifying the error model both mathematically and computationally.

Spatial autocorrelation, on the other hand, is more easily corrected using covariograms (Chilès and Delfiner 2012). For computational efficiency we used the Gaussian covariogram, which generates a covariance matrix $\boldsymbol{V}_t$ (to replace $\sigma_t^2 \boldsymbol{I}$ above) based on σ_t^2 and a pair of correlation range parameters, $\boldsymbol{\rho}_t$. In this model, the logarithm of the likelihood function for observations of $\boldsymbol{\phi}_t$, given $\boldsymbol{\phi}_{t-1}$ and $\boldsymbol{X}_t$ is proportional to:

$$\mathcal{L}\left(\boldsymbol{\theta}_t \mid \boldsymbol{Z}_t\right) = -\log\left(\det(\boldsymbol{V}_t)\right) - \boldsymbol{Z}_t' \boldsymbol{V}_t^{-1} \boldsymbol{Z}_t \quad \text{where } \boldsymbol{\theta}_t = \left(\boldsymbol{\theta}_{\phi_t}, \boldsymbol{\theta}_{D_t}, \boldsymbol{\theta}_{V_t}\right) \tag{8}$$

with $\boldsymbol{Z}_t$ as defined in (7), and model parameters $\boldsymbol{\theta}_t$ organized into components of attack dynamics, $\boldsymbol{\theta}_{\phi_t} = (\kappa_t, \lambda, \epsilon_t, \beta_t)$; error, $\boldsymbol{\theta}_{V_t} = (\sigma_t^2, \boldsymbol{\rho}_t)$, and dispersal $\boldsymbol{\theta}_{D_t} = (\boldsymbol{\Delta}_{1,t}, \ldots \boldsymbol{\Delta}_{625,t})$; as in Table 2. The model can now be fitted to data by maximum likelihood estimation (MLE), which finds the maximizer of (8), called $\hat{\boldsymbol{\theta}}_t = (\hat{\boldsymbol{\theta}}_{\phi_t}, \hat{\boldsymbol{\theta}}_{D_t}, \hat{\boldsymbol{\theta}}_{V_t})$.

Our estimation method for $\boldsymbol{\theta}_t$ is based on the 2-step algorithm described in Crujeiras and Van Keilegom (2010), but with a blockwise approach to approximating the large number of parameters in $\boldsymbol{\theta}_{D_t}$. Each of the 625 pWMY kernels is fitted indendently to the data in its block, before being combined to form the nonstationary kernel matrix $\hat{\boldsymbol{D}}_t$. By assuming $\boldsymbol{D}_t \approx \hat{\boldsymbol{D}}_t$, estimation of the remaining parameters $\boldsymbol{\theta}_{\phi_t}$ and $\boldsymbol{\theta}_{V_t}$ then becomes straightforward using generalized least squares (GLS) based

468 methods (Chilès and Delfiner 2012). Simulations indicated that our approach yields
469 unbiased and reasonably precise estimates of $\boldsymbol{\theta}_t$ (Appendix 3).

470 3 Results

471 The estimated endemic densities and attack curve shapes in all three years (Figure
472 5) matched closely with ground surveys of our study area during the period 2001-
473 2005. We estimated ϵ_t, the endemic contribution, at 388, 279, and 566 (females/ha),
474 respectively, for the years 2006 – 2008. Note that these densities are well above
475 what is considered normal for the endemic phase (Safranyik and Carroll 2006), as
476 they represent populations before flight-establishment loss. After correcting for this
477 loss (97.5%), our estimates suggest a range of 7 – 14 attackers/ha in endemic-only
478 populations, similar to the ranges reported in Boone et al. (2011) and Bleiker et al.
479 (2014).

480 A density dependence in attack was detected in all years, with κ estimated at 1.69,
481 1.32, and 1.67. Note that the estimates in 2006 and 2008 very nearly matched the
482 value of 1.66 reported by Cooke and Carroll (2017) for pooled colonization curve
483 data from the preceeding years 2002-2003 and 2005 (Figure 5). This indicates that
484 not only is density dependence detectable from stand-level AOS data (in the absence
485 of failed attack counts) – supporting the findings of Goodsman et al. (2016) on Allee
486 effects – but also that the precise shape of the attack curve in (9) can be estimated
487 from aerial data on $\boldsymbol{\phi}_t$ and $\boldsymbol{H}_t$ alone. This includes both the Allee and compensatory
488 (crowd competition) effects highlighted in Figure 2.

