

**Risk rating for mountain pine beetle infestation of lodgepole pine forests  
over large areas with ordinal regression modelling**

Robertson, C.\*<sup>1</sup>, Wulder, M.A.<sup>2</sup>, Nelson, T.A.<sup>1</sup> and White, J.C.<sup>2</sup>

\*corresponding author:

Colin Robertson

Spatial Pattern Analysis & Research (SPAR) Laboratory, Dept of Geography, University  
of Victoria, PO Box 3050, Victoria, BC V8W 3P5, Canada

Email: colinr23@gmail.com, Ph. 1-250-885-8065

1) Spatial Pattern Analysis & Research (SPAR) Laboratory, Dept of Geography,  
University of Victoria, PO Box 3050, Victoria, BC V8W 3P5, Canada

2) Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, 506  
West Burnside, Victoria, BC V8Z 1M5, Canada

Submitted to: Forest Ecology and Management

Submission date: Feb 2008

# 1 **Abstract**

2 The mountain pine beetle *Dendroctonus ponderosae* Hopkins is endemic to lodgepole  
3 pine *Pinus contorta* var. *latifolia* Engelmann forests in western Canada. However, the  
4 current beetle epidemic in this area highlights the challenges faced by forest managers  
5 tasked with prioritizing stands for mitigation activities such as salvage harvesting,  
6 thinning, and direct control methods. In western Canada, the operational risk rating  
7 system for mountain pine beetle is based on biological knowledge gained from a rich  
8 legacy of stand-scale field studies. Due to the large spatial and temporal extents of the  
9 current epidemic, new research into large-area mountain pine beetle processes has  
10 revealed further insights into the landscape-scale characteristics of beetle infested forests.  
11 In this research, we evaluate the potential for this new knowledge to augment an  
12 established system for rating the short-term risk of tree mortality in a stand due to  
13 mountain pine beetle. New variables explored for utility in risk rating include direct  
14 shortwave radiation, site index, diameter at breast height, the temporal trends in local  
15 beetle populations, Biogeoclimatic Ecosystem Classification and beetle-host interaction  
16 variables. Proportional odds ordinal regression was used to develop a model for the  
17 Vanderhoof Forest District in west-central British Columbia. Prediction on independent  
18 data was assessed with the area under the receiver operator curve (AUC), indicating good  
19 discriminatory power (AUC = 0.84) for predicting damage due to mountain pine beetle.

20

21 Key words: mountain pine beetle, risk rating, landscape scale, monitoring, infestation

# 1 **Introduction**

2 Management of lodgepole pine *Pinus contorta* var. *latifolia* Engelmann forests in western  
3 Canada in recent years has been dominated by a large scale mountain pine beetle  
4 *Dedroctonus ponderosae* Hopkins epidemic that has impacted more than 9 million ha  
5 (Westfall, 2007). Management of any forest disturbance requires decision support tools  
6 that enable managers to predict future forest scenarios, set priorities, and evaluate  
7 management strategies. In the context of mountain pine beetle, forest managers must  
8 know the ability of a forest stand to support an epidemic mountain pine beetle population  
9 (i.e., susceptibility), and the possibility of host tree mortality as a result of an existing  
10 beetle infestation (i.e., risk). Susceptibility is determined using stand and site  
11 characteristics, independent of surrounding beetle population levels. Conversely, risk is  
12 determined by considering the susceptibility of the stand in the context of the local beetle  
13 population within the stand and in the vicinity of the stand (Bentz et al., 1993). A risk  
14 rating system is a specific decision support tool that is used to identify those forest stands  
15 on the landscape that are at greatest risk of timber losses as a result of a mountain pine  
16 beetle infestation (Shore et al. 2006).

17

## 18 ***Mountain pine beetle risk rating***

19

20 Many risk rating systems for mountain pine beetle have been developed over the past  
21 three decades. Safranyik et al. (1975) initially used weather station data to model and  
22 map the beetle outbreak hazard in western Canada. Amman et al. (1977) used stand  
23 characteristics such as elevation, age, and diameter at breast height (dbh) to develop a  
24 three-class risk classification system (i.e., low, moderate, high). Other risk rating systems

1 were developed that also relied on stand characteristics and adopted similar approaches to  
2 rating stand risk as a categorical variable (e.g., Mahoney 1978; Berryman 1978a; Stuart  
3 1984; Anhold and Jenkins 1987). Bentz et al. (1993) evaluated the accuracy of three  
4 categorical risk rating systems (Amman et al. 1977; Mahoney 1978; Berryman 1978a)  
5 and one continuous variable risk rating system (Schenk et al. 1980). All of the systems  
6 evaluated by Bentz et al. (1993) were found to provide poor estimates of pine mortality,  
7 primarily because they failed to consider spatial relationships between host stands and  
8 beetle populations. Furthermore, Bentz et al. (1993) concluded that the empirical  
9 development of these risk rating systems limits their portability to other geographic areas.  
10  
11 Shore and Safranyik (1992) introduced a continuous variable risk rating system which  
12 incorporated those elements of previous systems that had a strong theoretical basis for  
13 characterizing susceptibility and risk. Shore and Safranyik (1992) define risk as the short-  
14 term expectation of volume loss due to mountain pine beetle attack. There are two  
15 components to the Shore and Safranyik risk rating system: stand susceptibility, defined as  
16 the inherent characteristics of a stand of trees that affect its likelihood of attack by  
17 mountain pine beetle; and, beetle pressure, which is a measure of the size and proximity  
18 of the mountain pine beetle population to the stand. Susceptibility in Shore and Safranyik  
19 (1992) is determined using stand density, age, composition, and geographic location.  
20 Each of these variables has a direct link to biological processes associated with mountain  
21 pine beetle. Stand density effects tree competition for light and nutrients, so less dense  
22 stands tend to have larger, more vigorous trees; however, low density stands have a  
23 negative impact on the microclimate that is required to facilitate pheromone mediated

1 attacks, landing, and emergence rates (Bartos and Amman 1989). Intermediate stand  
2 densities of 750 to 1500 stems/ha are thought to be more conducive to beetle-induced tree  
3 mortality (Anhold and Jenkins 1987), giving rise to a nonlinear relationship between risk  
4 and stand density. Stand age relates to the beetle's preference for large diameter trees  
5 (Safrankyik et al. 1974), and stand age has an inverse relationship with tree vigor after  
6 maturity, which determines a tree's ability to resist infection by beetle-introduced fungi  
7 (Shrimpton 1973). Once physiological maturity has been reached, trees become  
8 susceptible to attack, although the rate and likelihood of attack is impacted by other  
9 variables such as climate (Shrimpton and Thomson 1983). Stand composition is included  
10 in the Shore and Safranyik model because stand risk relates to the amount of near term  
11 volume loss, so greater amounts of pine contribute to higher risk ratings. The composition  
12 variable is measured as the percentage of a stand's basal area that is composed of large  
13 diameter pine. The location factor incorporates the impact of geographic location on  
14 beetle survival. At higher elevations and northern latitudes, beetles are exposed to colder  
15 temperatures, thereby increasing winter mortality and disrupting the beetle's development  
16 cycle (Amman 1973).

17

18 Beetle pressure in Shore and Safranyik (1992) is based on the number of beetles within  
19 and proximal to the stand. Beetle population is estimated by the number of infested trees  
20 and the likelihood of attack is incorporated by the distance between the stand being rated  
21 and the infestation. The maximum distance at which beetles from surrounding areas can  
22 enter the stand being assessed is 3 km. Susceptibility variables and beetle pressure are  
23 combined multiplicatively to determine the overall risk rating for a stand of trees. Recent

1 model refinements have replaced discrete look-up tables with continuous equations for  
2 each of the susceptibility variables, the beetle population variable, and for the risk  
3 calculation in an effort to reduce the impact of class boundaries on final risk assessments  
4 (Shore et al. 2006). However, the fundamental elements of the Shore and Safranyik  
5 model remain unchanged: four equally weighted susceptibility variables combined with a  
6 spatial measure of the beetle population to determine a relative ranking of stand risk.

7  
8 With the explosive growth of mountain pine beetle populations in western Canada, the  
9 risk rating system has become a recommended forest planning tool (British Columbia  
10 Ministry of Forests 1995). However, there are considerable limitations to implementing  
11 the Shore and Safranyik system over large areas. Firstly, the relationships in the model  
12 are derived from field research over relatively small geographic areas (Shore and  
13 Safranyik 1992; Shore et al. 2000). Extrapolating these relationships to new geographic  
14 regions may neglect regional variations in mountain pine beetle processes. Secondly, the  
15 data inputs required for operational modelling across large areas often do not exist and  
16 substitute variables available in forest inventory data are generally poor replacements  
17 (Nelson et al. 2006). Hence, there is an information need for a decision support tool  
18 capable of assessing risk over large areas.

19

## 20 ***Recent research***

21

22 Research conducted over the last decade may be able to provide enhancements to existing  
23 risk rating systems. The growth of geographic information systems (GIS) as a tool for  
24 managing complex spatial and attribute information combined with increasing

1 efficiencies in automated and semi-automated data collection technologies has enabled  
2 forest managers and researchers to link theoretical and empirical knowledge of ecological  
3 processes (e.g., Blackburn and Milton 1996). Additionally, more advanced analytical  
4 methods for spatial data are being developed that facilitate spatially explicit analysis of  
5 forest disturbances across large spatial and temporal scales (e.g., Nelson and Boots 2005).  
6 These new results in mountain pine beetle research may enhance risk rating systems.

7

## 8 **Landscape scale red attack modelling**

9

10 A number of studies have investigated the potential for locating and estimating the  
11 severity of mountain pine beetle red attack damage of forested landscapes (red attack is a  
12 term used to describe the characteristic fading of an attacked tree's foliage, which  
13 typically occurs within 6 to 8 months following attack). Variables useful for predicting  
14 red attack damage over large areas may also be helpful for predicting areas at risk of  
15 beetle attack. While remotely sensed data have been used in many studies for detecting  
16 and mapping red attack damage at a range of spatial scales (Sirois and Ahern 1988;  
17 Franklin et al. 2003; Skakun et al. 2003; White et al. 2005; Wulder et al. 2006b), here we  
18 highlight studies that have used imagery in conjunction with ancillary terrain and  
19 radiation information (White et al. 2006; Wulder et al. 2006a; Coops et al. 2006).  
20 Elevation, slope, and direct radiation have all been significant predictors in logistic  
21 regression models of red attack damage (see Table 1). Additionally, Coops et al. (2006)  
22 investigated the relationship between probability of attack, forest structure, and forest  
23 susceptibility variables using regression tree models and found site index and slope to be  
24 most important for explaining variation in the probability of red attack. Negron and Popp

1 2004) used a similar approach for ponderosa pine (*Pinus ponderosa* Lawson) and found  
2 stand density index (SDI) and quadratic mean diameter to be most important for  
3 estimating the plot scale probability of infestation.

