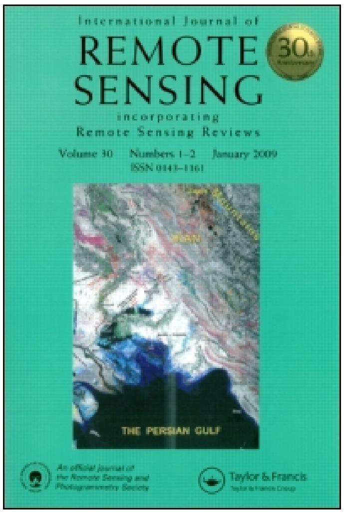

International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713722504

Integrating high resolution remote sensing, GIS and fuzzy set theory for identifying susceptibility areas of forest insect infestations

C. Bone [a]; S. Dragicevic [a]; A. Roberts [a]
[a] Department of Geography, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada

Online Publication Date: 10 November 2005

To cite this Article Bone, C., Dragicevic, S. and Roberts, A.(2005)'Integrating high resolution remote sensing, GIS and fuzzy set theory for identifying susceptibility areas of forest insect infestations',International Journal of Remote Sensing,26:21,4809 — 4828

To link to this Article: DOI: 10.1080/01431160500239180

URL: http://dx.doi.org/10.1080/01431160500239180

International Journal of Remote Sensing
Vol. 26, No. 21, 10 November 2005, 4809–4828

Taylor & Francis
Taylor & Francis Group

Integrating high resolution remote sensing, GIS and fuzzy set theory for identifying susceptibility areas of forest insect infestations

C. BONE*, S. DRAGICEVIC and A. ROBERTS
Department of Geography, Simon Fraser University, 8888 University Drive, Burnaby, BC, V5A 1S6, Canada

(*Received 21 January 2005; in final form 26 June 2005*)

The use of fuzzy set theory has become common in remote sensing and geographical information system (GIS) applications to deal with issues surrounding the uncertainty of geospatial datasets. The objective of this study is to develop a model that integrates the concept of fuzzy set theory with remote sensing and GIS in order to produce susceptibility maps of insect infestations in forest landscapes. Fuzzy set theory was applied to information extracted from multiple-year high resolution remote sensing data and integrated in a raster-based GIS to create a map indicating the spatial variation of insect susceptibility in a landscape. Variable-specific fuzzy membership functions were developed based on expert knowledge and existing data, and integrated through a semantic import model. The results from a case study on mountain pine beetle (*Dendroctonus ponderosae* Hopkins) illustrate that the model provides a method to successfully estimate areas of varying susceptibility to insect infestation from high resolution remote sensing images. It was concluded that fuzzy sets are an adequate method for dealing with uncertainty in defining susceptibility variables. The susceptibility maps can be utilized for guiding management decisions based on the spatial aspects of insect–host relationships.

1. Introduction

High resolution remote sensing images of forest landscapes can provide resource managers with important information regarding insect infestations. Images captured over multiple years can be used to investigate the spatial dynamics of a variety of forest pests, especially those that have significant effects on the forest canopy. While numerous attempts at detecting damage caused by insects have proved successful (Gimbarzevsky *et al.* 1992, Franklin *et al.* 2003, Roberts *et al.* 2003, Skakun *et al.* 2003, Riel *et al.* 2004), research that utilizes high resolution remote sensing data and the analytical tools of geographical information systems (GIS) for estimating the susceptibility of forests to various insects is not extensive. This is because forests are complex systems that are heterogeneous and continuously changing over space and time. Therefore, it is difficult to capture the dynamic nature of forests in remote sensing images at appropriate spatial and temporal resolutions to be able to fully analyse the data in a GIS. Understanding the complexity of forest systems is also limited by the lack of multiple year remote sensing data at an established study site to provide information on the nature of insect infestations through time. Furthermore, insect behaviour varies across space and time causing uncertainty in

*Corresponding author. Email: cbone@sfu.ca

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2005 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160500239180

quantitatively describing how specific variables affect susceptibility. Despite existing methods developed in remote sensing and GIS literature to deal with uncertainty of spatial data (Fisher 1999, Foody and Atkinson 2002, Zhang and Goodchild 2002) little has been done in relation to the heterogeneous and continuous nature of forest processes (Brown 1988a, Lowell 1996).

Uncertainty in geographical information science (GIScience) refers to the degree of inexactness when assigning precise values to both locations and attribute that define geographical data (Robinson 2003). For example, defining boundaries between forest stands in a GIS from remote sensing data may include uncertainty because the exact boundary between two stands cannot be precisely determined, or bias may be introduced when defining the stands as either deciduous, coniferous or mixed (Lowell and Gold 1995). This explains why the majority of susceptibility mapping to date, which integrates remote sensing and GIS, has focused on processes governed by larger spatial and temporal scales such as landslides (Van Westen 2000, Cevik and Topal 2003, Sarkar and Kanungo 2004), while less research has been performed on short-term dynamic processes with greater spatial heterogeneity such as insect infestations (Luther *et al.* 1997). Uncertainty can also be caused by a lack and insufficiency of relevant data that are necessary to study geographic phenomena. In the absence of sufficient data it is difficult to measure how various elements influence the spatial variation of concepts such as susceptibility. Therefore, there is a need for methods that acknowledge uncertainty in extracting information from remote sensing data for developing susceptibility maps of potential insect disturbance. In order to address these issues, the main objective of this study was to develop a model that integrates the concept of fuzzy set theory with remote sensing and GIS to define the susceptibility of different areas to insect infestation in forest landscapes.

Fuzzy set theory, developed by Zadeh (1965), has played a role in dealing with uncertainty in remote sensing and GIS since Bezdek (1984) introduced the fuzzy c-means algorithm (Robinson 2003). Fuzzy sets allow for partial membership to one or more classes, thus objects are represented by a value based on a membership continuum between 0 and 1. The membership function of an element x belonging to a fuzzy set A is represented by $\mu_A : U \rightarrow [0,1]$, where U is the universal set of x. This explains that the function associates a graded membership with each point x in U. The development of the fuzzy membership function is an important component of using fuzzy sets, and is accomplished using expert opinion and available data to define the function.

The use of fuzzy sets and fuzzy logic reasoning was found suitable to address the nature of geographic boundaries and the fact that spatial objects do not always precisely fit into the classes in which they are assigned by traditional remote sensing or GIS procedures. Specifically, the concepts of fuzzy sets have been employed for defining the spatial and attribute characteristics of geographic objects (Burrough 1996, Wang and Hall 1996), soil classes (Burrough *et al*. 1992, Davidson *et al*. 1994), temporal interpolation (Dragicevic and Marceau 2000), and enhancing classification of remote sensing images (Wang 1990, Zhang and Foody 1998, Brandtberg 2002, Lucieer and Kraak 2004). With regards to applications with forest landscapes, fuzzy sets have been used to deal with issues surrounding digitizing objects from thematic maps (Lowell and Gold 1995), identifying forest types (Brown 1998) and identifying individual trees (Brandtberg 2002). The challenge still exists to use fuzzy sets for developing a realistic representation of susceptibility to various spatial phenomena.

The proposed model for defining the susceptibility of areas in a forest first extracts information regarding susceptibility from multiple-year high resolution remote sensing data of a site to create a raster GIS database. High resolution data play an important role in this model because forest infestation is often studied at the individual tree scale. Variables that determine susceptibility to insect attack include the biological characteristics of a tree as well as its proximity to already infested trees; these types of variable can in some form be measured from high resolution images. After the remote sensing data interpretation, fuzzy set theory was used to define how various characteristics of a forest affect susceptibility to insect infestation, and to define the spatial elements of the variables involved in the infestation process. Information from each variable was integrated using a fuzzy operator to produce a final map defining levels of susceptibility of different areas.

While modifications will have to be made to the model to fit different forest insects in different geographical locations, the overall framework is applicable to numerous scenarios due to similar life cycle characteristics. In this study, the model was tested with a case study of the mountain pine beetle (*Dendroctonus ponderosae* Hopkins) in the central interior of British Columbia, Canada. A review of this insect's life cycle suggests that a large pine tree within a pure stand of pine trees close to a previous attack, and distanced from large constraints, is more likely to be susceptible than the opposite for each given variable. However, a problem arises when attempting to define how the terms *large*, *pure*, *close* and *distanced* affect susceptibility due to the issues regarding representation of reality from remote sensing images in a GIS. With fuzzy sets and fuzzy logic reasoning (Zadeh 1972), membership functions can be developed to explain how these variables affect susceptibility based on existing knowledge. Instead of defining each location in space as either, for example, a large tree or not a large tree, close to a previous attack or not close, etc., each location receives a membership based on the degree to which they represent such variables. As a result of combining these variables, the value representing susceptibility is continuous between 0 and 1 instead of defined as either susceptible or non-susceptible.

