

# Spatial Analysis of High Resolution Aerial Photographs to Analyze the Spread of Mountain Pine Beetle Infestations

Eric Fox<sup>1</sup>, Shivanand Balram<sup>1</sup>, Suzana Dragicevic<sup>1</sup> & Arthur Roberts<sup>1</sup>

<sup>1</sup> Department of Geography, Simon Fraser University, Burnaby, Canada

Correspondence: Shivanand Balram, Department of Geography, Simon Fraser University, 8888 University Drive, Burnaby, BC, V5A 1S6, Canada. Tel: 1-778-782-3702. E-mail: sbalram1@sfu.ca

Received: July 23, 2012 Accepted: August 10, 2012 Online Published: August 28, 2012

doi:10.5539/jsd.v5n9p106

URL: <http://dx.doi.org/10.5539/jsd.v5n9p106>

## Abstract

Mountain pine beetle (*Dendroctonus Ponderosae*) infestations in Western Canada have reached alarming proportions. The spread of attacks has significantly impacted pine tree stocks, the forest ecosystem in general and the overall socio-economic condition of residents in communities that depend on the forest industry. The ability to track these attacks and anticipate their trajectories will allow forest managers to target intervention measures and improve mitigation efforts. Spatial analysis methods are well suited to characterize the dynamics of the mountain pine beetle attacks. This study implemented a multi-scale spatiotemporal analysis to describe the patterns of three consecutive attack years (2000-2002) for a 19 km<sup>2</sup> study site in British Columbia, Canada. In the first stage, baseline complete spatial randomness (CSR) was examined using average nearest neighbour analysis and quadrat counts on infestation data extracted from 19 cm resolution black and white aerial photographs. In the second stage, global and local Moran's *I* spatial autocorrelation statistics, Getis-Ord General *G* global high/low clustering and Getis-Ord  $G_i^*$  test statistic techniques were used to evaluate the spread patterns over time. The results show a significant clustering of attack sites in the first two years, with considerable changes in the attack spread in the third year. These changes were independent of the two levels of spatial resolution used. Knowledge about infestation patterns can be used to anticipate needed social and environmental assistance for affected communities. Further, the information can aid overall sustainable forestry management.

**Keywords:** aerial photographs, cluster and outlier analysis, lodgepole pine, mountain pine beetle infestations, spatial analysis, spatial autocorrelation, temporal spread

## Resumen

Las infestaciones del escarabajo del pino de montaña (*Dendroctonus Ponderosae*) en el oeste de Canadá han alcanzado proporciones alarmantes. La propagación de los ataques ha afectado considerablemente las reservas de pinos, el ecosistema forestal en general y la condición socioeconómica de los residentes en las comunidades que dependen de la industria maderera. La capacidad de conocer y anticipar la trayectoria de estos ataques permitirá a los administradores forestales tomar las medidas necesarias para intervenir y mejorar los esfuerzos de mitigación. Los métodos de análisis de los patrones espaciales o recorridos de los escarabajos del pino de montaña son adecuados para caracterizar las dinámicas espaciales de la propagación de éstos. Esta investigación implementó una escala múltiple de análisis espaciotemporal para describir los patrones de ataque en tres años consecutivos (2000-2002) en un área de estudio de 19 km<sup>2</sup> en Columbia Británica, Canadá. En la primera etapa, se examinó el modelo de aleatoriedad espacial completa a través del análisis del promedio vecino más cercano y los conteos cuadrados de los datos extraídos de fotografías aéreas en blanco y negro de 19 cm de resolución. En la segunda etapa, las estadísticas de autocorrelación espacial de la *I* de Moran a nivel global y local, Getis-Ord General *G* global de alto/bajo agrupamiento y las técnicas de prueba estadística Getis-Ord  $G_i^*$  se utilizaron para evaluar los patrones de dispersión a través del tiempo. Los resultados muestran que existe un agrupamiento significativo de las áreas de ataque en los dos primeros años con cambios considerables en la propagación en el tercer año. Estos cambios son independientes de los dos niveles de resolución espacial utilizados. El conocimiento sobre los patrones de infestaciones del escarabajo del pino de montaña sirve para anticipar la asistencia social necesaria para las comunidades afectadas y al medio ambiente. Además, la información puede ayudar a la administración forestal sostenible en general.