4

## 5 **Data uncertainty, accuracy, and large-area spatial analysis**

6

7 Large-area application of forest risk models rely on operational data which often have  
8 varying levels of accuracy and completeness. For instance, detailed variables such as  
9 stand density and basal area by species are not available in most provincial or national  
10 forest inventory products. Nelson et al. (2006) explored the impact of operationally  
11 available data representative of large areas on the Shore and Safranyik susceptibility  
12 model. The authors found that the surrogate variables, due to lack of detail, were only  
13 moderately correlated with the true variables, thereby highlighting the importance of  
14 including operationally available variables in landscape scale models. In an investigation  
15 of the accuracy of the Shore and Safranyik risk model computed with similar large-area  
16 data, Dymond et al. (2006) found the risk index to be within 30% to 43% true positive for  
17 high risk ( $>5$ ), and 93% accurate for low risk ( $>0<5$ ) for predicting presence of  
18 infestation. However, the severity of infestation was not evaluated against the risk index.  
19 Wulder et al. (2006c) compared two methods for estimating the beetle pressure  
20 component of risk, the distance-based method used in Shore and Safranyik (1992) and a  
21 density-based method employing Voronoi polygons; the density-based estimate was  
22 found to have greater correspondence with infestation occurrence.

23

## 1 **Epidemiology and population dynamics**

2  
3 The current state of knowledge on biology and epidemiology of the mountain pine beetle  
4 is succinctly reviewed in Safranyik and Carroll (2006). New research indicates that the  
5 four stage population cycle of the mountain pine beetle (endemic, incipient-epidemic,  
6 epidemic, and post-epidemic) is based on complex interactions with the host tree and an  
7 assemblage of secondary bark beetles. Due to the nature of large area sampling of  
8 mountain pine beetle populations (identification of red attack damage by aerial survey), it  
9 is difficult to detect beetle populations at endemic levels, where spatially disparate  
10 assemblages of only a few trees are infested. The most important aspect of the population  
11 cycle for modelling risk is the epidemic threshold; the point at which the growth in the  
12 beetle population exceeds the stand's ability to resist mass attacks of its large, healthy  
13 trees (Berryman 1982b). It is the large diameter trees that promote exponential growth in  
14 brood production, thereby enabling an epidemic when suitable host and climatic  
15 conditions exist (Safranyik et al. 1974). In population dynamics, this is the shift from the  
16 one stable equilibrium (endemic) to another stable equilibrium (epidemic) (Berryman  
17 1978b). The critical threshold between these equilibria is governed by events that  
18 increase beetle populations (i.e., warm winter temperatures) or decrease stand resistance  
19 (i.e., drought). For modelling the risk of mountain pine beetle damage, it is important to  
20 represent the dynamic nature of this critical population threshold in order to accurately  
21 forecast future infestation advance or collapse (Raffa and Berryman 1986; Bentz et  
22 al.1993; Logan et al.1998; Nelson et al. 2007).

23

## 1 **Range expansion, geographic variation and novel habitats**

2  
3 A key contributor to sustaining the beetle epidemic currently on-going in western Canada  
4 is an increase in climatically suitable habitat, enabling beetle range expansion (Carroll et  
5 al. 2006). Historically, the range of the primary host species exceeded the limits to  
6 mountain pine beetle range imposed by climate conditions. This has important  
7 implications for modelling risk as it is common for processes (i.e., dispersal) at range  
8 margins, to differ from those in traditionally colonized areas (Thomas et al. 2001).  
9 Indeed, latitudinal variation in mountain pine beetle developmental rates has been shown  
10 by Bentz et al. (2001). This perhaps supports other recent research pointing to large scale  
11 spatial synchrony of epidemic beetle populations (Aukema et al. 2006). Mountain pine  
12 beetle spatial processes, such as dispersal, pheromone dynamics, and host selection are  
13 influenced, if not determined, by environmental factors. It might therefore be expected  
14 that spatial patterns representing these processes will vary in different environments.  
15 Robertson et al. (2007) found spatial patterns of dispersal processes to have different  
16 frequencies in different Biogeoclimatic Ecosystem Classification (BEC) subzones. It may  
17 be useful for large scale models of risk to identify regional variations in the spatial  
18 patterns used to represent dynamic beetle processes (i.e., Wu et al. 2005).

## 20 ***Goals and objectives of this research***

21  
22 Knowledge gained from recent mountain pine beetle research may be incorporated into  
23 the next generation of risk models. In this study, we aim to encapsulate some of this  
24 knowledge with the goal of developing a risk model with the following characteristics:

- 25 • An operationally viable model that may be used with existing large-area data.

- 1 • A robust model that improves upon existing models for characterizing risk of  
2 immediate damage by mountain pine beetles over large geographic areas.

3

## 4 **Study Area and Data**

5 The study area for our model development is the Vanderhoof Forest District, located in  
6 central British Columbia, Canada (Figure 1). The approximately 1.38 million ha  
7 Vanderhoof Forest District has experienced substantial losses of lodgepole pine due to  
8 mountain pine beetle attack. Forest inventory data is available for this area and mountain  
9 pine beetle monitoring data is spatially and temporally exhaustive, providing a  
10 comprehensive view of beetle population levels. This area is dominated by forests of  
11 lodgepole pine and spruce (*Picea engelmannii x glauca* Moench. Voss), with a median  
12 age of 105 years. There are three main BEC zones represented in Vanderhoof, ranging in  
13 elevation from 680 m to 1800 m: Sub Boreal Spruce (SBS), Sub Boreal Pine Spruce  
14 (SBPS), and Engelmann Spruce-Subalpine Fir (ESSF). The Vanderhoof Forest District is  
15 proximal to the epicenter of the current outbreak, which is thought to be just to the west  
16 near Entiako Park and Protected Area and Tweedsmuir Provincial Park (Aukema et al.  
17 2006).

18

19 Forest inventory information for the Vanderhoof Forest District conforms to the current  
20 provincial vegetation resource inventory (VRI) standards. VRI is a seamless spatial  
21 coverage of forest stands where attribute information is estimated by a combination of  
22 aerial photo interpretation and field plot verification (British Columbia Ministry of  
23 Sustainable Resource Management 2002). This dataset was last updated with harvest and

1 natural disturbances in 2002. Forest stands in the VRI, defined as homogenous units by  
2 photointerpreters, made up the unit of analysis for all modelling. Attributes from the  
3 inventory that were used in modelling included stand composition, age, density, dbh,  
4 crown closure, and site index.

5  
6 Mountain pine beetle populations were estimated from aerial overview survey (AOS)  
7 data collected in the study area as part of a province-wide forest health survey conducted  
8 annually throughout British Columbia (British Columbia Ministry of Forests 2000).  
9 Broad areas of red attack damage are delineated on 1:100,000 or 1:250,000 basemaps by  
10 trained observers (Wulder et al. 2006c). Severity codes are assigned to these areas to  
11 indicate the proportion of the area that is infested. These broad AOS data are collected  
12 primarily for strategic purposes and are used to direct the subsequent acquisition of more  
13 detailed survey information. In Vanderhoof, detailed surveys were conducted using  
14 helicopters equipped with GPS receivers, where locations of mountain pine beetle attack  
15 were recorded and assigned a severity. Information contained in the AOS and helicopter-  
16 GPS surveys were combined to produce a raster layer indicating the cumulative area of  
17 infestation in the Vanderhoof Forest District for each year from 1999 to 2005. The area of  
18 infestation values were averaged at the forest stand level. All model predictions were  
19 made for the year 2005 based on data up to and including 2004.

20  
21 The BEC system stratifies landscapes based on vegetation, soils, and climatic and site  
22 characteristics. BEC zones are characterized by a common regional climate. Subzones,  
23 the finest spatial unit in the BEC system, represent geographically related ecosystems

1 (Eng and Meidinger 1999). The BEC data used in this analysis were mapped at a scale of  
2 1:20,000 in 2003. The main BEC zone in the Vanderhoof Forest District is the Sub  
3 Boreal Spruce (SBS), making up 84% of the total area. Figure 1 illustrates where  
4 different subzones are located within the study area.

5

6 Elevation base data in the form of a digital elevation model were used for all topographic  
7 variables. This data set was obtained from the Government of Canada geographic data  
8 portal GeoBase, and conforms to Canadian Digital Elevation Data (CDED) 1:50,000  
9 standards. Elevation values are referenced to the Canadian Vertical Geodetic Datum 1928  
10 (DVG28). The data was resampled to 100 m grid cell resolution. The DEM was used to  
11 generate direct solar radiation using the methods of Kumar et al. (1997).

12

## 13 **Methods**

### 14 ***Modelling risk for large areas***

15

16 Our approach to modelling stands at risk of mountain pine beetle attack is based on Shore  
17 and Safranyik (1992). Our intent is to investigate the potential for model improvements  
18 based upon knowledge gained from recent research, while also overcoming some of the  
19 previously noted limitations when determining risk over large areas. Risk in Shore and  
20 Safranyik (1992) model is a continuous value indicating the relative risk of volume loss  
21 due to beetles and is based on basal area. One of the more difficult issues when dealing  
22 with the data constraints of landscape wide forest inventory is the selection of a  
23 dependent variable. Since basal area is often not available in forest inventories, our

1 dependent variable was calculated as the average value of percent mortality pixels for  
2 each polygon, scaled by the percentage of pine associated with the polygon. These values  
3 were then linked back to damage classes (Table 2), so they become the AOS severity  
4 classes scaled by the amount of pine in each VRI polygon, hereafter referred to as  
5 damage level. VRI polygons that did not contain any pine were excluded from the  
6 analysis.

7  
8 The modelling framework for the inclusion and structure of covariate variables was based  
9 on the breakdown of the mountain pine beetle-lodgepole pine system described by Raffa  
10 and Berryman (1986) and others, as either host/stand variables or beetle population  
11 variables. Simulations in Raffa and Berryman (1986) demonstrate how tree and beetle  
12 interactions influence the overall beetle population. Extending this idea to the stand scale,  
13 we capture this interaction by combining stand variables with beetle population variables  
14 (interaction variables). A list of new variables used in this analysis is presented in Table  
15 3.

### 17 **Stand Resistance Variables**

18  
19 Stand resistance typically indicates stand vigour, which is inversely related to the ability  
20 of mountain pine beetles to overcome a tree's defences. New variables related to the  
21 stand's ability to defend against beetle attacks included BEC subzone the stand is located  
22 within and the amount of annual direct shortwave radiation (SWR) (2006a). Other stand  
23 resistance variables that are known to be intimately linked with the state of the beetle

1 population, such as site index, crown closure, and dbh, are included instead as interaction  
2 variables.

3

#### 4 **Beetle Population Variables**

5

6 The integrated beetle population data contains cumulative area infested levels for each  
7 year from 1999 to 2005. From this data we estimated the spatial and temporal trends in  
8 beetle population levels. Beetle population variables provide information about the local  
9 characteristics of the infestation. Beetle population variables used in modelling included  
10 percent pine infested in 2004 (PPI), infestation in 2004 (INC0304), infestation in 2003  
11 (INC0203), duration of infestation (INFDUR) and the number of red attacked trees in  
12 2004 (NUMRED).

13

#### 14 **Interaction Variables**

15

16 Stand variables whose impacts on beetle populations are tied to the population state were  
17 modelled as interaction variables. We used a sigmoid function to scale interaction  
18 variables,  $w$ , on a scale  $[0,x]$ , where  $x$  is the percentage of pine in the stand.

$$19 \quad w = \frac{x}{1 + e^{((t+zr/2-z)/zr)}} \quad (1)$$

20 The threshold value of the variable  $z$  is denoted as  $t$  and the range of  $z$  that defines how  
21 quickly  $w$  reaches  $x$  is defined by  $zr$ . A constant  $c$  defines initial values and was  
22 determined experimentally from the data available. This sigmoid function represents the  
23 threshold nature of mountain pine beetle population interaction with variable  $z$  by

1 defining two parameters: the initial increase of the weight as variable  $z$  increases, and the  
2 value of  $z$  at which exponential increase in weight occurs. The beetle population variable  
3 defines the slope of the curve ( $zr$ ) and the threshold value ( $t$ ) of  $z$  is defined based on  
4 previous mountain pine beetle research. The highest weights are associated with stand  
5 conditions that promote beetle brood production when the population state is at epidemic  
6 levels. Variables modelled as interactions were crown closure (R\_C), site index (R\_SI),  
7 and quadratic mean diameter at breast height (R\_QDBH). Figure 2 presents the details of  
8 the interaction variable curves.