2. Methodology

This section is divided into two parts. The first presents the model framework for using a fuzzy set approach to obtain information from multiple-year high resolution remote sensing data and GIS operations in order to produce susceptibility maps of insect infestation in forest landscapes. The second part is a case study applying the model to mountain pine beetle infestations.

2.1 *Model for fuzzy susceptibility mapping*

The procedure that defines the model for developing susceptibility maps is illustrated in figure 1 where grey boxes indicate processes and white boxes represent the inputs and outputs of the processes. The first step includes identification of the variables responsible for susceptibility to insect infestation that can be measured from remote sensing images. In the second step, the initial-year image is interpreted in order to obtain information on the variables for a specific time represented by T_i; these variables are integrated in a raster-based GIS. This can be accomplished by either combining image interpretation techniques with ground truth data or by using existing image classifications. This results in multiple layers of information that correspond to different variables.

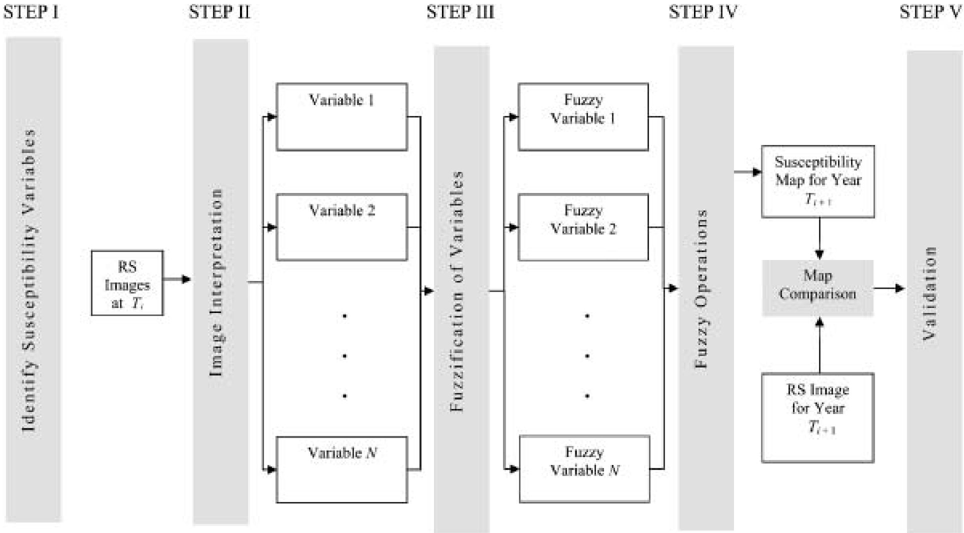

Figure 1. Model framework for development and validation of susceptibility maps.

For the third step, each layer enters a fuzzification process to assign a value that represents both the susceptibility variable and the positional uncertainty of intermediate objects. The susceptibility variables were fuzzified using specifically constructed fuzzy membership functions in order to assign a value to each variable that represents the degree to which they belong to a particular set that has an influence on susceptibility. The fuzzy functions are developed using the semantic import model (Burrough 1989), which uses accumulated expert knowledge of particular insects and the nature of infestations to assign the membership grade. Robinson (2003) provided a discussion on the different types of functions commonly used for geographic phenomena. The output from this stage is layers with information corresponding to each fuzzy variable.

With regards to positional uncertainty, forest stands are considered intermediate objects because the boundary between two different stands cannot be precisely determined. Therefore, the fuzzy area between stands can be defined based on its degree of belonging to either adjacent stand. A fuzzy membership function for belonging to a stand is developed by assigning those cells (i.e. pixels) that are definitely in the stand a value of 1 for full membership, and those cells that are definitely in the adjacent stand a value of 0 to represent non-membership. Each cell in the transition zone receives a fuzzy value $\mu(x_{tz})$ based on its inverse distance to the adjacent stand and the susceptibility of that stand to attack, which is given by the equation

$$\mu(x_{tz}) = \max\left\{ \left[\mu(x_i)\frac{1}{D_i}\right], \left[\mu(x_j)\frac{1}{D_j}\right] \right\} \tag{1}$$

where $\mu(x_i)$ and $\mu(x_j)$ are the fuzzy membership values of the nearest cell in adjacent stands i and j, respectively, which are joined by the transition zone, and D is the distance to the nearest cell of the respective adjacent stands. Each cell in the transition zone will have two values as it is a member of two stands; therefore, the maximum value is obtained in order to represent the higher level of susceptibility.

The fourth step includes the use of fuzzy operators to combine all fuzzy variables into one layer to represent the overall susceptibility in the landscape. The operator computationally obtains one value from the raster layers for each cell location for the output layer. GIS applications traditionally rely upon AND (i.e. minimum) or OR (maximum) operators; however, while their use remains valid for a variety of applications, minimum and maximum operators for combining susceptibility variables will reflect the best and worst case scenarios, respectively. This can limit the ability to detect the influence of the values from each layer (Robinson 2003), and can also create problems when calibrating values or performing a sensitivity analysis, because changes in values that are not the minimum or maximum remain unnoticed. A simple solution to this problem is to obtain the algebraic product (compensatory operator) of each cell location from the raster layers (Robinson 2003). This ensures that the susceptibility of each cell takes into account all the variables used in the operation. The result from this step is a single raster layer containing the susceptibility of each cell $\mu(Su)$ to insect infestation based upon the defined variables.

The fifth and final step of the susceptibility model represents the model validation procedure and can be performed if remote sensing data for consecutive years are available. The procedure can be validated by comparing the observed locations of attack during the first year of analysis represented by T_{i+1}, with the susceptibility map of T_{i+1} to determine if higher susceptibility areas experienced a higher frequency of insect attack than lower susceptibility areas. A statistical test such as a *large-sample test comparing two population proportions* (McClave and Sincich 2000) can be used to determine if a significantly higher proportion of attack occurred in areas of higher susceptibility than lower susceptibility.

2.2 *Case study: mountain pine beetle in British Columbia, Canada*

Mountain pine beetle is the most serious insect disturbance agent of mature lodgepole pine, *Pinus contorta* and ponderosa pine, *Pinus ponderosa*, in western North America (Safranyik 1988). Trees killed by this insect can be readily detected by high resolution remote sensing imagery as they exhibit a red colour in the year following a mortal attack (Wulder and Dymond 2004). The susceptibility model can be customized for mountain pine beetle susceptibility as illustrated in figure 2.

Step I—identify susceptibility variables. The first step is to identify the variables that define the forest's susceptibility to mountain pine beetle. A life cycle review demonstrates that species diversity within a stand (Amman and Baker 1972, Shore *et al.* 2000), distance to trees attacked in the previous year (Thomson 1991, Safranyik *et al.* 1999b), and the diameter of the trees that are susceptible to attack (i.e. the host tree) (Mitchell and Preisler 1991, Moeck and Simmons 1991, Perkins and Roberts 2003) are significant indicators of susceptibility to attack. Exploratory analysis of the data revealed that the distance to large deciduous stands also played a role in influencing areas attacked by the mountain pine beetle. These four variables can be used in the model to develop maps of forest susceptibility.

Step II—image interpretation. The high resolution multi-spectral aerial photographs used for this study were collected in 2002 and 2003 at a pixel resolution of 15 cm (Roberts *et al.* 2003). Ground truth data for the aerial photographs were collected in 2001 by the British Columbia Ministry of Forestry (BC MoF), and in 2002 by Simon Fraser University and BC MoF. The sites are located in the central interior of

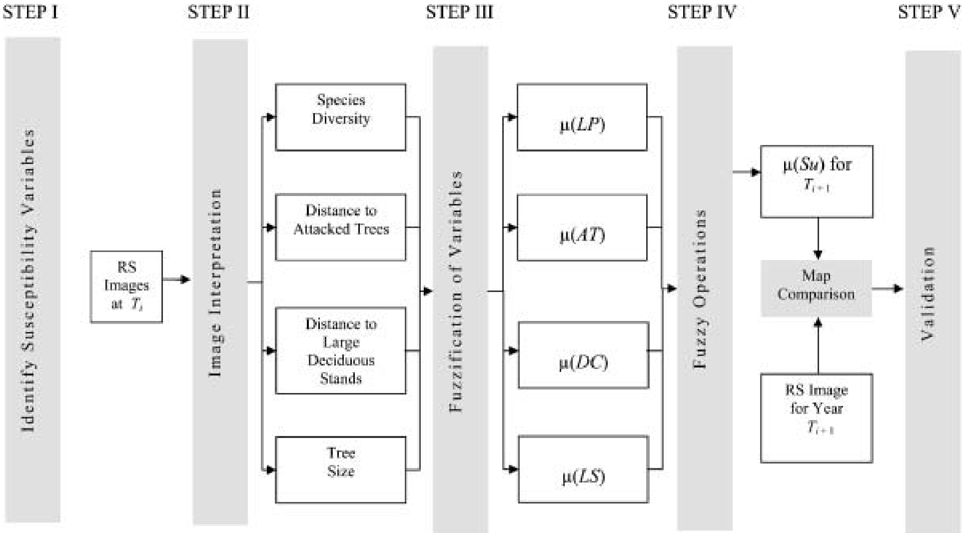

Figure 2. Model framework for development and validation of susceptibility maps specific to mountain pine beetle. The model combines the fuzzy variables of species diversity $\mu(LP)$, distance to attack $\mu(AT)$, distance to large deciduous stand $\mu(DC)$, and tree size $\mu(LS)$ to produce a single value for susceptibility $\mu(Su)$ to MPB attack.