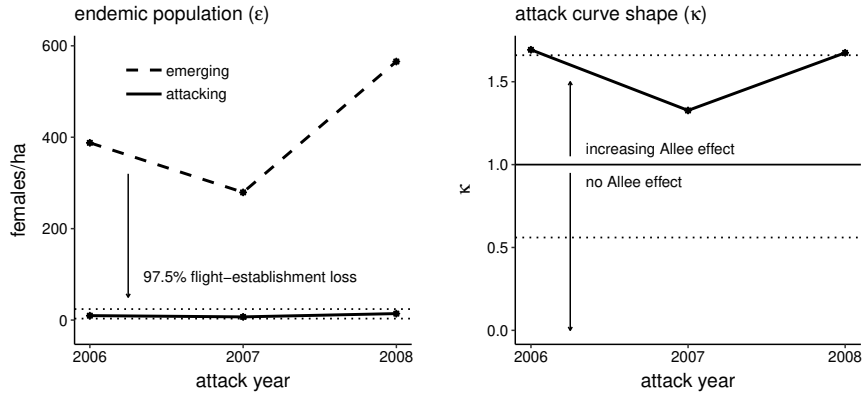

Fig. 5: Fitted attack parameters. At left, estimates of the endemic population and expected attack rates lying within the range (dotted lines) reported in Boone et al. (2011). At right, estimates of the attack curve shape compared with reference levels from Cooke and Carroll (2017) (dotted lines)

Estimates of stand susceptibility $a_{i,t}$ varied across the landscape, being spatially dependent on $\boldsymbol{x}_{i,t}$. Locations unsuitable to MPB (such as unforested areas) tended to assume extremely large $a_{i,t}$ values whereas areas with optimal habitat for MPB assumed much smaller ones.

Restricting our attention to optimal stands only – *i.e.* those having a density of 800-1500 stems/ha and aged > 80 yrs (Carroll et al. 2006), representing around 150,000 locations – the observed distribution of susceptibility values can be compared to empirical data from similar outbreaks. For example the modes of the estimated $m_{A_{i,t}}$ values over these optimal stands were centered at 336, 932, and 480 females/stem, for the years 2006-2008 respectively (Figure 6). This is reasonably consistent with

499 the 300-617 females/stem range observed in our study area by Safranyik and Linton
500 (1991) during a previous outbreak in 1984.

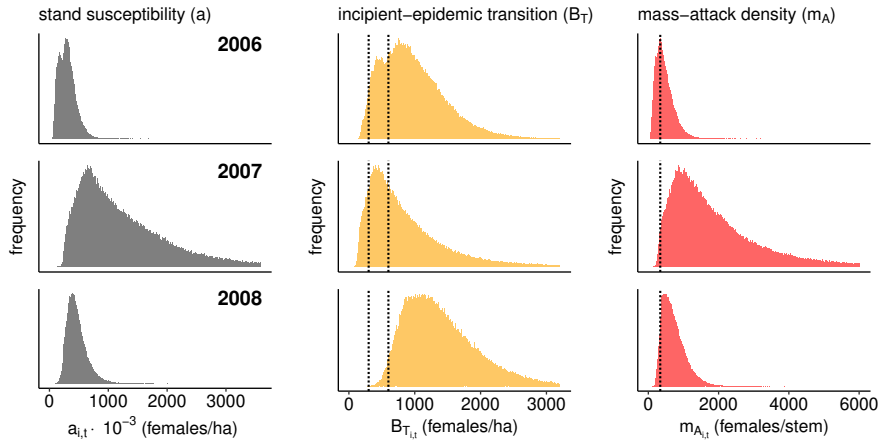

Fig. 6: Histograms of estimated susceptibility ($a_{i,t}$ left) in stands optimal for MPB in the years 2006-2008, and two associated quantities: (middle) the beetle pressure required for one mass attack per 15 ha, with dotted lines indicating an empirical range (Cooke and Carroll 2017); and the mass attack number (right), with a dotted line indicating the optimum of Raffa and Berryman (1983).

501 Using data on average diameters and attack heights for these optimal stands
502 (23cm, Carroll et al. 2006; and 11.36m, Safranyik and Linton 1991; respectively), we
503 estimated a typical bark area of 5.5 m^2/stem (Safranyik 1988, eq. 6). Our typical per-m^2
504 observed attack density ($m_{A_{i,t}}/5.5$) therefore lay in the range of 61-170 females/m^2.
505 Note that the lower end of this range (observed in 2006) coincides exactly with the
506 optimal attack density measured by Raffa and Berryman (1983) (Figure 6, right).

This shows that 2006 was a year of strong population growth for MPB, with a relatively low threshold for outbreak emergence (B_T), and mass attack numbers (m_A) centred at or near the optimum for brood production. Populations continued to expand through the next two years, with a large number of incipient epidemic transition events, followed by a collapse. Our model indicates that in optimal habitat, these events typically happened when MPB attack densities increased through the range 427-1114 of females/ha (the modes of the estimated $B_{T_{i,t}}$ by year; Figure 6, middle). This agrees with empirical observations by (Cooke and Carroll 2017) of a transition point in the 300-600 range during the five years leading up to 2006, and indicates that B_T values spiked as the epidemic neared collapse in 2008.