9

### 10 ***Ordinal Logistic Regression***

11

12 The dependent variable for our analysis, damage level, represents the amount of pine  
13 mortality due to mountain pine beetle in six distinct classes (none, trace, low, moderate,  
14 severe, very severe). The ordered nature of this variable warrants the use of an ordinal  
15 regression model. An ordinal model is an extension of binary logistic regression, an  
16 approach which has been successfully employed for predicting locations of mountain  
17 pine beetle red attack damage (Wulder et al. 2006a, Coops et al. 2006). In the ordinal  
18 model, as in binary logistic, the combination of linear predictor variables relates to the  
19 expectation of the dependent variable through a link function, usually the logit function.  
20 Since there are multiple ordered responses, multiple equations need to be resolved in  
21 ordinal regression (Guisan and Harrell 2000).

22

23 There are two commonly used ordinal regression models that deal with dependent  
24 variables derived from a continuous phenomenon that have been categorized. The

1 proportional odds (PO) model (Walker and Duncan 1967) is based on cumulative  
2 probabilities, and the continuation ratio (CR) model (Armstrong and Sloan 1989) is based  
3 on conditional probabilities. CR models are suited for situations where the dependent  
4 variable  $Y$  must pass through one category to reach the next (Guisan and Harrell 2000).  
5 The standard PO model assumes that the slopes of each independent variable  $X$  are equal  
6 for all levels of  $Y$ . Since we are modelling the dynamics between  $Y$  and  $X$  explicitly  
7 through the interaction variables and the progression of damage level need not pass  
8 through one level to reach the next (given our temporal resolution), we selected the PO  
9 model, defined as follows.

$$10 \quad P(Y \geq j | X) = \frac{1}{(1 + \exp[-(a_j + Xb)])} \quad (3)$$

11 Or rather, that the probability of the observed  $Y$  falling in a class greater than or equal to  
12 class  $j$  given the explanatory variables in  $X$  is similar to a logistic model where  $1 = Y \geq j$   
13 and  $0 = Y < j$  for all levels of  $j$  in the ordinal dependent variable  $Y$ . Thus in the case  
14 where  $j = 1$ , the PO model is equivalent to the logistic model. For additional levels of  $j$ ,  
15 the coefficients,  $b$ , stay the same while the intercept term,  $a$ , varies. All models were  
16 developed using functions in the Design package (Harrell 2001) for the statistical  
17 software R (Ihaka and Gentelman 1996). The PO model assumes that the independent  
18 variables vary linearly with the ordinal response variable. To check this assumption, we  
19 plotted the mean of each of the final predictor variables for each damage level against the  
20 expected value under ordinality. Confirmation of the PO assumption can be determined if  
21 the observed means are similar to the expected.

22

1 Despite the more relaxed assumptions for ordinal regression (similar to logistic)  
2 compared to ordinary least squares (OLS), one remaining requirement is that  
3 observations are independent. Since we know mountain pine beetle processes are  
4 structured spatially, we might also expect damage level to be spatially autocorrelated, and  
5 therefore there is the potential for inflated parameter estimates due to reduced degrees of  
6 freedom (Cliff and Ord 1981; Legendre 1993). Following the approach laid out in Bigler  
7 et al. (2005), we applied the Huber-White covariance estimator for cluster-correlated  
8 data, where each stand polygon was treated as a cluster (Huber 1967; White 1982).  
9 Correlated responses were corrected using the Huber-White method as implemented in  
10 the Design package for R (Harrell 2001). All reported model parameter estimates are  
11 corrected versions.

12

### 13 ***Model Development, Selection and Validation***

14

15 Models were developed using a random sample of approximately 50% of the polygons in  
16 the Vanderhoof VRI data ( $n = 23,683$ ). Edge polygons were excluded so as to avoid any  
17 edge effects in the neighborhood effect variable. Significance of predictor variables was  
18 assessed with the Wald  $\chi^2$ , which tests the null hypothesis that the coefficient is zero. The  
19 relative importance of variables was determined with  $\chi^2$ -*df* on significant variables in the  
20 model. The overall fit of the model was assessed with the Nagelkerke R-Square ( $R^2$ )  
21 (Nagelkerke 1991). Furthermore, model fit was also assessed with a measure of  
22 association for ordinal data, Goodman-Kruskal Gamma (Goodman and Kruskal 1954),  
23 which accounts for the ordered nature of the data (i.e., L predicted as M is less wrong  
24 than L predicted as VS); calculated as:

1  $\gamma = (P - Q)/(P + Q)$  (4)

2 Where  $P$  is the number of concordant pairs and  $Q$  is the number of discordant pairs, so  $\gamma$   
3 ranges from 0 to 1 where 1 indicates perfect prediction. Predictions no better than  
4 random will have  $\gamma$  near 0. In our analysis,  $P$  corresponds to the number of pairs  
5 classified in a class adjacent to the true damage level, while  $Q$  is the number of pairs  
6 where the predicted class is not adjacent to the true damage level. Ties, or correctly  
7 predicted classes, are ignored. Gamma can be interpreted as contribution of the  
8 independent variables in reducing errors incurred when predicting the response randomly.

9

10 To assess our model as a means of developing a large-area risk index, we compared a  
11 baseline model where damage level is related to the Shore and Safranyik Pine Risk Index  
12 (PRI), a modification of the 1992 risk rating system, to the best model obtained by adding  
13 the additional variables outlined in Table 3. Additionally, the component variables of PRI  
14 were also included separately during model development to assess their individual impact  
15 on risk. This allowed us to determine the relative enhancements additional variables may  
16 play in risk rating for large areas. Potential collinearity in variables was assessed with  
17 Pearson's  $r$  for each pairwise combination of variables. For values of  $r > 0.7$ , we tested  
18 variables separately during model development. As a final test to determine the impact of  
19 the interaction variables on model results, they were replaced by untransformed versions  
20 and models were re-estimated and assessed.

21

22 Model testing on independent data (i.e., the remaining 50% of the VRI polygons;  $n =$   
23 23,729) was also undertaken to assess the predictive accuracy of the adjusted model. The

1 final adjusted model used for prediction used significant variables with coefficients  $\geq$   
2 0.01. Prediction accuracy was assessed for each level of the dependent variable to assess  
3 the variation in model sensitivity to different damage levels. Prediction accuracy was  
4 measured by calculating sensitivity, specificity and classification accuracy for each  
5 damage level based on a 50% probability threshold.

6  
7 We also assessed prediction accuracy with the area (AUC) under the receiver operator  
8 characteristic (ROC) curve for each set of predictions (Fielding and Bell 1997; Manel et  
9 al. 2001; Allouche et al. 2006; Coudun and Gegout 2007; Wunder et al. 2007). The ROC  
10 is a plot of the true positive rate on the  $y$  axis against the false positive rate on the  $x$  axis.  
11 The AUC is a threshold independent measure of overall accuracy with values from 0.5,  
12 indicating a random model, to 1, indicating perfect prediction. The AUC avoids the need  
13 for an arbitrary selection of a single probability threshold and provides information about  
14 the nature of both sensitivity and specificity (Manel et al 2001; Allouch et al. 2006). ROC  
15 plots for each predicted damage level are presented to assess the trade off between  
16 sensitivity and specificity. Hosmer and Lemeshow (2000) suggest the following general  
17 guide to interpreting the AUC: a value  $>0.9$  is outstanding prediction,  $0.8 - 0.9$  is an  
18 excellent model, and  $0.7 - 0.8$  is an acceptable model.

19

## 20 **Results**

### 21 ***Baseline model***

22

23 The fit of the baseline model is summarized in Table 4a. Overall, PRI in 2004 related  
24 weakly to forest damage level in 2005. PRI was a significant predictor variable ( $\chi^2 =$

1 2603.66,  $p < 0.0001$ ) of forest damage, yet the model explained a small proportion of the  
2 overall variability ( $R^2 = 0.11$ ). The ordinal association measure was low, with  $\gamma = 0.26$ .  
3 The adjusted parameter estimate for PRI indicates a very slight increase (1.54%) in the  
4 odds of an increase in damage level will occur with one unit increase in PRI. The baseline  
5 model was not assessed for predictive accuracy on the independent data.

6

### 7 ***Adjusted model***

8

9 Pearson's correlation matrix of all new variables revealed collinearity in NUMRED and  
10 INC0304. NUMRED was subsequently dropped from model development. The fit of the  
11 adjusted model incorporating additional variables is summarized in Table 4b. With all  
12 variables, the model was moderately successful (Nagelkerke  $R^2 = 0.47$ ,  $\gamma = 0.57$ ) at  
13 explaining the variation in damage level. The relative importance of model variables is  
14 presented in Figure 3. Increased probability of an increase in damage level of forests in  
15 2005 was associated mostly with R\_C, INC0203, BEC subzones in the ESSF zone,  
16 MSxv, SBPSdc, SBSdw, SBSmc, R\_SI and BP. The plots of the independent variable  
17 means against the expected for each damage level indicated that the PO assumption was  
18 justified for most variables (Figure 4). Parameter estimates are provided in Table 4b. Re-  
19 estimating the full model with interaction variables replaced by the untransformed  
20 variables from the VRI (site index, crown closure, quadratic mean dbh) yielded a weaker  
21 model with much less explanatory power (Nagelkerke  $R^2 = 0.21$ ,  $\gamma = 0.37$ ).

22

## 1 **Prediction Accuracy**

2  
3 The adjusted model performed reasonably well when using a 50% probability of damage  
4 threshold for the cutoff (Table 5). The percentage of correctly classified cases ranged  
5 between 75%-93%. Examining the prediction performance for positive against negative  
6 cases revealed widely varying prediction accuracies. The sensitivity, or true positive rate,  
7 ranged from 1%-98%, while the specificity, or true negative rate, ranged from 48%-  
8 100%. Sensitivity declined with increasing damage level, while specificity showed the  
9 opposite trend. The variability in these measures however demonstrates their sensitivity  
10 to prevalence, common to most measures derived from contingency tables (Allouche et al  
11 2006).

12  
13 Predictions made on cumulative damage levels yielded an average AUC of 0.84. This  
14 falls in the range of an 'excellent' predictor. The ROC curves for each damage level are  
15 presented in Figure 5. Predictions of polygons with at least a trace amount of damage,  
16 analogous to predicting red attack presence, yielded an AUC of 0.93. Predictions made at  
17 other damage levels ranged from 0.80 to 0.83. All of these scores indicate good  
18 prediction accuracy model for the adjusted model.