British Columbia: site 1 centred at 53°38′45″ W and 123°26′20″ N, and site 2 centred at 53°29′07″ W and 125°06′40″ N (figure 3). The size of each site is approximately 750 m × 750 m in which there is minimal variation in elevation. The forest in these areas is dominated by lodgepole pine, and contained a relatively small composition of white spruce, *Picea glauca*, Douglas fir, *Pseudotsuga menziesii*, and trembling aspen, *Populus tremuloides*. The sites also contain roads and open areas that were previously created for forestry operations and forest management, which can be seen on the study site insets in figure 3.

The images were interpreted to specify the four variables and produce GIS raster-based information for (1) *Attacked Trees* variable—polygons indicating trees attacked by mountain pine beetle at T_{i} (2000), $T_{\mathrm{i}+1}$ (2001) and $T_{\mathrm{i}+2}$ (2002); (2) *Constraints* variable—polygons representing water bodies, open areas and deciduous stands; (3) *Species Diversity* variable—polygons defining stands of deciduous and coniferous trees and individual coniferous trees found in deciduous stands; (4) *Tree Size* variable—polygons defining the density of coniferous areas (i.e. high, medium and low). The values for density in this layer were used to construct information on tree size. After digitizing was completed, each layer was converted to raster files to perform cell-based operations. A resolution of 1 m was chosen during the conversion to raster files because the smallest significant tree crowns were estimated to occupy this area. The choice to use this spatial resolution inflicted some bias in the procedure because the original image resolution was significantly higher.

Step III—fuzzification. In this step specific fuzzy membership functions were developed to assign a value that represented both the susceptibility variable and the positional uncertainty of the intermediate objects.

The *Species Diversity* variable was used to define the proportion of host trees in a stand. A stand is considered more susceptible if it contains only lodgepole pine. First, all deciduous and coniferous stands were identified and digitized, and an

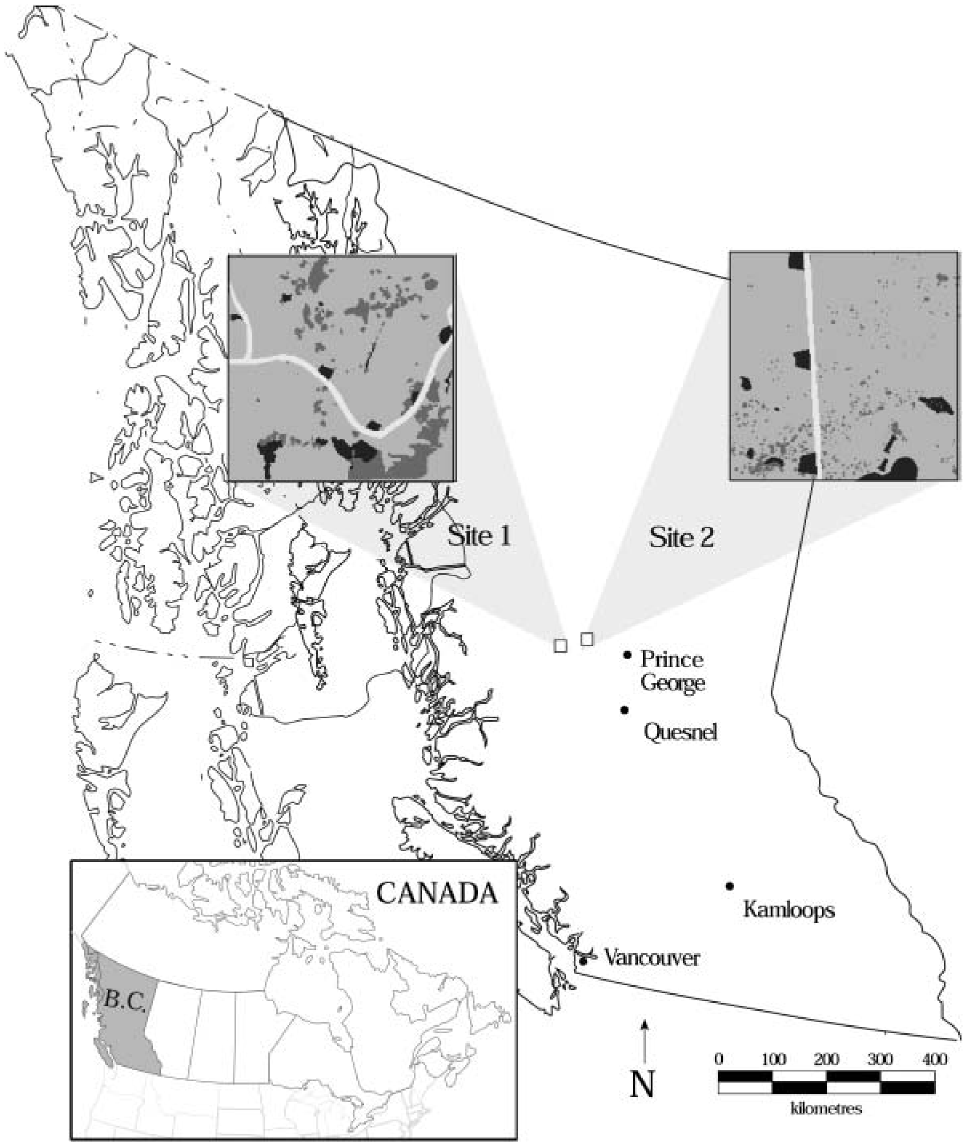

Figure 3. Study sites in central interior of British Columbia, Canada.

estimation of the proportion of each species in each stand was obtained from the ground truth data. A fuzzy membership function for pure lodgepole pine stand $\mu(LP)$ was then developed using expert knowledge. An increase in the proportion of host trees in a stand leads to a higher membership of $\mu(LP)$, which corresponds to an increase in susceptibility (Amman and Baker 1972, Thomson 1991, Shore and Safranyik 1992). Figure 4(*a*) illustrates the relationship that is defined by the location of d_1 and d_2 on the x-axis. A shift in d_1 to the right increases the influence of non-host trees to susceptibility. This means that susceptibility does not change as the proportion of lodgepole pine becomes higher until x reaches d_1. Conversely, a shift in d_2 to the left dampens the effect of diversity on susceptibility. Although sigmoid functions are commonly used to express this type of relationship in GIS

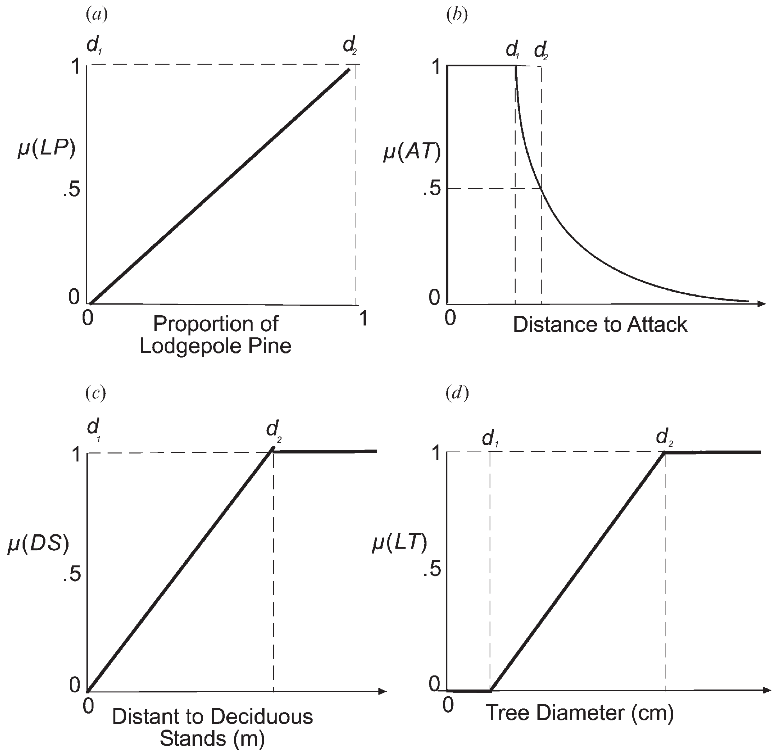

Figure 4. Potential fuzzy membership functions for susceptibility $\mu(Su)$ to mountain pine beetle attack based on (a) species diversity $\mu(LP)$, (b) distance to attack $\mu(AT)$, (c) distance to large deciduous stands $\mu(CD)$, and (d) tree size $\mu(LP)$. Values for d_1 and d_2 indicate the function parameters based on the location along the x-axis.