On dividing the $B_{T_{i,t}}$ values in Figure 6 by our estimates for ϵ_t, and taking medians, we find that a factor of $2.5 - 3.2$ increase in the endemic population was typically sufficient to initiate an outbreak. These findings support the observation of Carroll et al. (2006) that the incipient-epidemic transition point seems to occur at a level slightly above the density required to mass-attack a single pine. Our model expresses this quantity by the ratio $B_{T_{i,t}}/m_{A_{i,t}}$, whose median values (in optimal MPB habitat) were 2.2, 0.5, and 2.0 in the years 2006-2008, respectively.

Flight events under the fitted model are summarized by the block-wise redistribution kernel estimates. Our pWMY kernels identified a large number of highly directed (anisotropic) dispersal events in all years. The grid of fitted dispersal kernel parameters ($\hat{\boldsymbol{\theta}}_{D_{k,t}}$) that generate $\hat{D}_t$ (Figure 7) resembles a smooth vector field, raising some interesting questions as to the driving forces behind these patterns.

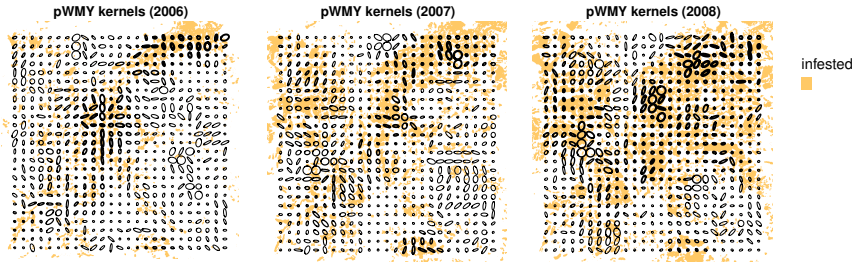

Fig. 7: Diffusion ellipses summarizing the angle and effective range corresponding to each of the 625 fitted pWMY parameter sets used to construct $\hat{\boldsymbol{D}}_t$ for each year. Each ellipse inscribes a contour of constant density for dispersal from its centre. Line thickness is scaled to match the estimated number of MPB displaced, emphasizing major outbreak centres. Infestations from the previous year are shaded to indicate the spatial distribution of source populations. Uninfested areas tended to produce small ellipses – these should be viewed as uninformative, as the model had no data from which estimate flight patterns in those blocks.

The combination of these stationary fitted kernels to form the nonstationary kernel ($\hat{D}_t$) brings into focus a complex landscape of MPB movement patterns (Figure 8), illustrating how detailed information on beetle pressure can be recovered from AOS data by rethinking the usual modelling assumptions about dispersal.

Note that our model was constructed for parameter inference, rather than predictions of future outbreak locations. However, our methodology for estimating beetle pressure could easily be adapted to serve a forecasting role. We illustrate the idea in Figure 9, where the empirical value of $B_T = 450$ (the midpoint of the range reported in Cooke and Carroll 2017) is used as threshold for outbreak development. The plot

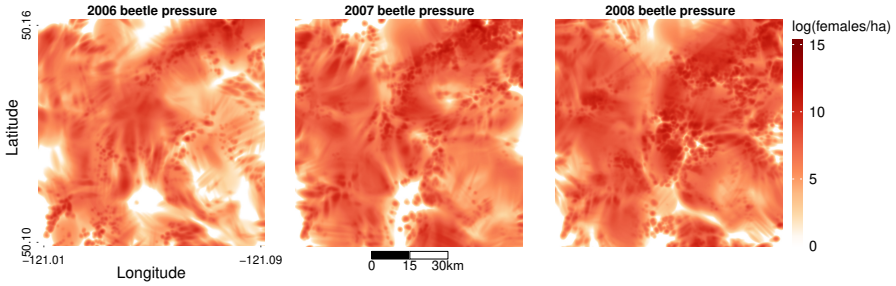

Fig. 8: Heatmaps of $\log(\lambda \hat{\boldsymbol{D}}_t\ (\boldsymbol{\phi}_{t-1} \odot \boldsymbol{H}_{t-1}))$, the fitted beetle pressure values arising from flight events in the years 2006-2008 (excluding endemic MPB). $\hat{\boldsymbol{D}}_t$ is the moving average of predictions from a 25×25 grid of local stationary models, each fitted to a local subset of the data

shows how our model delineates infested areas under two different scenarios; the first with no endemic population, and the second with ϵ_t set to its estimated value from 2006. Notice that neither $a_{i,t}$, κ_t nor $\boldsymbol{\theta}_{V_t}$ is needed for this classification.

The true positive rate in the training year 2006 was 93.5%, and in the forecast for 2007 it improved to 98.0%. By including the endemic population in our beetle pressure estimates, the contours of the infestation predictions broadened, sometimes by several kilometres. This improved detection rates substantially (true positive rate in 2006 without the endemic component: 71%; and in 2007: 84%).