19  
20 The spatial distributions of predicted probabilities for each cumulative damage level are  
21 presented in Figure 6. Polygons in each map indicate the probability value that it is equal  
22 to or above the specified damage level. There is large spatial variation in damage level. A  
23 large patch of non susceptible pine is discriminated fairly well at the low damage level.  
24 Prediction probabilities for very severe are quite low, indicating underprediction for areas

1 where very severe damage existed. However, overall predictive accuracy for the very  
2 severe class was high because so much of the area was classified as true negative (Table  
3 5). The 50% probability threshold used to derive contingency table scores in Table 5 was  
4 too low to capture the predictions of very severe damage output by the model. This  
5 highlights the difficulty of using a fixed probability threshold for predicting multiple  
6 ordinal classes. Depending on infestation level or user information need, the threshold  
7 may be made more or less conservative. The probability values may also be mapped to  
8 spatially identify trends in infestation levels and as an aid to planning or mitigation  
9 activities.

10

## 11 **Discussion**

12 The improvement of the adjusted model over the baseline model highlights the  
13 importance of additional variables for assessing the risk of mountain pine beetle  
14 infestation over large areas. The components of PRI such as tree age, location, stand  
15 density, and stand composition are derived from many field experiments and have been  
16 validated empirically (Safranyik et al 1974 and others), yet when used for modelling  
17 stand risk in this analysis, the PRI did not perform well, either as an index, or when  
18 broken out into its individual factors. Only one (BP) out of the five most important  
19 predictors of damage level was from the baseline model (Figure 3). Reasons for this  
20 disparity are likely based on data quality rather than biological factors associated with the  
21 beetle. Specifically, performance is likely impacted by how well different variables are  
22 represented in the forest inventory and the nature of mountain pine beetle aerial detection

1 surveys. For example, crown closure may be more accurately estimated than stems per  
2 hectare via air photo interpretation. Additionally, the nonlinear relationship between risk  
3 and stem density in PRI, representing intraspecific competition yielding weaker, less  
4 productive beetle broods (Shore and Safranyik 1992), may not be important for modelling  
5 the overall trend of the infestation over large areas. Stands with many moderately  
6 productive trees or fewer very productive trees, as long as they are above the diameter  
7 threshold, will contribute to an increased beetle population the following year.

8

9 The objective of this research was to incorporate knowledge gained from recent mountain  
10 pine beetle research, and new methodological approaches, to identify variables important  
11 for predicting risk of mountain pine beetle infestation over large areas. The adjusted  
12 model (Nagelkerke  $R^2 = 0.47$ ,  $\gamma = 0.57$ ) demonstrated an improved fit over the baseline  
13 model (Nagelkerke  $R^2 = 0.11$ ,  $\gamma = 0.26$ ), indicating the importance of some of the new  
14 variables in predicting risk over large areas. The adjusted model performed well when  
15 used for predicting mountain pine beetle damage level on independent data (AUC =  
16 0.84). The ROC curves for model predictions at different cumulative damage levels in  
17 Figure 5 show that the model predicts best for at least trace amounts, which is similar to a  
18 binary logistic model predicting locations of red attack. Interestingly, the next best  
19 predictions were for polygons of with at least severe and at least very severe damage  
20 levels.

21

22 The use of ordinal regression modelling provides a framework for exploring the  
23 probability of occurrence of different levels of mountain pine beetle infestation. The PO

1 ordinal model specifies constant regression coefficients and regression intercepts that  
2 vary with different levels of the dependent variable. We applied the adjusted model to  
3 independent data and assessed prediction accuracy for different damage levels. Using the  
4 AUC as our measure of predictive accuracy avoided the need to select an arbitrary  
5 probability threshold. This is important for mountain pine beetle risk assessment because  
6 the main source of beetle population data, aerial overview surveys, is collected in ordinal  
7 mortality classes. We were able to explore the variation in these classes, and build a new  
8 model of risk for large areas that predicts new occurrences of mortality classes. In our  
9 analysis, mortality classes were scaled by the percentage of pine in each polygon to yield  
10 damage level.

11

12 Probabilities derived from the adjusted model for cumulative levels of damage were  
13 mapped for the entire study area (Figure 6). Each map shows the probability that each  
14 polygon is at greater than or equal to a given damage level. Examining the difference  
15 between these maps sheds light onto the spatial distribution of mountain pine beetle risk  
16 over a very large area. While beyond the scope of this current research, spatial  
17 investigation of the differences between these probability maps may yield insights into  
18 the importance of various model input variables (Wulder et al. 2007).

19

20 The interaction variables R\_C and R\_SI were both important predictors of damage level.  
21 The s-curve functions defining the interaction variables appear to have captured the  
22 dynamics of damage level increases. When interaction variables were replaced with  
23 untransformed versions, the model fit declined markedly. R\_C was the most important

1 variable in the model (Figure 3). Increases in R\_C were associated with an increased  
2 probability of an increase in damage level. This supports results from Powell et al.  
3 (2000), which found crown closure to be positively related to stand risk to mountain pine  
4 beetle attack. The interaction between beetle population level and crown closure indicates  
5 that high crown closure, high percentage of pine, and a large increase in infestation in the  
6 year previous is positively associated with increases in damage level. Perhaps modelling  
7 crown closure in this way captures the period during which stand resistance declines, yet  
8 phloem thickness remains sufficient to sustain the epidemic (Berryman 1982b), or  
9 indicates that beetle populations have grown to the point where even an effective host  
10 resistance is no defense against beetle mass-attack (Safranyik and Carroll 2006) . For  
11 predicting risk over large areas, R\_C may be a more useful variable for risk rating than  
12 variables based on stand density. This is important because crown closure is more  
13 commonly available than stand density in forest inventory data

14

15 The site index interaction variable, R\_SI, was also an important predictor of damage  
16 level. Site index has been identified in previous models of red attack damage as an  
17 important variable at the stand level (Coops et al. 2006). In our model, R\_SI was  
18 inversely related to the increases in beetle damage. This suggests that high site index  
19 combined with high percentage of pine is associated with lower odds of a stand  
20 sustaining an increased damage level. This is expected if site index accurately reflects the  
21 site conditions of the stand and contributes to increased tree vigor and concomitant stand  
22 resistance to beetle attack. Furthermore, site index is a difficult variable to include in  
23 models of large area risk because it generally refers to site productivity of the leading

1 species. Thus, where lodgepole pine is not the leading species, site index does not  
2 represent site productivity for lodgepole pine. In the Vanderhoof Forest District,  
3 lodgepole pine is a leading species in over 50% of stands. Species composition is less of  
4 an issue in our model because the upper bound of R\_SI is defined by the percentage of  
5 pine, so stands where site index describes a non-pine leading species will have a  
6 maximum R\_SI value less than 0.50.

7  
8 BEC subzones were also important factors for predicting damage level. Negative  
9 associations were found in the SBPSmc and the SBSdk. The SBPSmc occurs along the  
10 south eastern edge and south western edges of the study area (Figure 1). This zone is  
11 dominated by dry lodgepole pine forests, and the site productivity in the SBPS is  
12 generally low. Stressed and stunted trees in this zone may be limiting the broods of  
13 beetles, and the cold climate of the SBPS may be causing greater mortality of beetles in  
14 this area relative to other zones (Steen and Demarchi 1991). The SBSdk occurs  
15 throughout study area at lower elevations, following the Nechako river valley. Here,  
16 conditions for lodgepole pine are more favourable with higher site productivity, yet  
17 mixed forests are also more common (Meidinger et al. 1991). Since a high damage level  
18 is related to the amount of pine in each stand, the risk level may not increase past a  
19 certain level in these areas. All other subzones had positive associations with damage  
20 level. The importance of climate and location in the mountain pine beetle system is well  
21 known. The BEC subzones in our model of stand risk offer a suitably scaled stratification  
22 of the landscape for exploring regional variation in mountain pine beetle processes, and  
23 subsequently, for predicting the future course of epidemics.

1

2 Interestingly the most important beetle population variable in the model was the increase  
3 in infestation in 2003, two years previous to the prediction year. Over large areas,  
4 locations previously infested are more likely to be infested than locations that have not  
5 already subjected to attack (White et al. 2006). It is possible that the winter conditions  
6 (brood over-wintering success) in 2002 were similar to those in 2004, resulting in similar  
7 outcomes.

8

9 Mitigating the spread of mountain pine beetles into new areas may be facilitated by  
10 focusing efforts on known areas of infestation to effect local reductions in beetle  
11 populations (as well as the proportion of beetles eligible for long range transport). The  
12 effort and expense in seeking endemic infestation levels at the leading edge, which may  
13 be less “dangerous” as fewer beetles exist locally for mass attack, and the proportion of  
14 beetles that do engage in long range transport (estimated at 2.5% by Safranyik et al.  
15 1992) will not be sufficient in number to overwhelm the defenses of a mature pine  
16 (especially in climatically difficult, new range, environments). Much research has  
17 indicated that short range dispersal is the primary means of infestation spread over large  
18 areas. Perhaps the largest, most intense landscape infestations should be the focus of  
19 management activities aimed at reducing the spread of beetle populations into new areas.

20

21 Modelling beetle-host interactions at the stand level with interaction variables also  
22 improved our model. Incorporating ecological complexity into models of risk requires a  
23 trade off with accuracy. Our formulation captures some of the dynamics of this ecological

1 complexity, yet remains relatively simple to implement with operational data. Future risk  
2 assessment models may benefit from exploring the concept of interaction variables  
3 further.

4

5 The inclusion of BEC subzones in the model provides evidence that incorporating  
6 information about local ecological characteristics is vital to any model of mountain pine  
7 beetle infestation. For British Columbia, the BEC system offers a suitably scaled  
8 classification which could serve this purpose. However in order to incorporate the full  
9 BEC system into a mountain pine beetle risk model, detailed analysis of the suitability of  
10 each subzone for mountain pine beetle is required. One option would be to map mountain  
11 pine beetle climate suitability classes (Carroll et al. 2006) to BEC subzones and use these  
12 as a component in a large area risk index. Alternatively, regional variation in mountain  
13 pine beetle populations may be too complex to model accurately with one model. A better  
14 approach may be to develop region specific models of risk.

15

## 16 **Conclusion**

17 In this research we developed a model using operationally available data for predicting  
18 mountain pine beetle damage over a large area. Our objective was to identify new  
19 variables related to mountain pine beetle risk, with the aim of ultimately developing an  
20 index, similar to the Shore and Safranyik PRI, which could easily be applied to large  
21 areas throughout the mountain pine beetle range. The results of our model have several  
22 implications for predicting the spatial pattern of mountain pine beetle infestations.

23

1 The addition of variables describing the temporal trend in beetle population improved our  
2 ability to predict future patterns of infestation. Detection and mapping of beetle red attack  
3 damage levels have been topics of extensive research and many sources of data are  
4 available to accurately characterize past infestations (AOS, satellite remotely sensed data,  
5 helicopter-GPS, and air photo interpreted products). Most beetle monitoring programs  
6 have been collecting data for many years and for many locations, multiple years of survey  
7 data have been compiled. Additionally, the use of various remotely sensed data sources  
8 facilitates retrospective analyses for spatial and/or temporal gaps in the historical survey  
9 record. The integration of various sources of survey data allows temporal and spatial  
10 trends in the beetle infestation to be extracted and used for risk assessment. In our  
11 analysis, the beetle population from two years previous to the prediction year was most  
12 important for discriminating damage level.