**Palabras clave:** fotografías aéreas, análisis de agrupamiento y valor atípico, pino contorto, infestaciones del escarabajo del pino de montaña, análisis espacial, autocorrelación espacial, propagación temporal

## 1. Introduction

### 1.1 Background

Forest insect infestations create a disruptive footprint on the natural and human environment. In western North America, the mountain pine beetle (*Dendroctonus Ponderosae*) has been the main focus of attention. Regionally, the forested areas of the central interior of British Columbia, Canada including the Cariboo, Bulkley-Nechako and Fraser-Fort George regional districts have seen a dramatic spread of mountain pine beetle infestations with some of the most concentrated areas being around Prince George and Quesnel (Nelson et al., 2007; Safranyik & Wilson, 2006; Wulder et al., 2006). Locally, communities in rural British Columbia have witnessed significant social and economic changes as large acres of forest stocks have become depleted from widespread mountain pine beetle attacks (Safranyik & Wilson, 2006). Research to understand the spread patterns and processes of the mountain pine beetle is important because of the environmental impact on the pine tree species and local ecosystems as well as the financial impact on the livelihood of communities that depend on forest resources (Safranyik & Wilson, 2006; Wulder et al., 2006). The capability to map these attacks and anticipate their spatial and temporal trajectories will allow forest managers to target high risk communities for more effective sustainable intervention measures.

The objective of this study is to identify and quantify the spatial character and distribution of mountain pine beetle attacks on lodgepole pine stands at a case study site in British Columbia, Canada. Spatial analysis techniques are applied at multiple spatial scales to infestation data for three consecutive years using a vector-based geographic information system (GIS). Insights about the spread dynamics are inferred from the infestation patterns by comparing changes in the spatial distributions of infected trees over time, and evaluating the extent to which changes in the specified spatial scale influence the output results.

Mountain pine beetle attacks evolve in well established seasonal stages. In the early stages of an infestation, affected trees appear in highly clustered groups and as mountain pine beetle populations increase in number over consecutive growing seasons, the attacks coalesce to form larger patches (spot growth) and spread out to other locations (spot proliferation) (Safranyik & Wilson, 2006). Additionally, attack sites are based on the length of flight time of the species, the ability of trees to resist attacks, and the return of beetle populations to partially attacked trees (stripped attacks) from previous seasons (Safranyik & Wilson, 2006). Furthermore, host selection depends on a series of biological processes of the beetle including a mix of random landings, visual orientation of tree suitability and the odors of host material (Safranyik & Wilson, 2006). Due to the relative complexity of attack sites and beetle ecology, it can be to accurately predict attack locations (Safranyik & Wilson, 2006). The research literature has used mostly spatial analysis and modeling techniques to analyze the infestations.

In this study, spatial analysis techniques, as a core component of geographic information science, are used to describe and explain the spatial distribution and organization of pine beetle infestations across the landscape. Quantitative indices are available to characterize and classify particular distributions and patterns formed by a phenomenon (Boots & Getis, 1988). These indices assist in linking the ecology and behavior of the events being investigated. In addition, techniques to measure spatial autocorrelation and spatial dependence in environmental data have emerged for pattern recognition and quantitative description of the relationships between the location of the phenomena and associated attribute values (Haining, 2003; Overmars et al., 2003). Furthermore, multi-scale and multi-temporal data are crucial in spatial analysis procedures because they aid in understanding how geographic processes change over time and also indicate the stability of the phenomena at varying spatial resolutions (Nelson et al., 2007; Overmars et al., 2003). These tools enable spatial patterns to be easily discernible and support interpretations within the study area context. The results add to our understanding of the mountain pine beetle behavior, and quantify the infestation spread patterns in a real landscape. The following sections present an overview of the spatial analysis techniques, describe the study site and methods used, present the results and interpretations, and provide conclusions.

### 1.2 Spatial Data and Analysis

The data for spatial analysis and modeling can be obtained from a variety of sources including remotely-sensed imagery converted for use in a GIS which has been the main approach for research on mountain pine beetle infestations (Bone et al., 2005; Nelson et al., 2007; Wulder et al., 2006). Remote sensing technology, including satellite imagery and aerial photography, have been widely used in recent years as a pragmatic tool to observe provincial mountain pine beetle movement in order to map the size of epidemics and develop solutions to prevent their spread (Coggins et al., 2008; Wulder et al., 2006). The use of high resolution aerial imagery in

particular has been useful in monitoring tree stand health in the face of severe and widespread pine beetle attacks (Bone et al., 2006; Coggins et al., 2008; Erfanifard et al., 2009; Wulder et al., 2006).