4 Discussion

The S-shaped colonization curves that characterize the non-linearity of MPB attack dynamics (*e.g.* Raffa and Berryman 1983; Boone et al. 2011) are usually fitted to field data on individual attacked trees, so they relate attack density to the mortality among

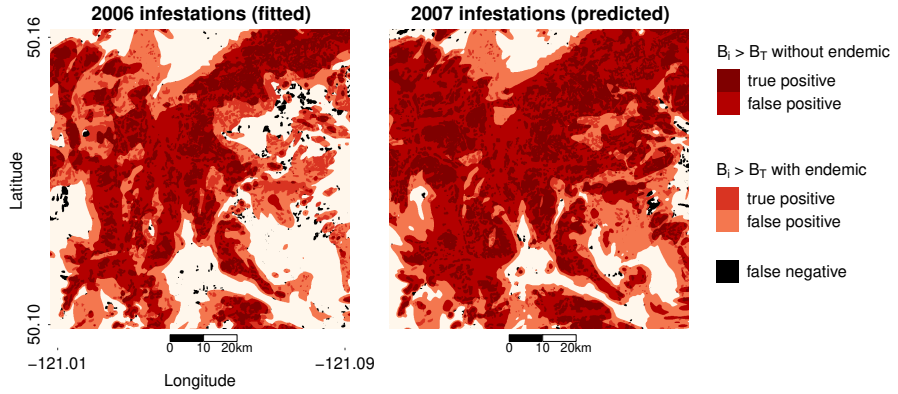

Fig. 9: Infested locations and next-year forecasts using the equation $\hat{\boldsymbol{B}}_t = \hat{\epsilon}\boldsymbol{I} + \hat{\boldsymbol{D}}_t\tilde{\boldsymbol{B}}_t$ from Section 2.5. Using the fitted values of $\hat{\epsilon}_t$ and $\hat{\boldsymbol{D}}_t$ from the training year 2006 (left), locations were classified as infested (shaded) if the predicted beetle pressure exceeded $B_T = 450$. Using these same parameters along with the observed attack damage and pine density in 2006, we then predicted infestations in 2007 (right). For comparison, an endemic-free estimate is also plotted (darker shaded regions) by replacing $B_{i,t}$ with $B_{i,t} - \epsilon_t$. The effect is to withdraw the contours of infested areas inward, considerably limiting the estimated spread.

pines undergoing attack. This is a conditional probability model. For example, the model of Cooke and Carroll (2017) has the form:

$$\text{logit}\left(\Pr(\text{pine mortality} \mid \text{attack})\right) = A + \kappa\log(N_a) = (A - \kappa\log(c)) + \kappa\log(B) \quad (9)$$

where A is an dimensionless intercept; and N_a is the number of stems attacked within the study plot, which we expect to scale according to $cN_a \approx B$ with the attack density B (in females/ha).

Our model, however, is based on aerial data, from which failed attacks cannot be resolved. In (2), we therefore related B to the *unconditional* probability of stand level mortality Pr(pine mortality | attack) Pr(attack), which we called ϕ. Notice that when Pr(attack) = 1, both the red-top model of Heavilin and Powell (2008) and our generalization (1) coincide exactly with (9). In reality, attack rates will be much lower, so in the high-level description (1) we assumed that the logit-linear relationship (9) remains after aggregating mortality data at the 1 hectare scale. Our results supported this assumption, with estimates of κ in close agreement with the field data reported by Cooke and Carroll (2017).

In Section 2.1 we showed how, via stand-susceptibility (a), this κ value is mathematically linked to the mass attack number (m_A) and the incipient-epidemic transition point (B_T). Our comparison of point estimates for these parameters with empirical data from previous years showed reasonably good agreement, supporting the theory behind formulae (3) and (4). This illustrates one way in which our model can be used to study ecological questions about MPB attack dynamics at the level of the individual tree, while using only (stand-level) AOS data for parameter fitting.

For instance, the observed increase in B_T in 2008, along with the elevated m_A levels in 2007-2008, can be attributed to host depletion, as MBP tend to select pine of a certain phloem, size, and vigour class for colonization (Shrimpton and Thomson 1985; Cole and McGregor 1983; Raffa and Berryman 1983). As the preferred hosts become scarce, MPB are thought to balance increasing fitness costs by first intensifying mass attacks on the few that remain (Lewis et al. 2010), thus effectively increasing m_A

above its optimal level. Similarly, a scarcity of suitable mass-attack targets can be expected to make spontaneous eruptions from the endemic phase less likely.

Furthermore, our results on ϵ_t shed a mathematical light on how outbreaks might sporadically arise across the landscape – if environmental conditions were to double or triple the number of injured/weakened pines available to the endemic population, this could allow it to grow to the point of exceeding B_T in the absence immigrating MPB – in accordance with the theory of Berryman (1978), and the explanation of Cooke and Carroll (2017) as to the origin of the outbreaks analysed in Section 3.