applications, a linear function was used because the rate at which $\mu(LP) \rightarrow 1.0$ that would define a sigmoid function could not be adequately determined due to the lack of sufficient data. This statement also applies for the variables *Tree Size* and *Distance to Constraints*, which also use linear fuzzy membership functions as described below. The fuzzy membership function for $\mu(LP)$ is defined by

$$\mu(X) = \left\{ \begin{array}{ll} 0 & \text{for } x < d \\ \frac{x-d_1}{d_2-d_1} & \text{for } d_1 \leq x \leq d_2 \\ 1 & \text{for } x > d_2 \end{array} \right\} \tag{2}$$

where $X = LP$.

The *Attacked Trees* variable was used to build layers representing trees attacked in T_i, T_{i+1} and T_{i+2} (the latter two are used to develop susceptibility maps for subsequent years and for validation). A standard distance function was performed on the layer containing trees attacked in T_i. This provides the distance between each location in space and the nearest attacked tree. Next, a fuzzy membership function was constructed to assign each cell a value representing the degree to which it is close to a cell attacked in the previous year $\mu(AT)$. Thomson (1991) explained that mountain pine beetle first attack trees in a small patch, then move outwards to other areas, then outside the stand. This suggests a negative relationship between susceptibility and distance to attacked trees. This is supported by Shore and Safranyik's (1992) rating system where the likelihood of attack is considered greater the closer a tree is to the nearest infestation. However, the results from some studies suggest that this relationship is not linear (Safranyik *et al.* 1989), which is potentially due to a constant occurrence of attack within a certain distance from previously attacked trees, followed by a non-linear decrease as distance increases. This type of relationship can be represented using a j-shaped function shown in figure 4(*b*) (Burrough 1989), which is expressed as

$$\mu(AT) = \left\{ \begin{array}{ll} \dfrac{1}{1+\left(\frac{x-d_2}{d_2-d_1}\right)^2} & \text{for } x > d_1 \\ 1 & \text{for } x < d_1 \end{array} \right\} \tag{3}$$

where x is the distance between an infested tree and a susceptible host, d_1 is the location on the x-axis where $\mu(AT)$ begins to deviate from 1, and d_2 indicates the position on the curve where $\mu(AT)=0.5$. As d_1 shifts to the left, the distance of constant occurrence of attack decreases, and the slope also shifts to the left. A shift of d_2 in either direction changes the nature of the curve.

The *Distance to Constraints* variable was used to obtain information on the impact of large deciduous stands on susceptibility. Observations from images for both sites indicated that the attack was relatively low close to the large deciduous stands. This was characterized by a positive relationship between increasing infestation and distance. An analytical procedure was performed to determine the distance over which large deciduous stands affect the mountain pine beetle, and the nature of this relationship. This information was then used to build a fuzzy membership function in order to assign a value to each cell representing proximity to large deciduous stands $\mu(DS)$. Figure 4(*c*) illustrates a linear function that can be explained by equation (2), where $X=DS$. As d_1 moves to the right, the distance over which the beetle does not attack near large deciduous stands increases. As d_2 moves to the left, attack will occur closer to large deciduous stands.

The *Tree Size* variable was used to develop information regarding the overall tree size of a stand. The larger the trees in a stand, the more susceptible that stand will be to attack. The frequency of tree size is generally a function of stand density, which can be visually interpreted from the images. Stands exhibiting high density are younger stands that have yet to experience competitive exclusion, thus there are many trees of relatively small diameter. Conversely, low-density stands contain larger and older trees that have out-competed other trees, which in turn decrease the number of trees. Ground truth data were used to confirm if the visually defined levels of stand density contained different tree size distributions, and individual stands were then digitized and defined accordingly.

Next, a fuzzy membership function was developed that represented membership for the class of large trees $\mu(LT)$. The susceptibility of a tree to mountain pine beetle infestation increases as diameter increases (Shore and Safranyik 1992, Hindmarch and Reid 2001). This was supported by the ground truth data. The value of $\mu(LT)$ is represented by a positive linear fuzzy function (figure 4(*d*)) which is explained by equation (2), where $X=LT$. Shifting d_1 to the left increases the size range of trees that are susceptible to attack. Moving d_2 in either direction will affect the size of trees that are most susceptible. The degree of membership values for $\mu(LT)$ was then used in equation (4) to obtain a single fuzzy value for belonging to a stand with large trees $\mu(LS)$:

$$\mu(LS) = 2\left[\sum(p\mu(LT))\right] \tag{4}$$

where p is the proportion of each tree size in the frequency distribution provided by the ground truth data. The products of p and $\mu(LT)$ are summed to represent the susceptibility given by the entire distribution of trees. The sum is multiplied by two to exaggerate the difference between stand susceptibility.

The final process in Step III was to transform the discrete boundaries of the forest stands into fuzzy boundaries. Applying equation (1) to the $\mu(LP)$ and $\mu(LS)$ layers ensured that all forest stands contained fuzzy boundaries that represented their intermediate nature. The cells in the transition zone of the layer with information on $\mu(LP)$ and $\mu(LS)$ are denoted as $\mu(LP_{tz})$ and $\mu(LS_{tz})$, respectively.

Step IV—fuzzy operator. The four layers with fuzzy values were then combined using a compensatory operator in order that the fuzzy membership value of each variable is expressed as a final susceptibility value $\mu(Su)$. Using a traditional AND or OR operator would only reflect the best or worst case scenario, respectively, and would not account for the effects of the other variables. The fuzzy operator for $\mu(Su)$ is expressed as

$$\mu(Su) = [\mu(LP), \mu(LP_{tz})]\mu(DS)\mu(AT)[\mu(LS), \mu(LS_{tz})] \tag{5}$$

Either $\mu(LP)$ or $\mu(LP_{tz})$ is used in the equation depending on whether the cell is located directly within a stand or within a transition zone, as is the same for $\mu(LS)$ and $\mu(LS_{tz})$. The result was a single map for each site showing the estimated degree of susceptibility $\mu(Su)$ of the various site areas in the forests for T_{i+1}.

Step V—validation. In order to validate the susceptibility maps for each site, the raster layer containing observed locations of insect attack for the year T_{i+1} was overlaid on the map of susceptibility at T_{i+1} to visually and statistically determine if the maps appropriately estimated areas of susceptibility. Next, a one-tailed large-sample test comparing two population proportions (McClave and Sincich 2000) was used to determine if a significantly higher proportion of cells were attacked in areas of higher susceptibility than areas of lower susceptibility; the null hypothesis states that no difference exists in the amount of trees attacked between levels of susceptibility. In order to test the difference between susceptibility levels, the values $\mu(Su)$ from 0.0 to 1.0 were divided into four levels of susceptibility based on equal intervals: *zero susceptibility*, $\mu(Su)=0.0$; *low* susceptibility, $\mu(Su)\in[0.01,0.34]$; *medium* susceptibility, $\mu(Su)\in[0.35,0.67]$; and *high* susceptibility, $\mu(Su)\in[0.68,1.00]$. The observed attack locations in areas of high susceptibility were compared with areas of medium susceptibility, areas of medium susceptibility were compared with areas of low susceptibility, and low susceptibility was compared to zero susceptibility. The test statistic is described as

$$z = \frac{p_1 - p_2}{\sigma_{(p_1 - p_2)}} \tag{6}$$

where p_1 and p_2 are the proportion of trees attacked in the adjacent susceptibility levels. Therefore, for the first test p_1=high susceptibility and p_2=medium susceptibility; for the second test p_1=medium susceptibility and p_2=low susceptibility; and for the third test p_1=low susceptibility and p_2=zero susceptibility. The denominator $\sigma_{(p_1 - p_2)}$ represents the standard deviation of the sampling distribution.