A specific implementation of spatial analysis is spatial pattern analysis, which is a set of techniques that are either pattern descriptors or pattern analyzers. Pattern descriptors use spatial statistics (e.g. mean center, standard distance) to summarize the distribution of data locations. Pattern analyzers on the other hand use quantitative measurements (e.g. nearest neighbour, quadrat count, K-function) and indices (e.g. Getis-Ord General G, Getis-Ord  $G_i^*$ , Moran's I, Geary C) to evaluate if the data distribution is random, dispersed or clustered. In this study, various pattern analyzers were used to examine connections between the spatial locations and attribute values of the mountain pine beetle infestations.

The first stage of spatial pattern analysis explores overall patterns using indices developed from the quadrat analysis and the nearest neighbour analysis methods. While these approaches have limitations such as sensitivity to quadrat specification and study area boundary definition, they provide a first approximation comparison between an observed point pattern and one generated from a complete spatial randomness (CSR) process. For confirmation, the Ripley's  $K$ -function is used to discern patterns based on the spatial location of features at multiple distance-band intervals. The Ripley's  $K$ -function is often used to study the spatial distribution of plants and insects in order to explain ecological processes (Spooner et al., 2004; Vadrevu & Badarinath, 2009). Some authors demonstrated that Ripley's  $K$ -function provides important information about the spatial organization of a variety of plant and animal ecology ranging from lowland shrubs and forest fire events to the movement of spruce budworms (Hering et al., 2009; Lynch & Moorcroft, 2008; Spooner et al., 2004; Vadrevu et al., 2008).  $K$ -function analysis is advantageous as it provides a measurement to test the extent of spatial clustering at defined distance intervals in order to overcome the shortcomings of the global nearest neighbour analysis (Lynch & Moorcroft, 2008). Despite the various  $K$ -function analysis approaches, such as the network  $K$ -function to avoid non-homogenous environments where Euclidean distance cannot be used, the method relies on user-defined distance intervals (Spooner et al., 2004).

The second stage of spatial pattern analysis considers the spatial attribute distributions and their inter-distance clustering. Beyond standard cluster and outlier analysis methods, the Getis-Ord General G index and the Getis-Ord  $G_i^*$  test statistic can be used to characterize the internal clustering within the data (Getis & Ord, 1992; Tobler, 1970). Getis-Ord cluster analysis procedures have been widely used to map the clustering of various plant species such as weeds at local and regional scales (Laffan, 2006; Mueller-Warrant et al., 2008). The General G clustering statistic operates at a global level in the study area and detects whether clustering is present within the data as well as indicating above average or below average clusters more strongly (Mueller-Warrant et al., 2008). The Getis-Ord General G test statistic measures the extent to which low and high value attribute data tend to cluster across the entire dataset while the Getis-Ord  $G_i^*$  statistic is a local measurement of hot spot and cold spot locations (Mueller-Warrant et al., 2008; Nelson et al., 2007). A key challenge for the General G global measure and the Getis-Ord  $G_i^*$  statistic is their susceptibility to edge effects at the study area boundaries causing an exaggeration of similarities and differences between the high and low values (Mueller-Warrant et al., 2008). Furthermore, other researchers have found spatial autocorrelation is dependent on the level of data aggregation and the scale of analysis (Odland, 1988; Overmars et al., 2003). Smoothed data obtained by averaging of increased aggregation level produces higher levels of autocorrelation for the Moran's  $I$  index (Overmars et al., 2003). However, the extent to which variation exists in the spatial dependency among the data based on relatively small changes in level of spatial resolution is less clear (Odland, 1988). Moreover, when data are aggregated to different scales acquired from medium or low resolution, coarser scale remote sensing imagery may have different changes in spatial autocorrelation than that of high resolution data (Odland, 1988; Overmars et al., 2003). The combination of the derived indices described previously, together with their assessment at multiple scales, assisted in describing the spatial distribution of mountain pine beetle attacks over time in order to derive evidence-based conclusions about spread patterns at the study site.

## 2. Method

### 2.1 Data and Study Site

The data were acquired from multispectral aerial photographs collected on 14<sup>th</sup> August 2002. This imagery was rectified to determine the relative geometry and absolute location for use in remote sensing and GIS processing software. Supervised classification methods were used to identify land use features within forested areas including open spaces, roads, forest tree types and infected trees in order to produce a classification map. The classification accuracy was determined to be suitable based on comparison with ancillary map data and expert opinion about ground features of the study site. In addition, near vertical camera orientation as well as the

relatively small study area and assumed flat ground topography reduced the need for image orthorectification. This classified data in GIS format forms the basis from which to conduct the spatial analyses techniques. Despite the availability of imagery at only one point in time, using manual interpretation and automated classification based on multispectral pixel values, the associated time of specific mountain pine beetle attacks was readily identified. Ground truthing procedures were also implemented to spatially reference and adjust the data for errors in preparation for GIS analysis.