In-flights of MPB are equally important to understanding MPB outbreak dynamics. This is clear from the large number of spatial regression studies pointing to beetle pressure as the single most significant factor in outbreak development (*e.g.* Aukema et al. 2008; Preisler et al. 2012; Sambaraju et al. 2012). However there remains little consensus in the modelling literature on how best to represent beetle pressure mathematically.

As we explained in Section 2.5, beetle pressure simply expresses our modeling assumptions about MPB dispersal; Different modelling approaches handle this problem in different ways. With few exceptions (such as Powell and Bentz 2014; Powell et al. 2018) forecasting models tend to reconstruct beetle pressure in a heuristic way, by defining infestation indicator variables that are summed over local spatial neighbourhoods (see *e.g.* Shore et al. 2000; Aukema et al. 2008; Robertson et al. 2009; Kunegel-Lion et al. 2019). Many attack dynamics regression models also employ this trick (*e.g.* Zhu et al. 2010; Preisler et al. 2012; Sambaraju et al. 2012; Kärvemo et al. 2014), and indeed a stationary and isotropic kernel-based representation (as in

Heavilin and Powell 2008; Goodsman et al. 2016) is simply a refinement that finds a biology-based shape (and range) for the filter. Our method refined this idea further, in a novel way, by introducing directedness and location-dependence by means of a weighted combination of stationary kernels.

As we observed in a previous study (Koch et al. 2020b), the precision gained through the use of anistropic kernels appears to far outweigh the drawbacks associated with the introduction of additional dispersal parameters. Moreover, we believe our refined flight model shows promise not only in formulating beetle pressure (as we do here), but as a tool for studying nonstationary dispersal processes more generally. Future work might look for connections between $\boldsymbol{\theta}_{D_{k,t}}$ and environmental drivers such as prevailing wind direction, as a means of studying the dispersal process itself. For example, one could analyse whether patterns of directionality might arise in reaction to population density, both of beetles and hosts, similar to work by Powell and Bentz (2014) (we thank an anonymous reviewer for these suggestions).

Though we did not analyse the kernel parameters ($\boldsymbol{\theta}_{D_t}$) in detail, it is worth remarking that in most of the pWMY kernels a leptokurtic pattern of dispersal was favoured over the simpler Gaussian model of bio-diffusion. This highlights the versatility of the pWMY in modelling different flight mechanisms (Koch et al. 2020b), and suggests that a wide range of MPB flight behaviours are realized across the landscape: including both the fat-tailed patterns, proposed by Goodsman et al. (2016) and Turchin and Thoeny (1993); and the Gaussian, suggested by Heavilin and Powell (2008).

Note that the model-fitting procedure of Section 2.6 was constructed to study attack dynamics (at least) one year after they occur, not to predict them in future

summers, nor is our estimate of stand susceptibility $a_{i,t}$ (as a log-linear function of local covariates) intended for extrapolation. A more judicious choice of covariates whose values can be projected in time (combined with a significance-based covariate selection) would be needed in a predictive risk model for MPB damage. Nonetheless we think the framework in (7) – and in particular the nonstationary approach to dispersal – will be helpful in building model-based solutions to management and forecasting problems.

We illustrated this briefly in the next-year classification example of Figure 9. Note that our high detection rates lie near the level mentioned in Fettig et al. (2014) for stabilizing outbreaks by mitigation measures (such as cut and burn). However with high recall comes a high false positive rate (low precision); Moreover the 2007 prediction required information on pre-dispersal density that is typically not available until *after* the attack summer being predicted – recall that $\tilde{\boldsymbol{B}}_t$ is derived from crown fade data with a one-year lag. One possible solution would be to iterate equation (7) with simulated error to produce a suite of multi-year forecasts under various scenarios of stand susceptibility and process error, an idea we plan to explore in future work.

Figure 9 illustrates an important consequence of the ubiquity of endemic MPB in their natural range: it increases the potential for outbreaks to spread. The potential for range expansion may be therefore be underestimated if the endemic contribution to MPB outbreaks is ignored. This will be of particular relevance in contemporary areas of concern, such as the Boreal forest in Alberta (Safranyik et al. 2010). The establishment of endemic populations in these areas should be closely monitored, as

our results show that they have the potential to accelerate the spread of outbreaks, and thus speed the range expansion of the MPB.