The null hypothesis for a test is rejected if the test statistic produces a *z*-score greater than 1.96 for a 95% confidence level.

The entire procedure for this case study was repeated replacing the T_i attack locations raster layer with the T_{i+1} attack locations layer to develop a susceptibility map for T_{i+2}.

3. Results

3.1 *Fuzzy membership functions*

The values for d_1 and d_2 were assigned to each variable based on expert opinion and information from research literature. For membership in $\mu(LP)$, d_1 =0 and d_2=1 were selected to represent a 0.1 increase in susceptibility with every 10% increase in lodgepole pine trees in the stand. Therefore, the influence of the presence of host trees was directly proportional to the number of host trees present.

With regards to membership in $\mu(AT)$, this study used information from a collection of resources to derive a j-shaped function where d_1=50 m and d_2=60 m. This function explains that susceptibility due to dispersal behaviour is high over the first 50 m, and then decreases non-linearly over the remaining dispersal range.

The values for d_1 and d_2 for $\mu(DS)$ were determined from an analysis of ground truth data. Large deciduous stands were defined as any stand containing only deciduous trees that is at least 17 m^2 in size. Figure 5 reveals the relationship between the number of trees attacked and distance to large deciduous stands. This information suggests that for membership in $\mu(DS)$, d_1=0 because the attacked cell itself is not susceptible, and d_2=50 m as at this distance observations of attack become independent from the distance to large deciduous stands.

Finally, information regarding membership in $\mu(LT)$ was gathered from the literature and also collected from ground truth data (see figure 6). As previous literature suggests (Amman and Baker 1972, Shore and Safranyik 1992, Preisler and Mitchell 1993, Mata *et al.* 2003), larger trees were attacked more frequently than

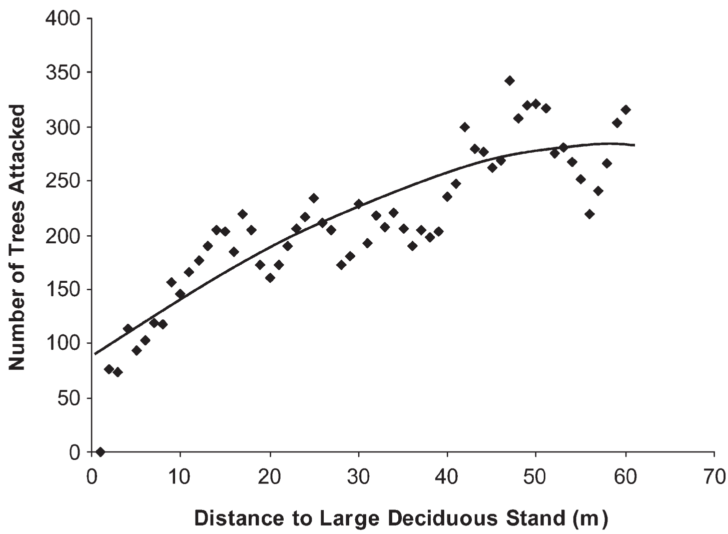

Figure 5. Relationship between number of trees that were observed to be attacked in the remote sensing images and distance to large deciduous stands.

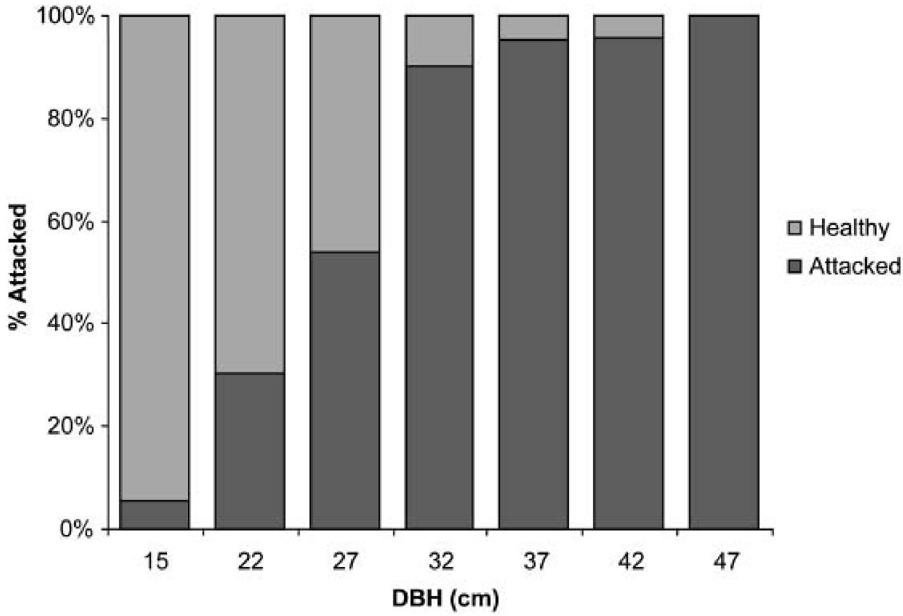

Figure 6. Ground truth data for the proportion of trees that received heavy and moderate attack by mountain pine beetle per size class based on diameter of the tree at breast height (DBH).

smaller trees. Although very large trees were under-represented in the ground truth data, there did exist an obvious linear increase for those age classes that were sufficiently represented. This suggests a linear fuzzy membership function for $\mu(LS)$. No trees in the ground truth data less than 15 cm in diameter at breast height were attacked, therefore $d_1=14$ cm. All trees in the ground truth data larger than 46 cm in diameter were attacked, therefore $d_2=47$ cm. This information was used to determine a single value for the susceptibility of a stand $\mu(LS)$ depending on the overall size of trees in the stand. Figure 7 illustrates the proportion of the different tree sizes in stands defined as high density, medium density and low density. The results from the GIS overlay procedure produced the fuzzy susceptibility maps for T_{i+1} (2001) and T_{i+2} (2002) shown in figure 8. These maps were then used with the original remote sensing data for model validation.

3.2 *Susceptibility map validation*

In order to validate the model, the T_{i+1} and T_{i+2} attack layers were overlaid onto the appropriate susceptibility maps for both sites. Figure 9 provides a visual comparison between the estimated areas of susceptibility and the observed locations of attack. The locations of attack in the high resolution remote sensing images can be observed in figure 10. Small areas of each site were magnified to illustrate the patterns of mountain pine beetle attack over the two years. The images from 2002 show red trees that were killed by mountain pine beetles in 2001, and grey trees killed by mountain pine beetles before 2001. The 2003 images show red trees that were killed by mountain pine beetles in 2002, and grey trees killed by mountain pine

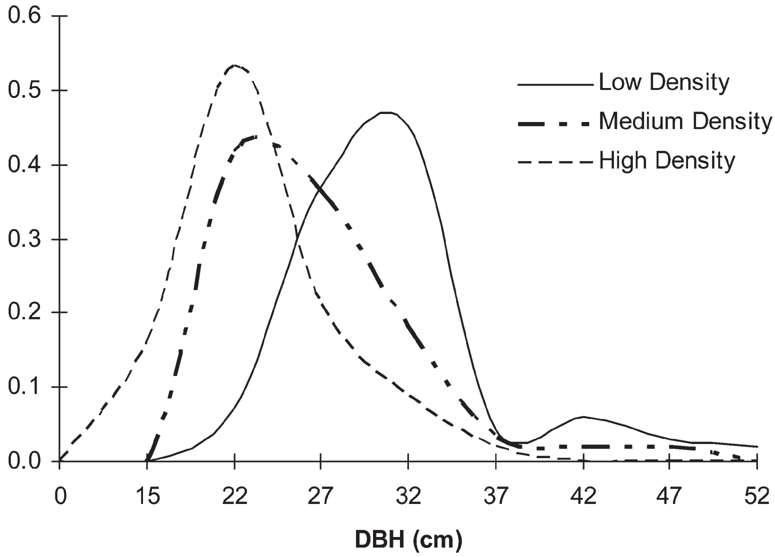

Figure 7. Frequency of tree sizes for visually defined levels of stand density.

beetles before 2002. The images illustrate that larger lodgepole pine trees are attacked first and more frequently as time proceeds.

Next, the proportion of attacked cells in each susceptibility class was tabulated. Figure 11 shows the proportion of cells attacked in each susceptibility class, which indicated that the high susceptibility class had the highest proportion of attack, and the proportion gradually decreased with a decrease in susceptibility rating. The one exception was for the zero susceptibility class for Site 2 at T_{i+2}.