The study site, Nazko, is an uninhabited forest area in the Interior Plateau of the northwestern Cariboo regional district (Figure 1). The coordinates of the study site centroid is 53°10'45" N and 123°26'20" W (or 470550 mE and 589250 mN, UTM Zone 10N). Topographic effects are assumed to be minimal. This site was selected because of its proximity to the First Nations settlement of Nazko located about 23 km to the southwest. The nearest moderately-sized community is the town of Quesnel, located approximately 67 km to the east of the site.

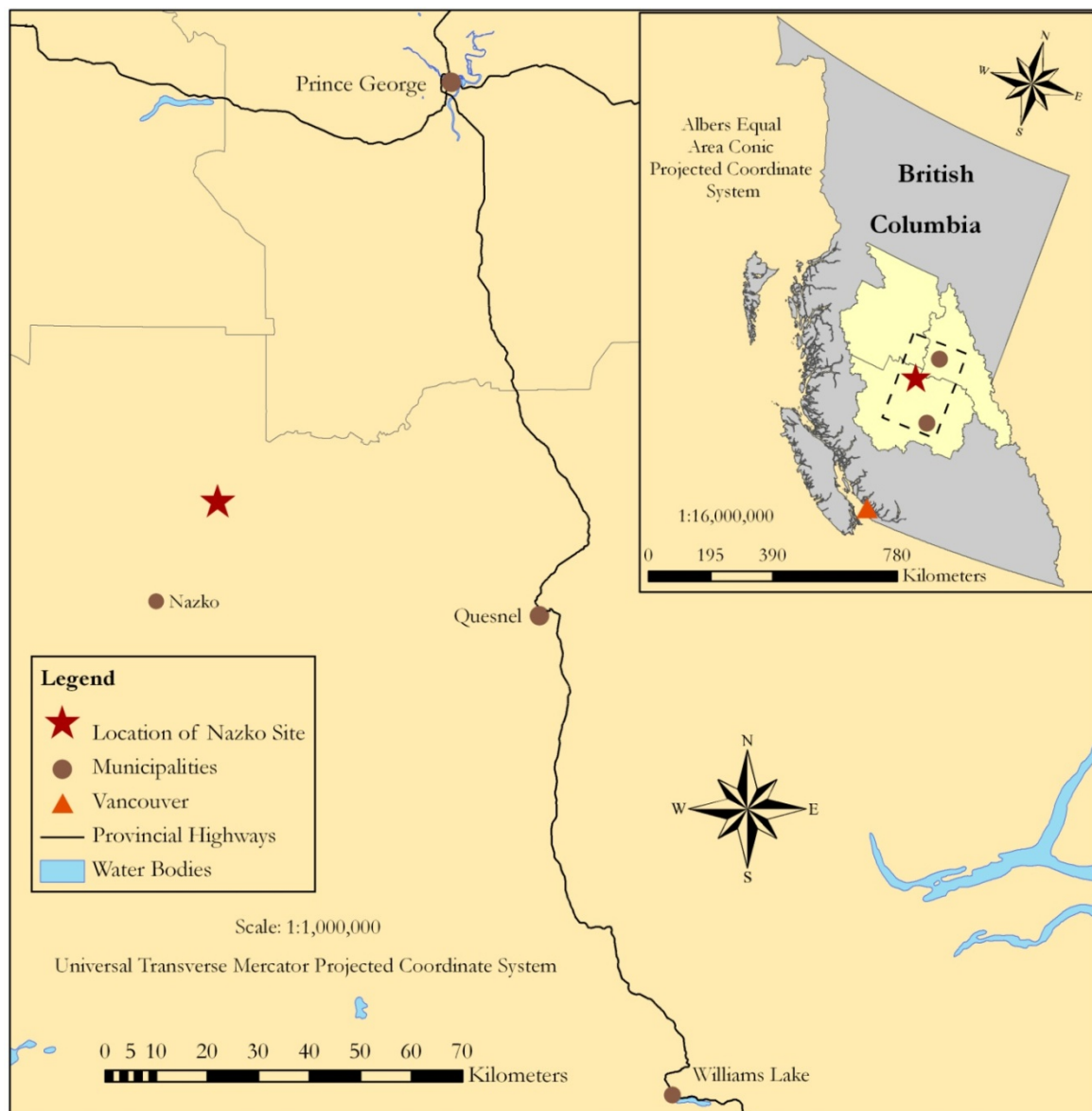


Figure 1. Map of the study site within the northern Cariboo of the Central Interior