References

Ainslie B, Jackson PL (2011) Investigation into mountain pine beetle above-canopy dispersion using weather radar and an atmospheric dispersion model. Aerobiologia (Bologna) 27(1):51–65, DOI 10.1007/s10453-010-9176-9

Amman GD (1984) Mountain Pine Beetle (Coleoptera: Scolytidae) Mortality in Three Types of Infestations. Environ Entomol 13(1):184–191, DOI 10.1093/ee/13.1.184

Aukema BH, Carroll AL, Zhu J, Raffa KF, Sickley TA, Taylor SW (2006) Landscape level analysis of mountain pine beetle in British Columbia, Canada: Spatiotemporal development and spatial synchrony within the present outbreak. Ecography (Cop) 29(3):427–441, DOI 10.1111/j.2006.0906-7590.04445.x

Aukema BH, Carroll AL, Zheng Y, Zhu J, Raffa KF, Dan Moore R, Stahl K, Taylor SW (2008) Movement of outbreak populations of mountain pine beetle: Influences of spatiotemporal patterns and climate. Ecography (Cop) 31(3):348–358, DOI 10.1111/j.0906-7590.2007.05453.x

Beaudoin A, Bernier PY, Guindon L, Villemaire P, Guo XJ, Stinson G, Bergeron T, Magnussen S, Hall RJ (2014) Mapping attributes of Canada's forests at moderate resolution through kNN and MODIS imagery. Can J For Res 44(5):521–532, DOI 10.1139/cjfr-2013-0401

Berryman AA (1974) Dynamics of Bark Beetle Populations: Towards a General Productivity Model 1. Environ Entomol 3(4):579–585, DOI 10.1093/ee/3.4.579

667 Berryman AA (1978) Towards a theory of insect epidemiology. Res Popul Ecol
668 (Kyoto) 19(2):181–196, DOI 10.1007/BF02518826

669 Bleiker KP, O'Brien MR, Smith GD, Carroll AL (2014) Characterisation of attacks
670 made by the mountain pine beetle (Coleoptera: Curculionidae) during its endemic
671 population phase. Can Entomol 146(3):271–284, DOI 10.4039/tce.2013.71

672 Boone CK, Aukema BH, Bohlmann J, Carroll AL, Raffa KF (2011) Efficacy of tree
673 defense physiology varies with bark beetle population density: A basis for positive
674 feedback in eruptive species. Can J For Res 41(6):1174–1188, DOI 10.1139/x11-
675 041

676 Carroll AL, Aukema BH, Raffa KF, Linton DA, Smith GD, Lindgren BS (2006)
677 Mountain pine beetle outbreak development: the endemic - incipient epidemic
678 transition. Can For Serv Mt Pine Beetle Initiat Proj 1:27

679 Chen H, Walton A (2011) Mountain pine beetle dispersal: spatiotemporal patterns
680 and role in the spread and expansion of the present outbreak. Ecosphere 2(6):art66,
681 DOI 10.1890/es10-00172.1

682 Chen H, Jackson PL, Ott PK, Spittlehouse DL (2015a) A spatiotemporal pattern
683 analysis of potential mountain pine beetle emergence in British Columbia, Canada.
684 For Ecol Manage 337(4):11–19, DOI 10.1016/j.foreco.2014.10.034

685 Chen H, Jackson PL, Ott PK, Spittlehouse DL (2015b) A spatiotemporal pat-
686 tern analysis of potential mountain pine beetle emergence in British Columbia,
687 Canada. For Ecol Manage 337:11–19, DOI 10.1016/j.foreco.2014.10.034, URL
688 http://dx.doi.org/10.1016/j.foreco.2014.10.034

689 Chilès JP, Delfiner P (2012) Geostatistics: Modeling Spatial Uncertainty: Second
690 Edition, vol 2. John Wiley & Sons, Hoboken, NJ, DOI 10.1002/9781118136188
691 Cole WE, McGregor MD (1983) Estimating the rate and amount of tree loss
692 from mountain pine beetle infestations. Tech. Rep. INT-318, US Department
693 of Agriculture, Forest Service, Intermountain Forest and Range˜. . . , DOI
694 10.5962/bhl.title.68709
695 Cole, W and Amman G (1969) Mountain pine beetle infestations in relation to lodge-
696 pole pine diameters, vol 95. US Dept. of Agriculture, Forest Service, Intermountain
697 Forest & Range
698 Cooke BJ, Carroll AL (2017) Predicting the risk of mountain pine beetle spread to
699 eastern pine forests: Considering uncertainty in uncertain times. For Ecol Manage
700 396:11–25, DOI 10.1016/j.foreco.2017.04.008
701 Crujeiras RM, Van Keilegom I (2010) Least squares estimation of nonlinear spatial
702 trends. Comput Stat Data Anal 54(2):452–465, DOI 10.1016/j.csda.2009.09.014
703 Fettig CJ, Gibson KE, Munson AS, Negrón JF (2014) A comment on "Management for
704 mountain pine beetle outbreak suppression: Does relevant science support current
705 policy?". Forests 5(4):822–826, DOI 10.3390/f5040822
706 Garlick MJ, Powell JA, Hooten MB, McFarlane LR (2011) Homogenization of Large-
707 Scale Movement Models in Ecology. Bull Math Biol 73(9):2088–2108, DOI
708 10.1007/s11538-010-9612-6
709 de la Giroday HMC, Carroll AL, Aukema BH (2012) Breach of the northern Rocky
710 Mountain geoclimatic barrier: Initiation of range expansion by the mountain pine
711 beetle. J Biogeogr 39(6):1112–1123, DOI 10.1111/j.1365-2699.2011.02673.x