13

14 Through this research we have demonstrated the utility of including readily available data  
15 sets for representing beetle population trend, interaction with hosts, and an stratification  
16 of the landscape to aid in operational risk rating of mountain pine beetle infestation. We  
17 follow Shore et al. (2006), in acknowledging the stochastic nature of the prediction  
18 mountain pine beetle infestation likelihood. The complex interactions between forest  
19 conditions, climate, and insects have many random or poorly understood elements, and as  
20 a result, any improvement in model predictive ability, especially those facilitated through  
21 readily available additional support data, are welcomed and of use to the science and  
22 management communities interested in mountain pine beetle movement and impacts.

## 1 **Acknowledgements**

- 2 This research was funded by the Government of Canada as part of the Mountain Pine
- 3 Beetle Program administered by Natural Resources Canada, Canadian Forest Service,
- 4 with additional information available at: <http://mpb.cfs.nrcan.gc.ca/>.

## 1 **References**

2 O. Allouche, T. Asaf, K. Ronen, 2006. Assessing the accuracy of species distribution  
3 models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*,  
4 43, 1223-1232.

5

6 G.D. Amman, 1973. Population changes of the mountain pine beetle in relation to  
7 elevation. *Environmental Entomology*, 2, 541-547.

8

9 Amman, G.D., McGregor, M.D., Cahill, D.B., Klein, W.H., 1977. Guidelines for  
10 reducing losses of lodgepole pine to the mountain pine beetle in unmanaged stands in the  
11 Rocky Mountains. USDA For. Serv. Gen. Tech. Rep. INT-136.

12

13 J. Anhold, M. Jenkins, 1987. Potential mountain pine beetle (Coleoptera: scolytidae)  
14 attack of lodgepole pine as described by stand density index. *Environmental Entomology*,  
15 16, 738-742.

16

17 B.G. Armstrong, M. Sloan, 2001. Ordinal regression models for epidemiologic data.  
18 *American Journal of Epidemiology*, 129, 191-204.

19

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

1

2 Bartos, D.L., Amman, G.D., 1989. Microclimate: An alternative to tree vigor as a basis  
3 for mountain pine beetle infestations. Res. Pap. INT-400, USDA For. Serv. 10p.

4

5 B. Bentz, G. Amman, L. Logan, 1993. A critical assessment of risk classification systems  
6 for the mountain pine beetle. *Forest Ecology and Management*, 61, 349-366.

7

8 B.J. Bentz, J.A. Logan, J.C. Vandygriff, 2001. Latitudinal variation in *Dendroctonus*  
9 *ponderosae* (Colyoptera: Scolytidae) development time and adult size. *Canadian*  
10 *Entomologist*, 133, 375-387.

11

12 Berryman, A.A. 1978a. A synoptic model of the lodgepole pine/mountain pine beetle  
13 interactions and its potential application in forest management. In: Theory and practice of  
14 mountain pine beetle management in lodgepole pine forests, ed. A.A. Berryman, G.D.  
15 Amman, and R.W. Stark, 98-105. Moscow, ID: University of Idaho.

16

17 A.A. Berryman, 1978b. Towards a theory of insect epidemiology. *Researches on*  
18 *Population Ecology*, 19, 181-196.

19

20 Berryman, A.A. 1982a. Population dynamic of bark beetles. Pages 264-314 In: J. B.  
21 Mitton and K.B. Sturgeon, eds. Bark beetles in North American conifers: a system for the  
22 study of evolutionary biology. University of Texas Press, Austin, TX. 527 p.

23

1 A.A. Berryman, 1982b. Mountain pine beetle outbreaks in Rocky Mountain lodgepole  
2 pine forests. *Journal of Forestry*, 80, 410-413.  
3

4 C. Bigler, D. Kulakowski, T.T. Veblen, 2005. Multiple disturbance interactions and  
5 drought influence fire severity in Rocky Mountain subalpine forests. *Ecology*, 86, 3018-  
6 3029.  
7

8 G.A. Blackburn, E.J. Milton, 1996. Filling the gaps: remote sensing meets woodland  
9 ecology. *Global Ecology and Biogeography Letters*, 5, 175-191.  
10

11 British Columbia Ministry of Forests. 1995. Bark Beetle Management Guidebook Table  
12 of Contents. Forest Practices Branch.  
13

14 British Columbia Ministry of Forests. 2000. Forest health aerial overview survey  
15 standards for British Columbia, Version 2.0. Forest Practices Branch and Canadian Forest  
16 Service, Victoria, BC, 45 p.  
17

18 British Columbia Ministry of Sustainable Resource Management. 2002. Vegetation  
19 Resources Inventory: Photo Interpretation Procedures, Version 2.4.  
20 <http://www.for.gov.bc.ca/hts/vri/standards/>.  
21

22 Carroll, A.L., Regniere, J., Logan, J.A., Taylor, S.W., Bentz, B.J., Powell, J.A., 2006.  
23 Impacts of climate change on range expansion by the mountain pine beetle. *Natural*

1 Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, BC.  
2 Mountain Pine Beetle Initiative Working Paper 2006-14. 20 p.  
3  
4 Cliff, A., Ord, J.K., 1981. Spatial Processes Models and Applications. Pion Limited,  
5 London, 266p.  
6  
7 N.C. Coops, M.A. Wulder, J.C. White, 2006. Integrating remotely sensed and ancillary  
8 data sources to characterize a mountain pine beetle infestation. *Remote Sensing of*  
9 *Environment*, 30, 83-97.  
10  
11 C. Coudun, J.C. Gegout, 2007. Quantitative prediction of the distribution and abundance  
12 of *Vaccinium myrtillus* with climatic and edaphic factors. *Journal of Vegetation Science*,  
13 18, 517-524.  
14  
15 C.C. Dymond, M.A. Wulder, T.L. Shore, T.A. Nelson, B. Boots, B. Riel, 2006.  
16 Evaluation of risk assessment of mountain pine beetle infestations. *Western Journal of*  
17 *Applied Forestry*, 21, 5-13.  
18  
19 M. Eng, D. Meidinger, 1999. A method for large-scale biogeoclimatic mapping in British  
20 Columbia. Province of British Columbia, Victoria, BC.  
21  
22 A.H. Fielding, J.F. Bell, 1997. A review of methods for the assessment of prediction error  
23 in conservation presence / absence models. *Environmental Conservation*, 24, 38-49.

1

2 S. Franklin, M.A. Wulder, R. Skakun, A. Carroll, 2003. Mountain pine beetle red-attack  
3 forest damage classification using stratified Landsat TM data in British Columbia,  
4 Canada. *Photogrammetric Engineering & Remote Sensing*, 69, 283-288.

5

6 L.A. Goodman, W.H. Kruskal, 1954. Measures of association for cross classifications.  
7 *Journal of the American Statistical Association*, 49, 732-764.

8

9 A. Guisan, F.E. Harrell, 2000. Ordinal Response Regression Models in Ecology. *Journal*  
10 *of Vegetation Science*, 11, 617-626.

11

12 Harrell, F.E. 2001. Regression Modeling Strategies: With applications to linear models,  
13 logistic regression, and survival analysis. Springer, New York, 600 p.

14

15 Hosmer, D.W., and Lemeshow, S. 2000. Applied Logistic Regression. 2<sup>nd</sup> edn. John  
16 Wiley, New York, 392 p.

17

18 Huber, P.J. 1967. The behaviour of maximum likelihood estimates under non-standard  
19 conditions. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and  
20 Probability 1:221-233.

21

22 I. Ihaka, R. Gentleman, 1996. R: a language for data analysis and graphics. *Journal of*  
23 *Computational and Graphical Statistics*, 5, 299-314.

1

2 L. Kumar, A.K. Skidmore, E. Knowles, 1997. Modelling topographic variation in solar  
3 radiation in a GIS environment. *International Journal of Geographic Information*  
4 *Science*, 11, 475-497.

5

6 P. Legendre, 1993. Spatial autocorrelation: Trouble or new paradigm? *Ecology*, 74, 1659-  
7 1673.

8

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

11

12 J. Logan, B. Bentz, 1999. Model Analysis of Mountain Pine Beetle (Coleoptera:  
13 Scolytidae) Seasonality. *Environmental Entomology*, 28, 924-934.

14

15 Mahoney, R. 1978. Lodgepole pine-mountain pine beetle risk classification methods and  
16 their application. In *Theory and practice of mountain pine beetle management in*  
17 *lodgepole pine forests*, ed. D.M. Baumgartner, 106-113. Pullman, WA.

18

19 S. Manel, H.C. Williams, S.J. Ormerod, 2001. Evaluating presence-absence models in  
20 ecology: the need to account for prevalence. *Journal of Applied Ecology*, 38, 921-931.

21

1 Meidinger, D., J. Pojar, and W.L. Harper. 1991. Sub-Boreal Spruce Zone. In Ecosystems  
2 of British Columbia, 209-221. Special Report Series 6, British Columbia Ministry of  
3 Forests, Victoria, BC.

4

5 N.J. Nagelkerke, 1991. A note on a general definition of the coefficient of determination.  
6 *Biometrika*, 78, 691-692.

7

8 J.F. Negron, J.B. Popp, 2004. Probability of ponderosa pine infestation by mountain pine  
9 beetle in the Colorado Front Range. *Forest Ecology and Management*, 191, 17-27.

10

11 T. Nelson, B. Boots, 2005. Identifying insect infestation hot spots: an approach using  
12 conditional spatial randomization. *Journal of Geographical Systems*, 7, 291-311.

13

14 T. Nelson, B. Boots, M.A. Wulder, T. Shore, L. Safranyik, T. Ebata, 2006. The impact of  
15 data on the rating of forest susceptibility to mountain pine beetle infestations. *Can. J.*  
16 *For. Res.*, 36, 2815-2825.

17

18 W.A. Nelson, A. Potapov, M.A. Lewis, A.E. Hunsdorfer, F. He, 2007. Balancing  
19 ecological complexity in predictive models: A reassessment of risk models in the  
20 mountain pine beetle system. *Journal of Applied Ecology*, 45, 248-257.

21

1 K.F. Raffa, A.A. Berryman, 1986. A mechanistic computer model of mountain pine  
2 beetle populations interacting with lodgepole pine stands and its implications for forest  
3 managers. *Forest Science*, 32, 789-805.

4

5 C. Robertson, T. Nelson, B. Boots, 2007. Mountain Pine Beetle Dispersal: The spatial-  
6 temporal interaction of infestations. *Forest Science*, 53, 395-405.

7

8 Safranyik, L, DM Shrimpton, and HS Whitney. 1974. Management of lodgepole pine to  
9 reduce losses from the mountain pine beetle. Forestry Technical Report 1. Pacific  
10 Forestry Centre, Canadian Forestry Service, Victoria, BC, 25 p.

11

12 Safranyik, L., Shrimpton, D. and Whitney, H. 1975. An interpretation of the interaction  
13 between lodgepole pine, the mountain pine beetle, and its associated blue stain fungi.  
14 Managment of Lodgepole Pine Ecosystems. In: Baumgartner, D.M. (ed.), Managment of  
15 Lodgepole Pine Ecosystems Symp. Proc., Washington State Univ. Coeprative Extension  
16 Service, pp. 406-428.

17

18 Safranyik, L., Carroll, A. 2006. The biology and epidemiology of the mountain pine  
19 beetle in lodgepole pine forests. In: The mountain pine beetle, a synthesis of biology,  
20 management, and impacts on lodgepole pine, 3-66. Victoria, BC: Natural Resources  
21 Canada, Canadian Forest Service, Pacific Forestry Centre.