## 2.2 Data Preparation

The original image acquisition resolution of 19.02 cm was preserved for the image processing using ERDAS ER Mapper 2010 and the spatial analysis using ESRI ArcGIS v.9.3. Further, a multi-scale approach was used to

compare and contrast differences in spatial analysis measures. Since the original imagery was captured at a very fine resolution and the study site was approximately 19.6 km<sup>2</sup>, a balance between image detail and computer processing performance resulted in the selection of an aggregation size of times four to conduct the secondary examination at a derived spatial resolution of about 76.09 cm. This scale size can provide meaningful comparison to the original scale without compromising the spatial heterogeneity of the study area (Nelson et al., 2007). In order to analyze the study area at an even coarser spatial resolution additional data sources and a larger study area would be required both of which were not available for this evaluation.

Infestation data for three years were interpreted from the single multispectral aerial photograph. Infected trees were characterized by the year in which the attack took place based on multispectral brightness values and reflectance color of infected trees in comparison to healthy ones (Figure 2). Since the original imagery of the site was obtained during the summer of 2002, infected trees were delineated by image pixel colour into the three category years of 2000, 2001 and 2002 based on when the mountain pine beetle attacks occurred in the past. Spatial data on this phenomenon has only been acquired over a limited period of time in response to rapid mountain pine beetle infestation over the past fifteen to twenty years. In addition, the remote, rural nature of the case study also makes data inventories expensive and challenging. As a result, in the case of the Nazko site, extensive data collections with considerable temporal resolution were not available. Nevertheless, in these sparse data situations, the indirect extraction of temporal information from the spectral signatures of single-time imagery data is a reliable surrogate for multi-year image data.

The original raster data was vectorized before defining attribute variables and implementing spatial analysis. The total number of infected trees per polygon was used as the single associated variable when performing each of the seven index measurements with the exception of the Ripley's *K*-function analysis and initial exploratory data analysis procedures. The total number of trees attacked was determined by combining supplementary ground truth data, area calculations of attack sites, and manual interpretation of the original aerial imagery.