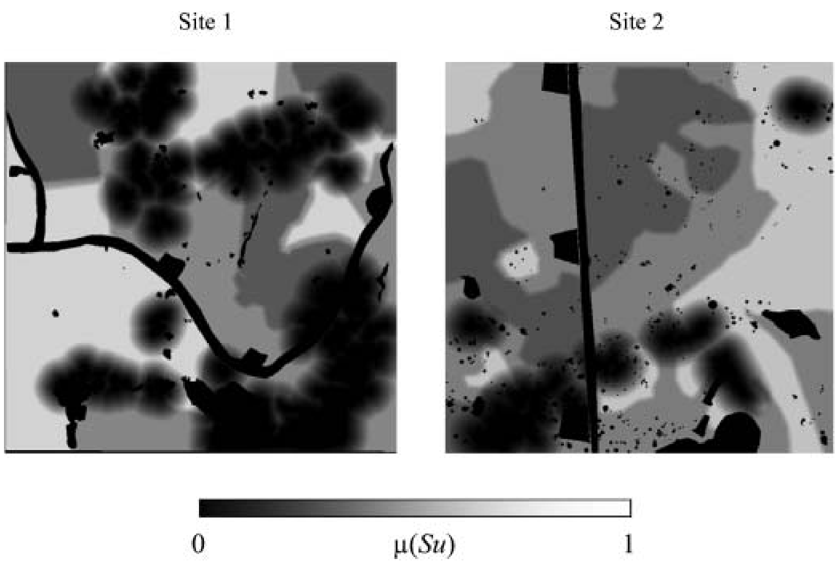

Figure 8. Susceptibility μ(*Su*) maps, resulting from fuzzy operations (see figure 2) for Site 1 and Site 2. Dark tones indicate no susceptibility (i.e. roads and clear cuts) and light tones indicate high susceptibility.

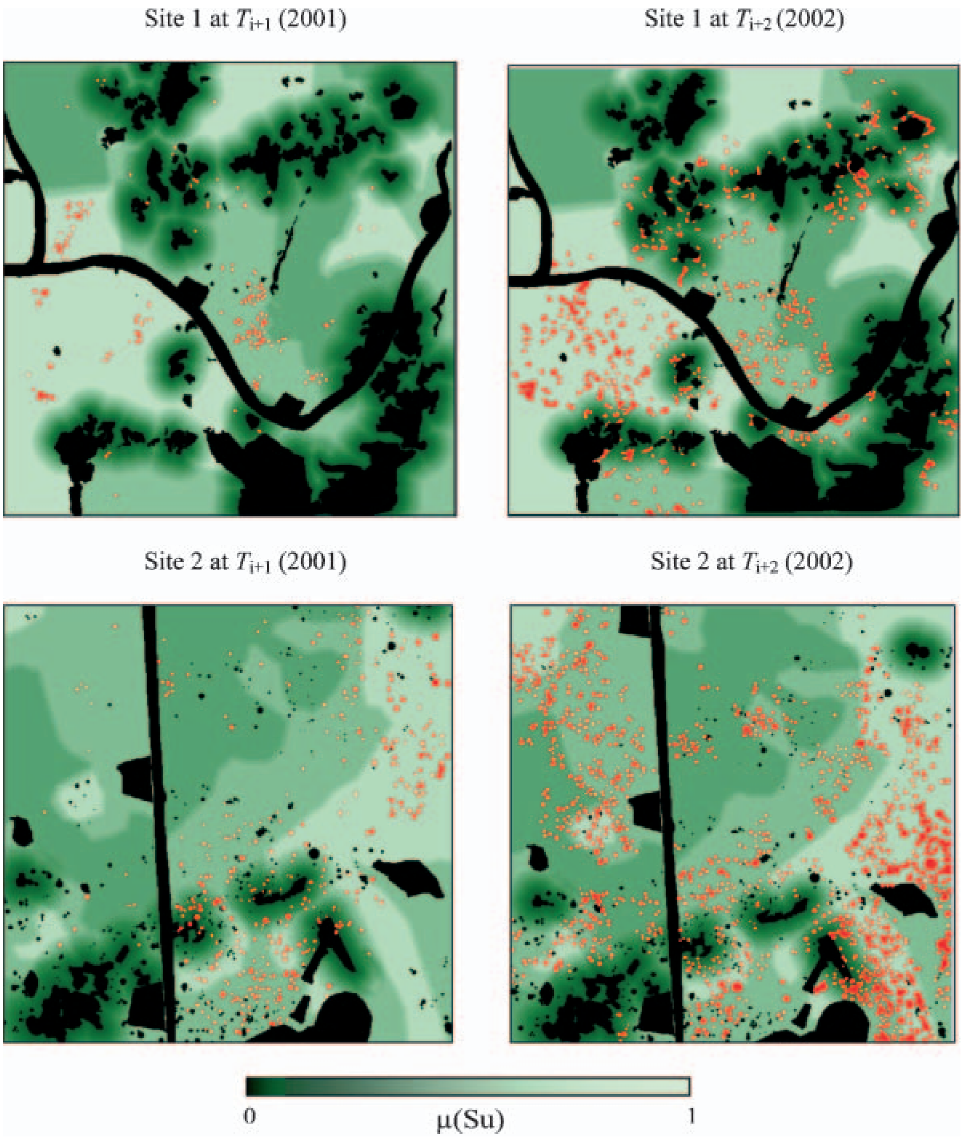

Figure 9. Visual validation of susceptibility maps (shown in figure 7) using observed attack locations from aerial imagery for Site 1 – T_{i+1}, Site 1 – T_{i+2}, Site 2 – T_{i+1}, Site 2 – T_{i+2}. Red cells indicate trees attacked by mountain pine beetle. Indicated years are for the year of attack which was prior to the year of detection on the remote sensing imagery.

The significance of these observations was statistically evaluated in the test for the hypothesis (p_1-p_2). The z-scores from the test are presented in table 1. In order to reject the null hypothesis with 95% confidence, z>1.96. These results indicate that most of the higher susceptibility classes contain a significantly greater proportion of cells that were attacked by mountain pine beetle than lower susceptibility classes, with the exception of the difference between the medium and low susceptibility classes and low and zero susceptibility classes for Site 2 in T_{i+1}. Therefore, the null hypothesis was rejected for 10 of the 12 tests.