22

1 J.A. Schenk, R.L. Mahoney, J.A. Moore, D.L. Adams, 1980. A model for hazard rating  
2 lodgepole pine stands for mortality by mountain pine beetle. *Forest Ecology and*  
3 *Management*, 3, 57-58.

4

5 Shore, T., Safranyik, L. 1992. Susceptibility and risk rating systems for the mountain pine  
6 beetle in lodgepole pine stands. Information Report No. BC-X-336. Forestry Canada,  
7 Pacific and Yukon Region, Pacific Forestry Centre, Victoria, BC, p.12.

8

9 T.L. Shore, L. Safranyik, J.P. Lemieux, 2000. Susceptibility of lodgepole pine stands to  
10 the mountain pine beetle: testing of a rating systems. *Canadian Journal of Forest*  
11 *Research*, 20: 44-49.

12

13 Shore, T.L., B.G. Riel, L. Safranyik, and A. Fall. 2006. Decision support systems. In: The  
14 mountain pine beetle: A synthesis of biology, management, and impacts on lodgepole  
15 pine, 193-230. Victoria, B.C.: Natural Resources Canada, Canadian Forest Service,  
16 Pacific Forestry Centre.

17

18 D.M. Shrimpton, 1973. Age-and size-related response of lodgepole pine to inoculation  
19 with *Euophium clavigerum*. *Canadian Journal of Botany*, 51, 1155-1160.

20

21 D.M. Shrimpton, A.J. Thomson, 1983. Growth characteristics of lodgepole pine  
22 associated with the start of mountain pine beetle outbreaks. *Canadian Journal of Forest*  
23 *Research*, 13, 137-144.

1

2 J. Sirois, F.J. Ahern, 1989. An investigation of SPOT HRV for detecting recent mountain  
3 pine beetle mortality. *Canadian Journal of Remote Sensing*, 14, 104-110.

4

5 R. Skakun, M.A. Wulder, S. Franklin, 2003. Sensitivity of the Thematic Mapper  
6 enhanced wetness difference index to detect mountain pine beetle red-attack damage.  
7 *Remote Sensing of Environment*, 86, 433-43.

8

9 O. Steen, D.A. Demarchi, 1991. Sub-Boreal Pine-Spruce Zone. In *Ecosystems of British*  
10 *Columbia*, 195-208. Spec. Rep. Ser. 6: B.C. Ministry of Forests.

11

12 J.D. Stuart, 1984. Hazard rating of lodgepole pine stands to mountain pine beetle  
13 outbreaks in southcentral Oregon. *Can. J. For. Res.*, 14, 666-671.

14

15 C.D. Thomas, E.J. Bodsworth, R.J. Wilson, A.D. Simmons, Z.G. Davies, M. Musche,  
16 2001. Ecological and evolutionary processes at expanding range margins. *Nature*, 411,  
17 577-581.

18

19 S.H. Walker, D.B. Duncan, 1967. Estimation of the probability of an event as a function  
20 of several independent variables. *Biometrika*, 54, 167-179.

21

22 H. White, 1982. Maximum likelihood estimation of misspecified models. *Econometrica*,  
23 50, 1-25.

1

2 J.C. White, M.A. Wulder, D. Grills, 2006. Detection and mapping mountain pine beetle  
3 red-attack damage with SPOT-5 10-m multispectral imagery. *BC Journal of Ecosystems  
4 and Management*, 7, 105-118.

5

6 H.X. Wu, C.C. Ying, H.B. Ju, 2005. Predicting site productivity and pest hazard in  
7 lodgepole pine using biogeoclimatic system and geographic variables in British  
8 Columbia. *Ann. For. Sci.*, 62, 31-42.

9

10 M.A. Wulder, J.C. White, B. Bentz, M.F. Alvarez, N.C. Coops, 2006a. Estimating the  
11 probability of mountain pine beetle red-attack damage. *Remote Sensing of Environment*,  
12 101, 150-166.

13

14 M.A. Wulder, J.C. White, C.C. Dymond, T. Nelson, B. Boots, T.L. Shore, 2006b.  
15 Calculating the risk of mountain pine beetle attack: A comparison of distance- and  
16 density-based estimates of beetle pressure. *Journal of Environmental Informatics*, 8, 58-  
17 69.

18

19 M.A. Wulder, C.C. Dymond, J.C. White, D.G. Leckie, A.L. Carroll, 2006c. Surveying  
20 mountain pine beetle damage of forests: A review of remote sensing opportunities. *Forest  
21 Ecology and Management*, 221, 27-41.

22

- 1 M.A. Wulder, J.C. White, N.C. Coops, T. Nelson, B. Boots, 2007. Using local spatial  
2 autocorrelation to compare outputs from a forest growth model. *Ecological Modelling*,  
3 209, 264-276.
- 4
- 5 J. Wunder, B. Reineking, J-F. Matter, C. Bigler, H. Bugmann, 2007. Predicting tree death  
6 for *Fagus sylvatica* and *Abies alba* using permanent plot data. *Journal of Vegetation*  
7 *Science*, 18, 525-534.
- 8

Table 1 Overview of recent studies and variables for predicting mountain pine beetle red attack damage.

Publication	Scale	Primary Data Source	Dependent Variable	Significant Predictors	Method	Accuracy (95% CI)
Wulder et al. 2006	landscape	Landsat7, DEM 30m	Red attack presence	+elevation, -slope	Logistic	86% (5%)
White et al. 2006	landscape	SPOT5 10m, DEM 25m	Red attack presence	+direct radiation, -elevation	Logistic	71% (9%)
Coops et al. 2006	landscape	Landsat7, Landsat5, DEM 30m	Red attack presence	+direct radiation, +elevation	Logistic	69% (N/A)
Coops et al. 2006	landscape	VRI	Probability of red attack	Site index, slope, basal area, crown closure	CART	N/A
Coops et al. 2006	landscape	VRI	Probability of red attack	Location factor, age factor, basal area factor, density factor	CART	N/A
Negron and Popp 2004	Plot	Field plot	Probability of red attack	Stand density index, quadratic mean diameter	CART	N/A

Table 2 – Damage level classes are based on the amount of infestation and the percentage of pine. Percent area mortality in the AOS data is multiplied by the percentage of pine in each VRI polygon.

Damage Level	Criteria
None	0
Trace	$\leq 1$
Low	$>1 \leq 10$
Medium	$>10 \leq 29$
Severity	$>29 \leq 49$
Very Severe	$>49$

Table 3 – New Variables included in model development.

<b>Variable name</b>	<b>Category</b>	<b>Rationale</b>
BEC Subzone (BEC)	Stand Resistance	The beetle life cycle is principally determined by temperature. BEC zones may represent suitably scaled delineations for identifying variations in the productivity of beetle populations.
Annual direct shortwave radiation (SWR)	Stand Resistance	Changing climate may impact the effect of elevation on beetle populations. Direct radiation may indicate variability in the productivity of beetle populations more accurately than elevation alone.
Percent pine infested in 2004 (PPI04)	Beetle Population	The proportion of pine infested represents both beetles and the amount of pine remaining in the stand.
Damage level in 2004 (INC0304)	Beetle Population	Aerial increase in infestation in 2004 as recorded in aerial surveys.
Damage level in 2003 (INC0203)	Beetle Population	Aerial increase in infestation in 2003 as recorded in aerial surveys.
Number of years infestation present in 2004 (INFDUR)	Beetle Population	Indicator of the infestation trend.
Number of red attacked trees in stand in 2004 (NUMRED)	Beetle Population	Estimate of the local beetle population.
Site index (R_SI)	Interaction	High site indices represent productive stands and under epidemic conditions, productive beetle populations.
Crown closure (R_C)	Interaction	Stand canopies determine the amount of radiation received on the bole and impact pheromone mediated dispersal.
Quadratic mean diameter at breast height (R_QDBH)	Interaction	Mountain pine beetles attack large diameter pine disproportionately. Tree diameters above 25.4 cm are thought to be infestation sources, while those below this threshold are infestation sinks (Safranyik et al. 1974)

Table 4 – Model coefficients and estimates.

a) Baseline Model. Nagelkerke  $R^2 = 0.11$ , K-W  $\gamma = 0.26$ .

Parameter	coefficient	Wald	Sig.	S.E.
Y >= T	0.8941	37.73	0.000	0.024
Y >= L	-0.3184	-14.55	0.000	0.022
Y >= M	-1.4177	-61.07	0.000	0.023
Y >= S	-2.4484	-91.21	0.000	0.027
Y >= VS	-3.7628	-107.51	0.000	0.035
PRI	0.0153	51.03	0.000	0.000

b) Adjusted Model. Nagelkerke  $R^2 = 0.47$ , K-W  $\gamma = 0.57$ .

Parameter	coefficient	Wald	Sig.	S.E.
Y >= T	-0.8820	-12.83	0.000	0.069
Y >= L	-2.5582	-35.88	0.000	0.071
Y >= M	-3.9838	-54.66	0.000	0.073
Y >= S	-5.3274	-70.54	0.000	0.076
Y >= VS	-6.9153	-85.44	0.000	0.080
PRI	-0.0047	-9.45	0.000	0.000
BP	2.3129	30.87	0.000	0.075
Inc0203	-6.1958	-40.34	0.000	0.154
Inc0304	-1.1121	-5.15	0.000	0.216
R_QDBH	1.1073	9.67	0.000	0.114
R_C	4.8708	67.80	0.000	0.072
R_SI	-2.6736	-30.27	0.000	0.088
ESSFmvp	1.1219	2.20	0.028	0.510
ESSFxv	1.1601	3.11	0.002	0.373
MSxv	0.4191	1.89	0.059	0.222
SBPSdc	0.2374	2.00	0.046	0.119
SBPSmc	-0.6227	-10.57	0.000	0.059
SBSdk	-0.6039	-12.89	0.000	0.047
SBSdw	0.3819	8.47	0.000	0.045
SBSmc	0.2810	7.10	0.000	0.040

Table 5 – Prediction accuracy for different damage levels using 50% probability threshold. Sensitivity is the percentage of correctly predicted true positives, and specificity is the percentage of correctly predicted true negatives. Correct classification rate is the total percentage of correctly predicted cases.

<b>Damage level</b>	<b>Sensitivity</b>	<b>Specificity</b>	<b>Correct Classification Rate</b>
T	98%	48%	91%
L	84%	60%	76%
M	67%	80%	75%
S	38%	92%	81%
VS	1%	100%	93%

## Figure Captions.

Figure 1 – Biogeoclimatic ecosystem classification (BEC) subzones of the Vanderhoof Forest District.

Figure 2 – Interaction variable curves. For each, the maximum weight is defined by the percentage of pine in the stand. Additional parameters are based on the beetle population in 2004 (bp2004), the associated VRI variables, and threshold values. The threshold value indicates when rapid increase begins and  $x - \text{bp2004}$  defines the number of units in the VRI variable required to reach the max weight. If bp2004 is high, the weight increases faster than if bp2004 is low. Site index threshold was based on the most important value for predicting red attack damage in regression tree analysis in Coops et al. 2006. Quadratic mean diameter threshold was based on a similar minimum dbh threshold of attack in Safranyik et al. 1974 (10cm), and so that the weight increase is proportional to the percentage mortality increase for each unit increase in diameter (increase of 5 at peak growth period). Crown closure threshold of 20 was set experimentally as no theoretical values were found.