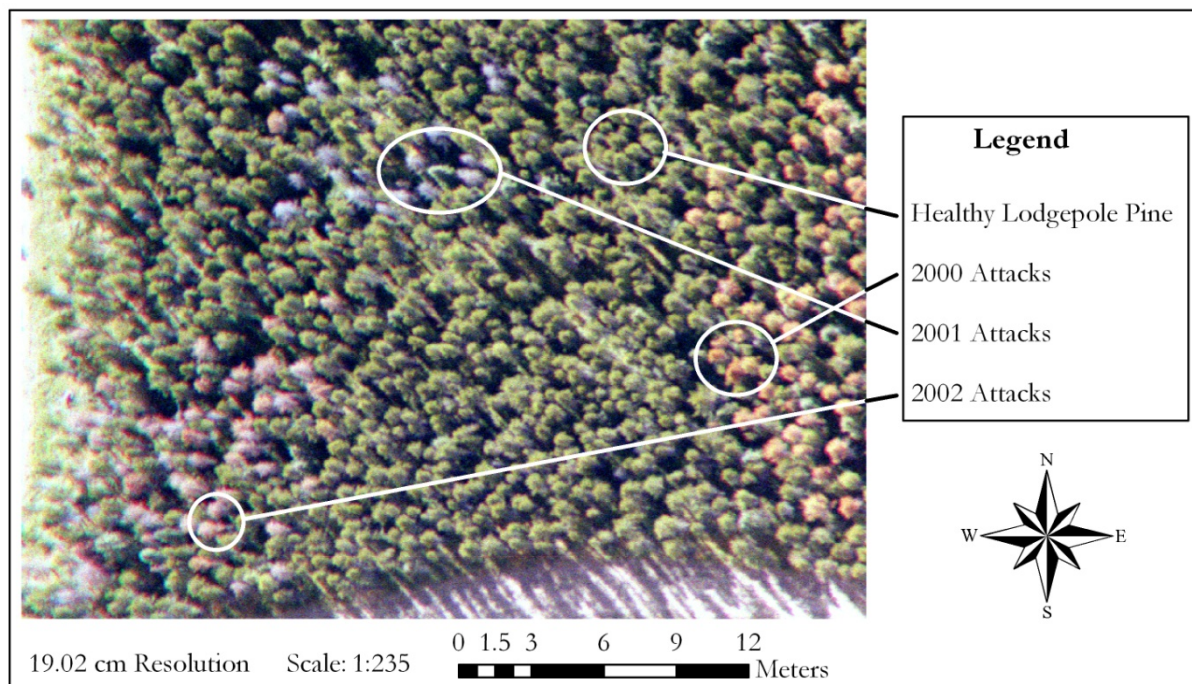


Figure 2. Sample area of the original visible band colour air photograph

### 2.3 Exploratory Spatial Data Analysis

A primary visual assessment for all three years and the two spatial resolutions showed a substantial change in the size of the attack sites between 2001 and 2002 (Figure 3). The attack site patches were not necessarily located in regions untouched in the previous two years, but the total area infested had more than doubled. Two types of exploratory data analysis measures were used to provide an initial baseline to describe the spatial arrangement of the attack sites. Average nearest neighbour analysis determine the spacing of features in a study site based on

measured distances between individual features and the nearest other feature (K. P. Vadrevu & Badarinath, 2009; K. P. Vadrevu et al., 2008). The mean nearest neighbour distance provides an index of spacing of features and is expressed as the ratio of the observed distance value divided by the expected distance value (Erfanifard et al., 2009; McGrew & Monroe, 2000). The standardized absolute nearest neighbour index ( $R$ ) is calculated as (McGrew & Monroe, 2000):

$$R = \frac{\overline{NND}}{\overline{NND}_R} \text{ and } \overline{NND}_R = \frac{1}{\sqrt{\frac{n}{A}}} \quad (1)$$

where  $\overline{NND}$  is the sum of all nearest neighbour distances divided by  $n$  identified as the total number of points. The average nearest neighbour distance for a random arrangement is  $\overline{NND}_R$ , where  $A$  represents the total area and  $n$  is the total number of point events.

Quadrat analysis measures the overall character of a point pattern distribution by calculating the frequency of points within a uniform area (Lloyd, 2007; Shiode, 2008). Instead of evaluating inter-point distances as in the case of nearest neighbour analysis, a square grid of cells is superimposed over the study site and point frequencies per cell are calculated (McGrew & Monroe, 2000; Shiode, 2008). An array consisting of 10 m x 10 m cells was selected as a suitable size to evaluate the 19.6 km<sup>2</sup> study area (Lloyd, 2007). The variance-mean ratio ( $VMR$ ) is calculated as (McGrew & Monroe, 2000):

$$VMR = \frac{VAR}{MEAN} \quad (2)$$

where  $VAR$  represents the variance of cell frequencies and  $MEAN$  represents the mean cell frequency and are calculated as (McGrew & Monroe, 2000):

$$VAR = \frac{\sum f_i X_i^2 - \left[ \frac{(\sum f_i X_i)^2}{m} \right]}{m-1} \text{ and } MEAN = \frac{n}{m} \quad (3)$$

where  $f_i$  is the frequency of cells with  $i$  attack sites,  $X_i$  is the number of attack sites per cell,  $n$  is the total number of attack sites and  $m$  is the total number of cells in the quadrat array.