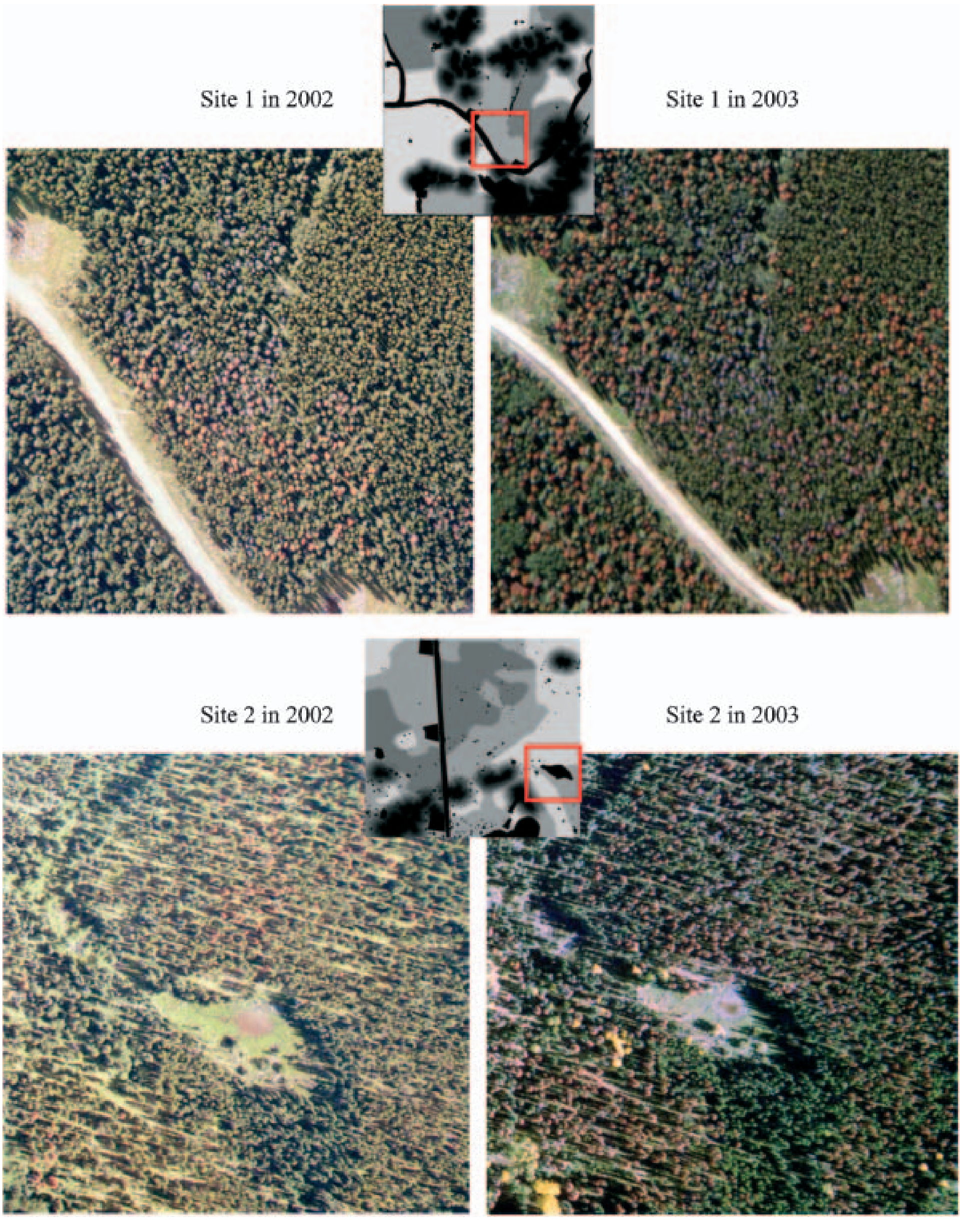

Figure 10. Magnified aerial imagery of Site 1 and Site 2 from 2002 and 2003. For the 2002 imagery, trees attacked in 2001 appear red; trees attacked before 2001 appear grey. For the 2003 imagery, trees attacked in 2002 appear red; trees attacked before 2002 appear dark red or grey. Different stands can be visually distinguished based on tree size; larger trees are more susceptible and therefore attacked earlier and more frequently than smaller trees.

4. Conclusion

The results indicate that the methods in this study were mostly successful at classifying levels of susceptibility to mountain pine beetle attack in the two sites. An issue of concern exists, however, with the locations of attack in areas of zero and low

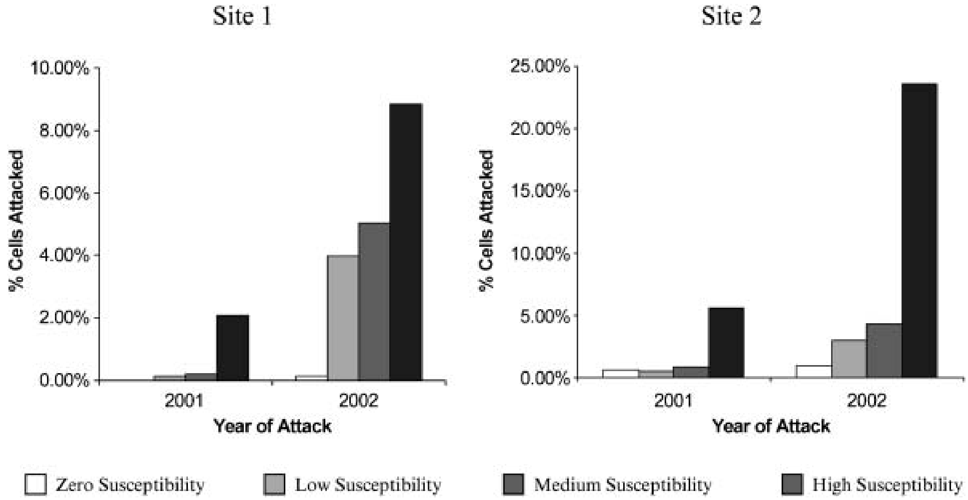

Figure 11. Proportion of cells attacked in each susceptibility class at T_{i+1} (2001 attack) and T_{i+2} (2002 attack) for Site 1 and Site 2.

Table 1. Proportion scores (z) for the *large-sample test comparing two proportions* (McClave and Sincich 2000). The z values indicate the proportional difference of mountain pine beetle attacked cells in adjacent susceptibility classes. With 95% certainty, higher susceptibility classes experience significantly more attack if $z>1.96$.

	T_{i+1} (2001)	T_{i+2} (2002)
Site 1		
High vs Medium	46.33	31.07
Medium vs Low	4.09	12.76
Low vs Zero	14.13	79.44
Site 2		
High vs Medium	12.77	165.81
Medium vs Low	1.33	15.34
Low vs Zero	−0.23	22.47

susceptibility for Site 2 at T_{i+2}, which were not significantly different from each other. A visual analysis of the original remote sensing images confirms that pine beetle attack did occur in areas estimated as zero susceptibility, which were mostly single or small groups of lodgepole pine trees that were located within large deciduous stands. Due to digitizing errors, these trees were not properly identified. Furthermore, there appeared to be several attacked trees that were located on the perimeter and also within open areas. It should also be noted that the presence of attacked trees in the low susceptibility areas did not come as a surprise because it was expected that trees of lower susceptibility will eventually be attacked once higher susceptibility trees are killed (Safranyik *et al.* 1999a).

The quality of the model stems from the ability to identify the degree of susceptibility of various areas coupled with acknowledging the uncertainty involved in the model development. Although some may argue that probability theory could have been used to produce values from 0 to 1, fuzzy set theory was considered necessary in order to deal with the level of available data and available information for defining susceptibility. The spatial and temporal extent of the data was a main contributing factor in limiting the knowledge of mountain pine beetle attack behaviour. With regards to the spatial extent, small study sites make it difficult to determine if there were attacked cells adjacent to the site boundary that could play an important factor in determining susceptibility. Furthermore, a limited spatial extent made the presence of clear-cut patches more significant. Clear-cuts occurred in some areas between the two years when the images were collected, which removed important data components.

The fact that high resolution remote sensing data on mountain pine beetle infestations exists for a limited temporal frame limits the ability to monitor and study insect infestations over time. Long-term projections of susceptibility can be restricted by lack of information regarding insect locations for each year. The success of the model proposed in this study was partly due to knowing the locations of insects from the ground truth data in the year previous to when susceptibility was estimated. Susceptibility maps can, however, play a significant role in long-term forecasting when used as data sources in spatial–temporal models. These models would require greater knowledge of insect population dynamics to estimate their infestation behaviour over periods of time. Spatial–temporal models are commonly used in analysing vegetation dynamics, but their use with datasets based on fuzzy sets has been largely unexplored.

Applications for developing susceptibility maps with fuzzy sets can go beyond insect dispersal to include phenomena such as susceptibility to wildfire, wind, diseases and invasive species by having adequate data and information regarding the development of these specific membership functions. Integrating fuzzy set theory with GIS can also act as a decision support tool for forest management as landscapes can be digitally manipulated to find optimal practices in the light of potential disturbances.

Acknowledgments

The first and second authors are thankful to the Natural Sciences and Engineering Research Council (NSERC) of Canada for full support for this study under the Discovery Grant Program. Acquisitions of high resolution datasets used in this study are funded from BC Forestry Innovation and Forestry Investment Account grants awarded to the third author. We gratefully acknowledge the research assistance provided by Jim Northrup, Yinbin Li, Richard Reich and Sebastian Wolf.

References

AMMAN, G.D. and BAKER, B.H., 1972, Mountain pine beetle influence on lodgepole pine stand structure. *Journal of Forestry*, **70**, pp. 204–209.

BEZDEK, J.C., EHRLICH, R. and FULL, W., 1984, FCM: The fuzzy c-means clustering algorithm. *Computers and Geosciences*, **10**, pp. 191–203.

BRANDTBERG, T., 2002, Individual tree-based species classification in high spatial resolution aerial images of forests using fuzzy sets. *Fuzzy Sets and Systems*, **132**, pp. 371–387.

BROWN, D.G., 1998, Mapping historical forest types in Baraga County Michigan, USA as fuzzy sets. *Plant Ecology*, **134**, pp. 97–111.

BURROUGH, P.A., 1989, Fuzzy mathematical methods for soil survey and land evaluation. *Journal of Soil Science*, **40**, pp. 477–492.

BURROUGH, P.A., 1996, Natural objects with intermediate boundaries. In *Geographic Objects with Intermediate Boundaries*, A.U. Frank (Ed.), pp. 177–188 (London: Taylor and Francis).

BURROUGH, P.A., MACMILLAN, R.A. and DEURSEN, W.V., 1992, Fuzzy classification methods for determining land suitability from soil profile observations and topography. *Journal of Soil Science*, **43**, pp. 193–210.