Figure 3 – Variables in new model. The measure of variable importance in the model,  $x^2 - df$ , takes into consideration the extra degrees of freedom in BEC subzone variable (categorical predictors), and is equivalent to the  $x^2$  test for continuous predictor variables.

Figure 4 – Plots of independent variable means for each ordinal class for adjusted model. The PO assumption is confirmed if the observed (solid) matches the predicted (dotted) under the assumption of proportional odds. This can be roughly confirmed for 4 out of the 6 variables, suggesting the PO model was appropriate.

Figure 5 – ROC curves for 2005 mountain pine beetle forest damage levels predicted from separate model alpha coefficients.

Figure 6 – Maps of predicted probabilities for each damage level.

Colin Robertson  
Department of Geography  
University of Victoria  
Mailing Address:  
University of Victoria  
Dept of Geography, PO Box 3050  
Victoria, BC  
V8W 3P5

February 4, 2008

Prof. P.M. Attiwill  
Editor  
*Forest Ecology and Management*

Dear Editor,

Please find attached our manuscript, titled: Risk rating for mountain pine beetle infestation of lodgepole pine forests over large areas with ordinal regression modelling.

This manuscript reports on an ordinal regression based approach to characterizing stands at risk of mountain pine beetle infestation using large-area, operational data. We review recent trends in mountain pine beetle research and incorporate relevant variables into a risk model.

Existing risk and susceptibility systems for mountain pine beetle were developed from stand scale field experiments and were designed with specific data requirements. Yet their use as models of risk across large landscapes, often with substitute variables, is cause for concern. In this research, we develop an operationally viable model for risk rating stands using standard inventory and forest health monitoring data. Further, our research encapsulates recent knowledge gained from the growing volume of large-scale forest disturbance research.

We appreciate consideration of our manuscript for publication in the *Forest Ecology and Management*.

Sincerely,

Colin Robertson

(Corresponding Author)

Cc: M.A. Wulder  
T.A. Nelson  
J.C. White

Figure1  
[Click here to download high resolution image](#)

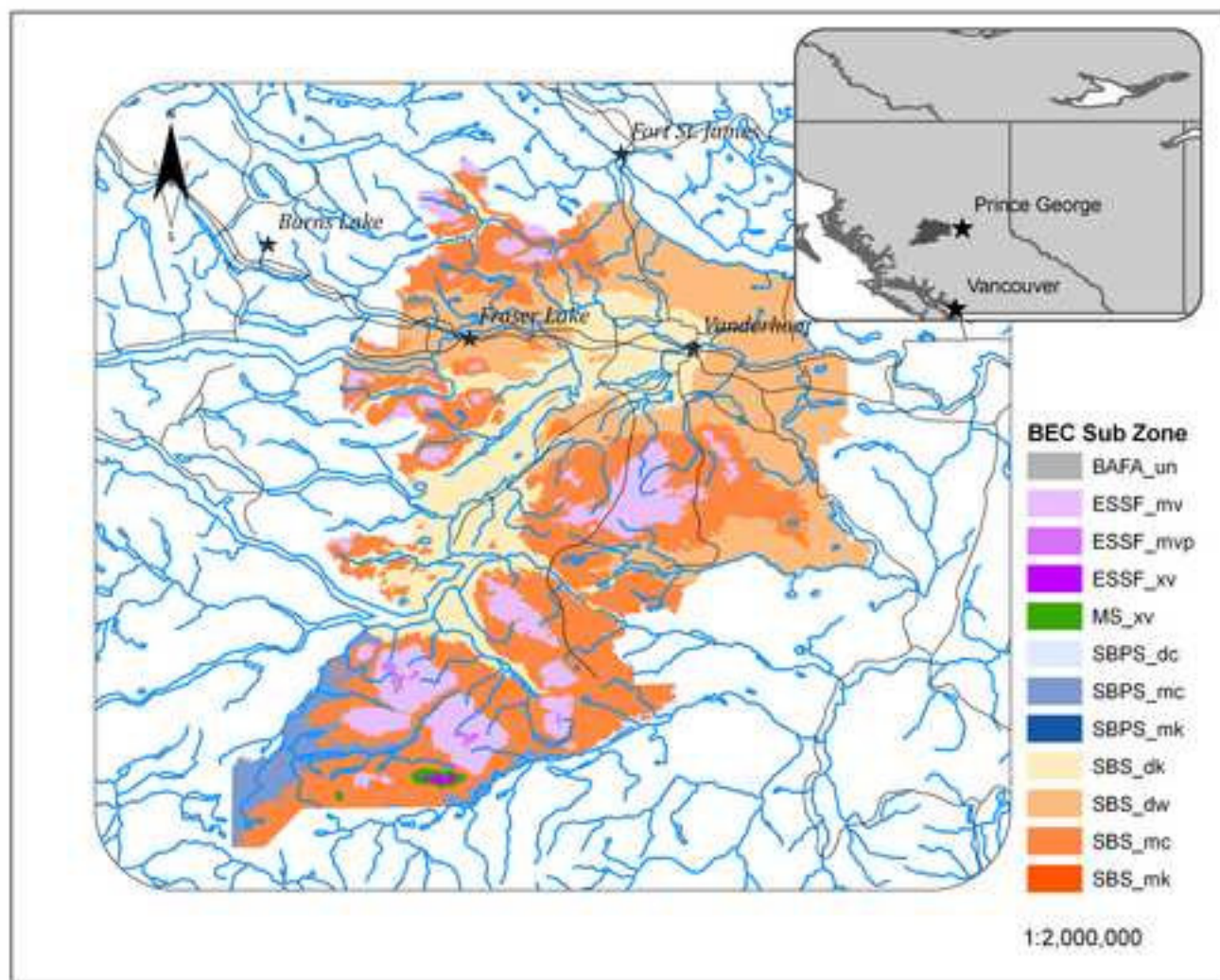


Figure2

[Click here to download high resolution image](#)

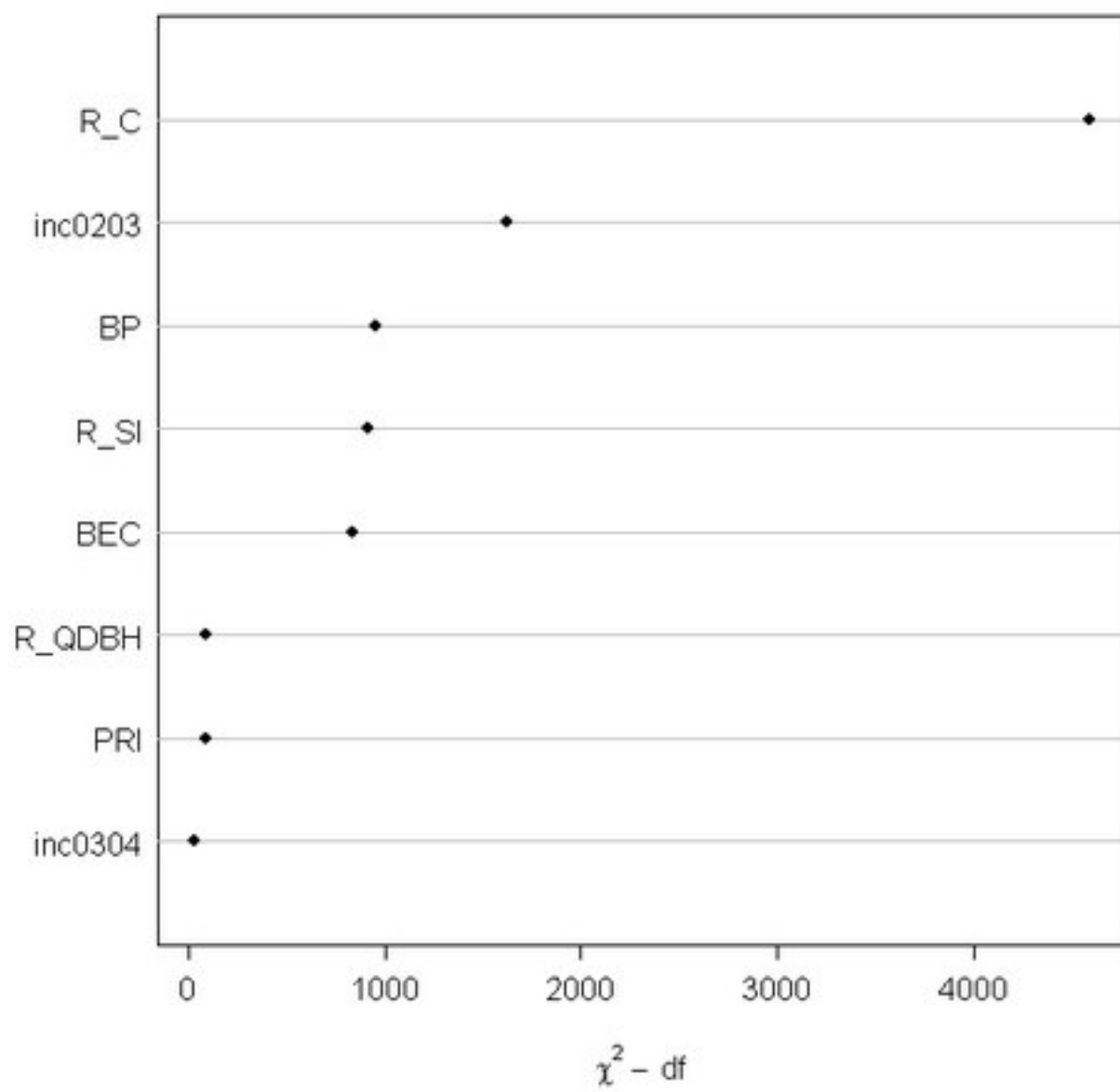


Figure3

[Click here to download high resolution image](#)