712 Goodsman DW, Koch D, Whitehouse C, Evenden ML, Cooke BJ, Lewis MA (2016)
713 Aggregation and a strong Allee effect in a cooperative outbreak insect. Ecol Appl
714 26(8):2621–2634, DOI 10.1002/eap.1404
715 Heavilin J, Powell J (2008) A novel method of fitting spatio-temporal models to data,
716 with applications to the dynamics of mountain pine beetles. Nat Resour Model
717 21(4):489–524, DOI 10.1111/j.1939-7445.2008.00021.x
718 Holling CS (1959) Some Characteristics of Simple Types of Predation and Parasitism.
719 Can Entomol 91(7):385–398, DOI 10.4039/Ent91385-7
720 Jackson PL, Straussfogel D, Lindgren BS, Mitchell S, Murphy B (2008) Radar ob-
721 servation and aerial capture of mountain pine beetle, Dendroctonus ponderosae
722 Hopk. (Coleoptera: Scolytidae) in flight above the forest canopy. Can J For Res
723 38(8):2313–2327, DOI 10.1139/X08-066
724 Kärvemo S, Van Boeckel TP, Gilbert M, Grégoire JC, Schroeder M (2014) Large-scale
725 risk mapping of an eruptive bark beetle - Importance of forest susceptibility and
726 beetle pressure. For Ecol Manage 318:158–166, DOI 10.1016/j.foreco.2014.01.025
727 Kautz M (2014) On correcting the time-lag bias in aerial-surveyed bark beetle in-
728 festation data. For Ecol Manage 326:157–162, DOI 10.1016/j.foreco.2014.04.010,
729 URL http://dx.doi.org/10.1016/j.foreco.2014.04.010
730 Koch D, Lele S, Lewis M (2020a) Computationally Simple Anisotropic Lattice Co-
731 variograms. Environ Ecol Stat DOI 10.1007/s10651-020-00456-2
732 Koch DC, Lewis MA, Lele SR (2020b) A unifying theory for two-dimensional spatial
733 redistribution kernels with applications in population spread modelling. J R Soc
734 Interface 17(170):20200434

735 Kunegel-Lion M, McIntosh RL, Lewis MA (2019) Management assessment of moun-
736 tain pine beetle infestation in cypress hills, SK. Can J For Res 49(2):154–163,
737 DOI 10.1139/cjfr-2018-0301
738 Lewis MA, Nelson W, Xu C (2010) A structured threshold model for mountain pine
739 beetle outbreak. Bull Math Biol 72(3):565–589, DOI 10.1007/s11538-009-9461-3
740 Limpert E, Stahel WA, Abbt M (2001) Log-normal distributions across the sciences:
741 Keys and clues on the charms of statistics, and how mechanical models resem-
742 bling gambling machines offer a link to a handy way to characterize log-normal
743 distributions, which can provide deeper insight into va. Bioscience 51(5):341–352
744 Nelson WA, Potapov A, Lewis MA, Hundsdörfer AE, He F (2008) Balancing eco-
745 logical complexity in predictive models: A reassessment of risk models in the
746 mountain pine beetle system. J Appl Ecol 45(1):248–257, DOI 10.1111/j.1365-
747 2664.2007.01374.x, URL http://doi.wiley.com/10.1111/j.1365-2664.2007.01374.x
748 Neubert MG, Kot M, Lewis MA (1995) Dispersal and pattern formation in
749 a discrete-time predator-prey model. Theor Popul Biol 48(1):7–43, DOI
750 10.1006/tpbi.1995.1020
751 Powell JA, Bentz BJ (2014) Phenology and density-dependent dispersal predict pat-
752 terns of mountain pine beetle (Dendroctonus ponderosae) impact. Ecol Modell
753 273:173–185, DOI 10.1016/j.ecolmodel.2013.10.034
754 Powell JA, Garlick MJ, Bentz BJ, Friedenberg N (2018) Differential dispersal and the
755 Allee effect create power-law behaviour: Distribution of spot infestations during
756 mountain pine beetle outbreaks. J Anim Ecol 87(1):73–86, DOI 10.1111/1365-
757 2656.12700

Preisler HK, Hicke JA, Ager AA, Hayes JL (2012) Climate and weather influences on spatial temporal patterns of mountain pine beetle populations in Washington and Oregon. Ecology 93(11):2421–2434, DOI 10.1890/11-1412.1