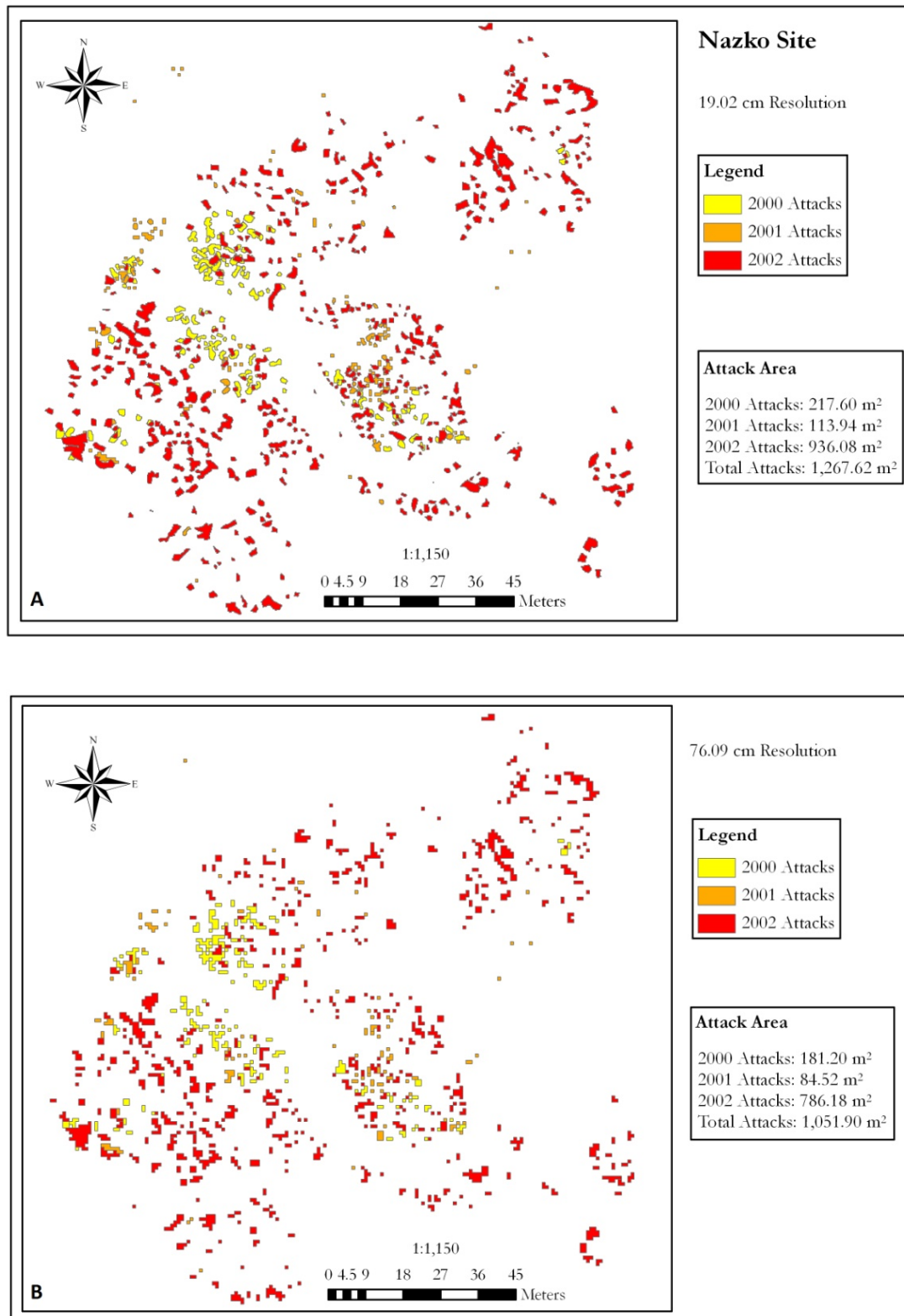


Figure 3a, b. Maps of each of the three attack years at 19 cm and 76 cm spatial resolutions

A chi-square two-tailed test statistic was used for hypothesis testing. These approaches analyze only spatial coordinates, but they give important indications about whether a Poisson distribution is in operation to produce random patterns (Fotheringham et al., 2002; Lloyd, 2007). These indices serve as the initial starting point in the

pattern identification process.

#### 2.4 Spatial Data Analysis

The Ripley's  $K$ -function is used to define global and local summaries of the spatial dependence within the data measured at different scales (Lloyd, 2007). Ripley's  $K$ -function is designed to identify the amount of spatial clustering within a multiple distance radii (K. P. Vadrevu et al., 2008). In this study, the amount of clustering or dispersion was measured for 10 distance bands with about a 3 m spread between bands (Lloyd, 2007). The residuals from the differences between the expected and observed values are compared to determine spatial dependence at each band interval (Hering et al., 2009; Spooner et al., 2004). The method offers an alternative to identify pattern distributions at multiple scales in comparison to the standardized value derived for the entire study area in the case of the mean nearest neighbour value (Lloyd, 2007; Vadrevu & Badarinath, 2009). Ripley's  $K$ -function is calculated as (Lloyd, 2007):

$$K(t) = n^{-2}A \sum_i^n \sum_{j \neq i} w_{ij} I_t(u_{ij}) \quad (4)$$

where  $t$  is the radius of a circle centered on a point in the pattern (band interval distance),  $n$  is the total number of points,  $A$  is the plot area,  $w_{ij}$  is a weighting factor to correct for edge effects, and  $I_t$  is a counter variable set to 1 if the distance  $u_{ij}$  between points  $i$  and  $j$  is  $\leq t$ , otherwise  $I_t = 0$ .