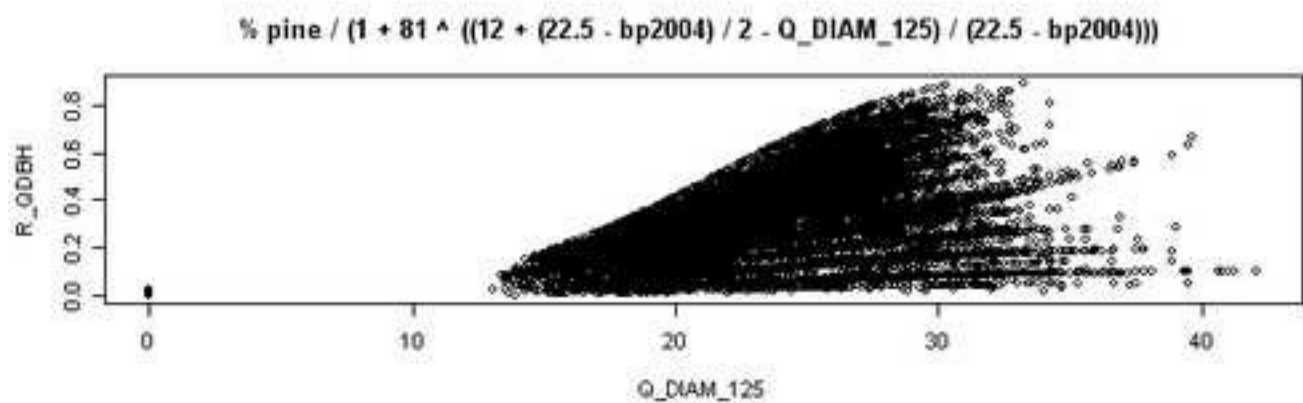
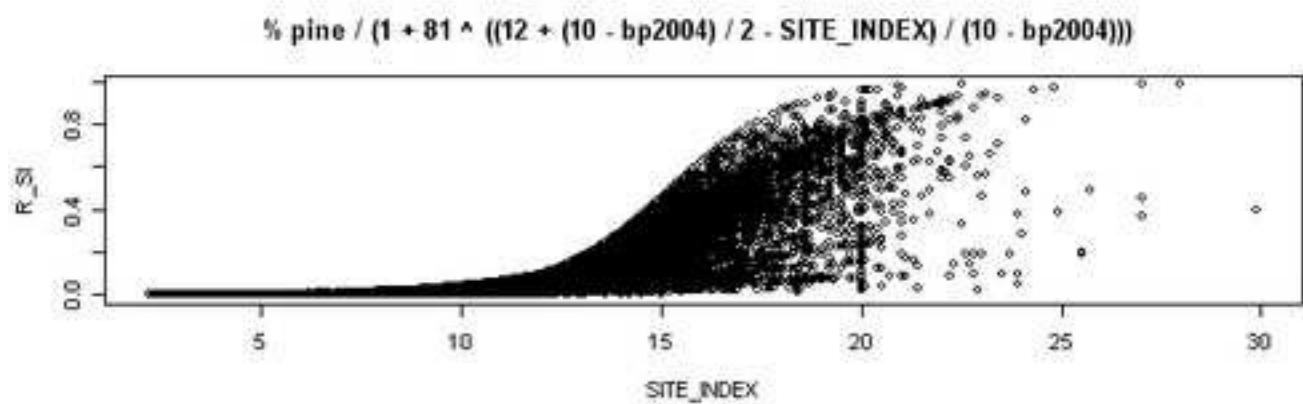
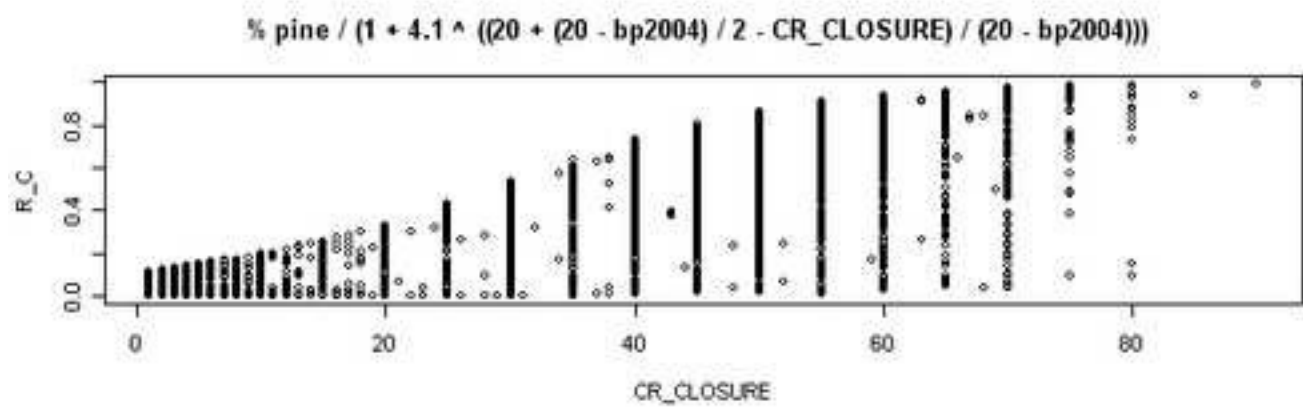


Figure4  
[Click here to download high resolution image](#)

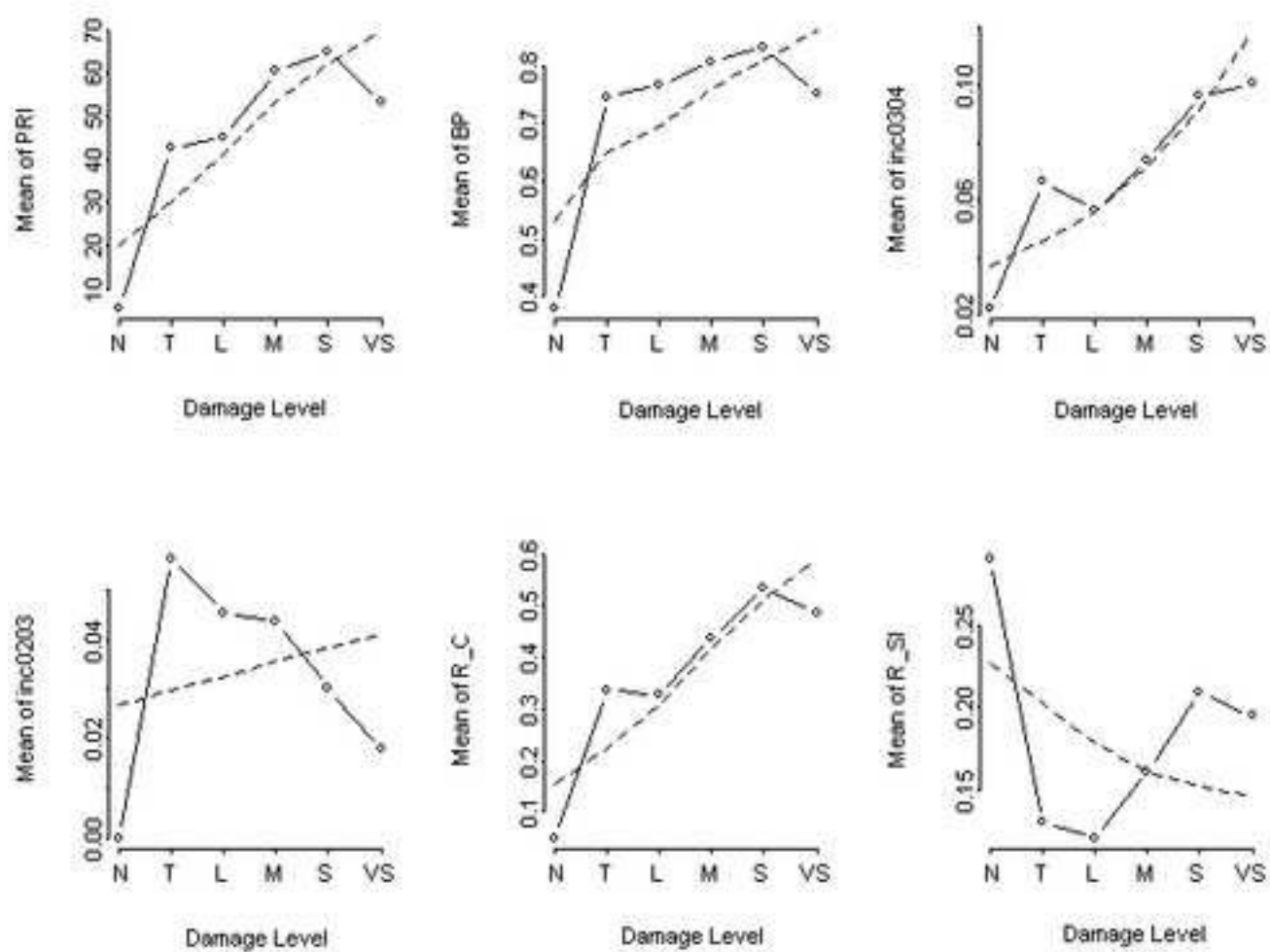


Figure5

[Click here to download high resolution image](#)

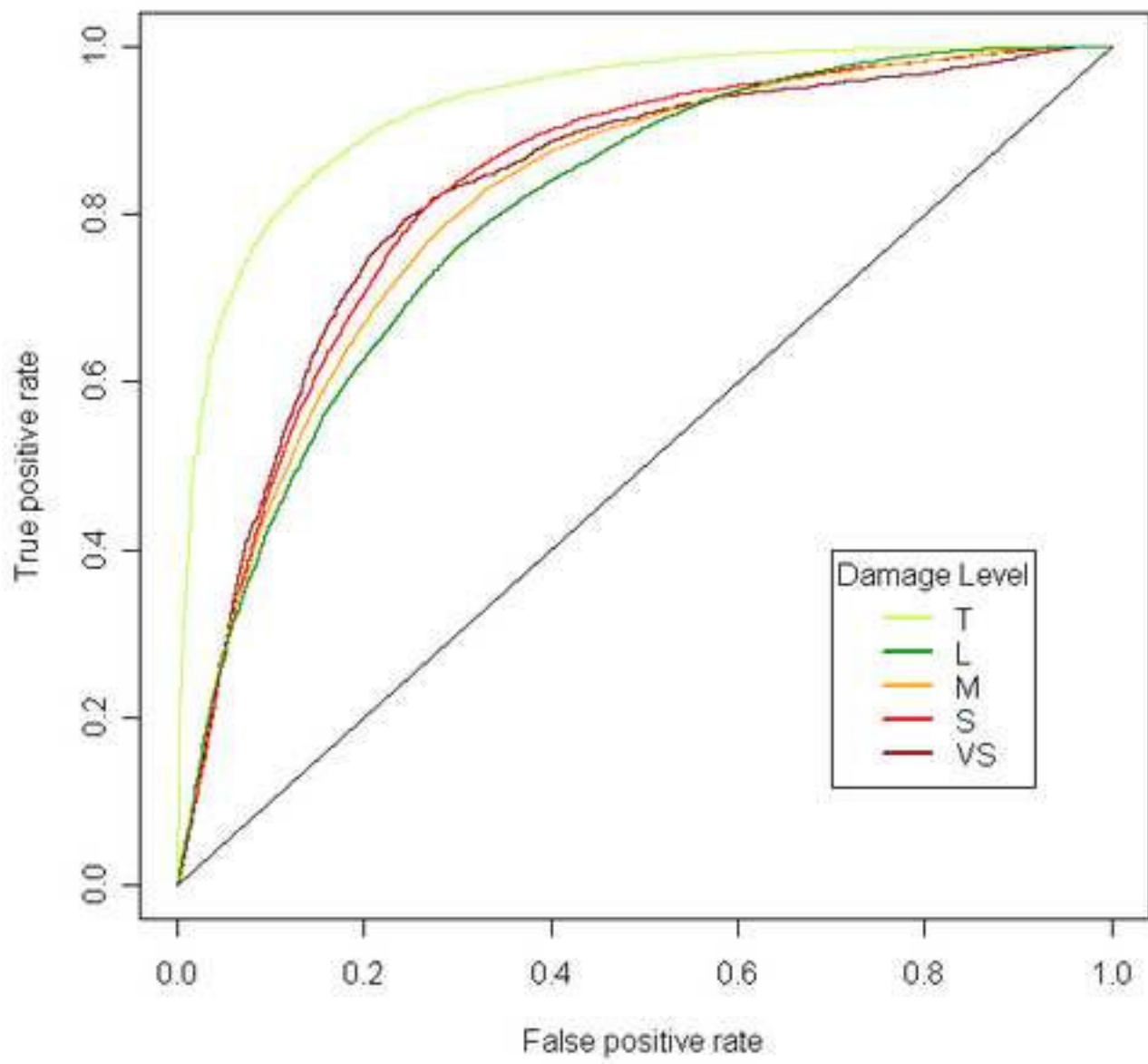


Figure6

[Click here to download high resolution image](#)