Preston FW (1948) The Commonness, And Rarity, of Species. DOI 10.2307/1930989

Raffa KF, Berryman AA (1983) The Role of Host Plant Resistance in the Colonization Behavior and Ecology of Bark Beetles (Coleoptera: Scolytidae). Ecol Monogr 53(1):27–49, DOI 10.2307/1942586

Reid RW (1962) Biology of the Mountain Pine Beetle, Dendroctonus monticolae Hopkins, in the East Kootenay Region of British Columbia I. Life Cycle, Brood Development, and Flight Periods1. Can Entomol 94(5):531–538, DOI 10.4039/Ent94531-5

Reyes PE, Zhu J, Aukema BH (2012) Selection of Spatial-Temporal Lattice Models: Assessing the Impact of Climate Conditions on a Mountain Pine Beetle Outbreak. J Agric Biol Environ Stat 17(3):508–525, DOI 10.1007/s13253-012-0103-0

Robertson C, Farmer CJ, Nelson TA, MacKenzie IK, Wulder MA, White JC (2009) Determination of the compositional change (1999-2006) in the pine forests of British Columbia due to mountain pine beetle infestation. Environ Monit Assess 158(1-4):593–608, DOI 10.1007/s10661-008-0607-9

Safranyik L (1988) Estimating attack and brood totals and densities of the mountain pine beetle in individual lodgepole pine trees. Can Entomol 120(4):323–331, DOI 10.4039/Ent120323-4

Safranyik L, Carroll AL (2006) The biology and epidemiology of the mountain pine beetle in lodgepole pine forests. Mt Pine Beetle – A Synth Biol Manag Impacts Lodg Pine 11:3–66, DOI 10.1673/031.011.12701, URL

http://www.ncbi.nlm.nih.gov/pubmed/22233385

Safranyik L, Linton DA (1991) Unseasonably low fall and winter temperatures affecting mountain pine beetle and pine engraver beetle populations and damage in the British Columbia Chilcotin region. J - Entomol Soc Br Columbia 88:17–21

Safranyik L, Carroll AL, Régnière J, Langor DW, Riel WG, Shore TL, Peter B, Cooke BJ, Nealis VG, Taylor SW (2010) Potential for range expansion of mountain pine beetle into the boreal forest of North America. Can Entomol 142(5):415–442, DOI 10.4039/n08-CPA01

Sambaraju KR, Carroll AL, Zhu J, Stahl K, Moore RD, Aukema BH (2012) Climate change could alter the distribution of mountain pine beetle outbreaks in western Canada. Ecography (Cop) 35(3):211–223, DOI 10.1111/j.1600-0587.2011.06847.x

Shegelski VA, Evenden ML, Sperling FA (2019) Morphological variation associated with dispersal capacity in a tree-killing bark beetle Dendroctonus ponderosae Hopkins. Agric For Entomol 21(1):79–87, DOI 10.1111/afe.12305

Shore TL, Safranyik L, Lemieux JP (2000) Susceptibility of lodgepole pine stands to the mountain pine beetle: Testing of a rating system. Can J For Res 30(1):44–49, DOI 10.1139/x99-182

Shrimpton DM, Thomson AJ (1985) Relationship between phloem thickness and lodgepole pine growth characteristics. Can J For Res 15(5):1004–1008, DOI 10.1139/x85-161

Taylor SW, Carroll AL, Alfaro RI, Safranyik L, Safranyik L, Wilson B (2006) The mountain pine beetle: A synthesis of biology. Canadian Forest Service, Victoria, BC, CA

804 Turchin P, Thoeny WT (1993) Quantifying dispersal of southern pine beetles with
805 mark-recapture experiments and a diffusion model. Ecol Appl 3(1):187–198, DOI
806 10.2307/1941801
807 Woodell SRJ, Peters RH (1992) A Critique for Ecology., vol 80. Cambridge University
808 Press, Cambridge, England, DOI 10.2307/2261026
809 Wulder MA, Dymond CC, White JC, Leckie DG, Carroll AL (2006) Surveying
810 mountain pine beetle damage of forests: A review of remote sensing opportunities.
811 For Ecol Manage 221(1-3):27–41, DOI 10.1016/j.foreco.2005.09.021
812 Wulder MA, White JC, Grills D, Nelson T, Coops NC, Ebata T (2009) Aerial Overview
813 Survey of the Mountain Pine Beetle Epidemic in British Columbia: Communication
814 of Impacts. BC J Ecosyst Manag 10(1):45–58
815 Yasuda N (1975) The random walk model of human migration. Theor Popul Biol
816 7(2):156–167, DOI 10.1016/0040-5809(75)90011-8
817 Zhu J, Huang HC, Reyes PE (2010) On selection of spatial linear models for lat-
818 tice data. J R Stat Soc Ser B Stat Methodol 72(3):389–402, DOI 10.1111/j.1467-
819 9868.2010.00739.x