CEVIK, E. and TOPAL, T., 2003, GIS-based landslide susceptibility mapping for a problematic segment of the natural gas pipeline, Hendek (Turkey). *Environmental Geology*, **44**, pp. 949–626.

DAVIDSON, D.A., THEOCHAROPOULOS, S.P. and BLOKSMA, R.J., 1994, A land evaluation project in Greece using GIS and based on Boolean and fuzzy set methodologies. *International Journal of Geographical Information Systems*, **8**, pp. 369–384.

DRAGICEVIC, S. and MARCEAU, D.J., 2000, A fuzzy set approach for modelling time in GIS. *International Journal of Geographical Information Science*, **14**, pp. 225–245.

FISHER, P., 1999, *Models of Uncertainty in Spatial Data. Geographic Information Systems: Principles and Applications* (New York: John Wiley).

FOODY, G.M. and ATKINSON, P., 2002, *Uncertainty in Remote Sensing and GIS* (Chichester: John Wiley).

FRANKLIN, S.E., WULDER, M.A., SKAKUN, R.S. and CARROLL, A.L., 2003, Mountain pine beetle red-attack damage classification using stratified Landsat TM data in British Columbia, Canada. *Photogrammetric Engineering and Remote Sensing*, **69**, pp. 283–288.

GIMBARZEVSKY, P., DAWSON, A.F. and SICKLE, G.A.V., 1992, Assessment of aerial photographs and multi-spectral scanner imagery for measuring mountain pine beetle damage. Information Report BC–X–333, Canadian Forestry Centre, Victoria, British Columbia.

HINDMARCH, T.D. and REID, M.L., 2001, Forest thinning affects reproduction in pine engravers (Coleoptera: Scolytidae) breeding in felled lodgepole pine trees. *Environmental Entomology*, **30**, pp. 919–924.

LOWELL, K., 1996, Modelling heterogeneity and change in natural forests. *Geomatica*, **50**, pp. 425–440.

LOWELL, K. and GOLD, C., 1995, Using a fuzzy surface-based cartographic representation to decrease digitizing efforts to natural phenomena. *Cartography and Geography Information Systems*, **22**, pp. 222–231.

LUCIEER, A. and KRAAK, M.-J., 2004, Interactive and visual fuzzy classification of remotely sensed imagery for exploration of uncertainty. *International Journal of Geographical Information Science*, **18**, pp. 491–512.

LUTHER, J.E., FRANKLIN, S.E., HUDAK, J. and MEADES, J.P., 1997, Forecasting the susceptibility and vulnerability of balsam fir stands to insect defoliation with Landsat Thematic Mapper data. *Remote Sensing of Environment*, **59**, pp. 77–91.

MATA, S.A., SCHMID, J.M. and OAKES, R.D., 2003, Growth of lodgepole pine stands and its relation to mountain pine beetle susceptibility. Research Paper RMRS-RP-42, USDA Forest Service, Washington, DC.

MCCLAVE, J.T. and SINCICH, T., 2000, *Statistics* (Upper Saddle River, NJ: Prentice-Hall).

MITCHELL, R.G. and PREISLER, H.K., 1991, Analysis of spatial patterns of lodgepole pine attacked by outbreak populations of the mountain pine-beetle. *Forest Science*, **37**, pp. 1390–1408.

MOECK, H.A. and SIMMONS, C.S., 1991, Primary attraction of mountain pine beetle, *Dendroctonus poderosae* Hopk. (Coleoptera: Scolytidae), to bolts of lodgepole pine. *Canadian Entomologist*, **123**, pp. 299–304.

PERKINS, D.L. and ROBERTS, D.W., 2003, Predictive models of whitebark pine mortality from mountain pine beetle. *Forest Ecology and Management*, **174**, pp. 495–510.

PREISLER, H.K. and MITCHELL, R.G., 1993, Colonization patterns of the mountain pine beetle in thinned and unthinned lodgepole pine stands. *Forest Science*, **39**, pp. 528–545.

RIEL, W.G., FALL, A., SHORE, T.L. and SAFRANYIK, L., 2004, A spatio-temporal simulation of mountain pine beetle impacts on the landscape. In *Mountain Pine Beetle Symposium: Challenges and Solutions*, Natural Resources Canada, Canadian Forestry Services, Pacific Forestry Centre, Information Report BC–X–399, Kelowna, British Columbia, pp. 106–113.

ROBERTS, A., DRAGICEVIC, S., NORTHRUP, J., WOLF, S., LI, Y. and COBURN, C., 2003, Mountain pine beetle detection and monitoring: remote sensing evaluations. BC Forest Innovation Investment Operational Research Report with reference to Recipient Agreement R2003-0205. Available online at: fii.forrex.org/reports/R2003-205.pdf (accessed 6 September 2005).

ROBINSON, V.B., 2003, A perspective on the fundamentals of fuzzy sets and their use in geographical information systems. *Transactions in GIS*, **7**, pp. 3–30.

SAFRANYIK, L., 1988, Mountain pine beetle: biology overview. In *Symposium on the Management of Lodgepole Pine to Minimize Losses to the Mountain Pine Beetle*, Kelispell, USA, 9–13 July 2003, pp. 9–13 (US Department of Agriculture).

SAFRANYIK, L., BARCLAY, H., THOMSON, A. and RIEL, W.G., 1999a, A population dynamics model for the mountain pine beetle, *Dendroctonus Ponderosae* Hopk. (Coleoptera: Scolytidae). Information Report BC–X–326, Pacific Forestry Centre, Victoria, British Columbia.

SAFRANYIK, L., SHORE, T.L. and LINTON, D.A., 1999b, Attack by bark beetles (Coleoptera: Scolytidae) following spacing of mature lodgepole fine (Pinaceae) stands. *Canadian Entomologist*, **131**, pp. 671–685.

SAFRANYIK, L., SILVERSIDES, L.H., MCMULLEN, L.H. and LINTON, D.A., 1989, An empirical approach to modeling the local dispersal of the mountain pine beetle (*Dendroctonus ponderosae* Hopk.) (Col., Scolytidae) in relation to sources of attraction, wind direction and speed. *Journal of Applied Entomology*, **108**, pp. 498–511.

SARKAR, S. and KANUNGO, D.P., 2004, An integrated approach for landslide susceptibility mapping using remote sensing and GIS. *Photogrammetric Engineering and Remote Sensing*, **70**, pp. 617–625.

SHORE, T.L. and SAFRANYIK, L., 1992, Susceptibility and risk rating systems for the mountain pine beetle in lodgepole pine stands. Information Report BC–X–336, Pacific Forestry Centre, Victoria, British Columbia.

SHORE, T.L., SAFRANYIK, L. and LEMIEUX, J.P., 2000, Susceptibility of lodgepole pine stands to the mountain pine beetle: testing of a rating system. *Canadian Journal of Forest Research—Revue Canadienne De Recherche Forestiere*, **30**, pp. 44–49.

SKAKUN, R.S., WULDER, M.A. and FRANKLIN, S.E., 2003, Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage. *Remote Sensing of Environment*, **86**, pp. 433–443.

THOMSON, A., 1991, Simulation of mountain pine beetle (*Dendroctonus ponderosae* Hopkins) spread and control in British Columbia. Information Report BC–X–329, Pacific Forestry Centre, Victoria, British Columbia.

VAN WESTEN, C.J., 2000, The modelling of landslide hazards using GIS. *Surveys in Geophysics*, **21**, pp. 241–255.

WANG, F., 1990, Improving remote-sensing image analysis through fuzzy information representation. *Photogrammetric Engineering and Remote Sensing*, **56**, pp. 1163–1169.

WANG, F. and HALL, G.B., 1996, Fuzzy representation of geographical boundaries in GIS. *International Journal of Geographical Information Systems*, **10**, pp. 573–590.

WULDER, M.A. and DYMOND, D., 2004, Remote sensing technologies for mountain pine beetle survey. In *Mountain Pine Beetle Symposium*, Natural Resources Canada, Canadian Forestry Services, Pacific Forestry Centre, Information Report BC–X–399, Kelowna, British Columbia, pp. 146–153.

ZADEH, L.A., 1965, Fuzzy sets. *Information and Control*, **8**, pp. 338–353.

ZADEH, L.A., 1972, A fuzzy-set-theoretic interpretation of linguistic hedges. *Journal of Cybernetics*, **2**, pp. 4–34.

ZHANG, J. and FOODY, G.M., 1998, A fuzzy classification method of sub-urban land cover from remotely sensed imagery. *International Journal of Remote Sensing*, **19**, pp. 2721–2738.

ZHANG, J. and GOODCHILD, M.F., 2002, *Uncertainty in Geographical Information* (London: Taylor and Francis).