Examining dependencies between locations and feature attributes can also provide valuable information about the spatial pattern. Spatial autocorrelation describes the correlation of the values of a single attribute related to the proximity of those values in geographic space (Suzuki, Kachi, & Suzuki, 2008; Tobin, 2004). Spatial autocorrelation relies on data interdependence and spatial statistics (Odland, 1988; Overmars et al., 2003). In this sense, the inverse of distance measurements assumes the attribute values follow a relationship where closer objects are assigned larger weights and farther objects are assigned smaller weights (Griffith, 2003; Odland, 1988). Positive spatial autocorrelation occurs when adjacent neighbourhood locations have similar attribute values (Almeida-Neto & Lewinsohn, 2004; Griffith, 2003; Mueller-Warrant et al., 2008). Spatial autocorrelation also uses the Pearson product-moment correlation coefficient and linear regression model to measure the strength and directional movement of the relationship (Almeida-Neto & Lewinsohn, 2004; Griffith, 2003). For this study, the Moran's  $I$  index of spatial autocorrelation was used at each of the three consecutive years to produce a global measurement of similarities or dissimilarities between neighbouring polygons (Fotheringham et al., 2002; Suzuki et al., 2008). The global Moran's  $I$  index is calculated as (Lloyd, 2007):

$$I = \frac{n \sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i \sum_j w_{ij} \sum_i (x_i - \bar{x})^2} \quad (5)$$

where  $n$  is the number of features measured,  $x_i$  is the value of feature  $i$ ,  $x_j$  is the value of feature  $j$ ,  $\bar{x}$  is the mean value of all features, and  $w_{ij}$  is the weight assigned to each pair of features  $x_i$ ,  $x_j$ .

Spatial autocorrelation can also be measured at the local level to evaluate the extent of autocorrelation within local neighbourhoods. Local measures capture the many local spatial variation and spatial dependency while global measurements provide only one set of values that represent the extent of spatial autocorrelation across the entire study area (Mueller-Warrant et al., 2008). The local Moran's  $I$  analysis identifies spatial clusters including low values in a low value neighbourhood (low-low clustering) and high values in a high value neighbourhood (high-high clustering) (Getis & Ord, 1992). The local Moran's  $I$  is calculated as (Griffith, 2003):

$$I_i = \left( \frac{z_i - \bar{z}}{\sigma^2} \right) \sum_{j=1, j \neq i}^n [w_{ij} (z_j - \bar{z})] \quad (6)$$

where  $z_i$  is the value of the variable  $z$  at location  $i$ ,  $\bar{z}$  is the value of the variable  $z$  at all other locations (where  $j \neq i$ ),  $\sigma^2$  is the variance of variable  $z$ , and  $w_{ij}$  is a weight defined as the inverse of the distance  $d_{ij}$  among locations  $i$  and  $j$  (Zhang et al., 2008). The local Moran's  $I$  analysis is often most valuable as a scatter plot displaying the relationship between the Moran's  $I$   $z$ -score and associated Moran's  $I$  index values (Griffith, 2003; Lloyd, 2007). In this graphical format the outliers and clusters can be easily identified in addition to the nature and strength of the correlation relationship (Lloyd, 2007).



In addition, methods developed by Getis and Ord (1992; 1996) not only provide hypothesis testing to determine whether clustering has occurred within a dataset, but also provide information on the extent to which above and below average values cluster more strongly and identify local concentrations of clustering (Laffan, 2006; Mueller-Warrant et al., 2008). The Getis-Ord General G high and low clustering analysis was done on each of the three years of data to estimate at a global level the extent to which high and low values cluster (Mueller-Warrant et al., 2008). The Getis-Ord General G high/low clustering is calculated as:

$$G = \frac{\sum_i \sum_j w_{ij}(x_i x_j)}{\sum_i \sum_j (x_i x_j)} \quad (7)$$

where  $x_i$  represents the value of feature  $i$ ,  $x_j$  represents the value of feature  $j$ , and  $w_{ij}$  is the weight assigned to each pair of features  $x_i, x_j$  (Getis & Ord, 1992; Mueller-Warrant et al., 2008). Also, the expected value for Getis-Ord General G for values distributed randomly is:

$$\frac{\sum_i \sum_j w_{ij}}{n(n-1)} \quad (8)$$

The Getis-Ord  $G_i^*$  test statistic is a local adaptation of global Getis-Ord General G and seeks to identify areas of hot and cold clustering based on local neighbourhood values (Getis & Ord, 1996; Laffan, 2006). The  $G_i^*$  statistic is calculated as the summation of the differences between local sample values and the mean, and is observed as standard normal distribution  $z$ -score values:

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j - W_i^* \bar{x}^*}{s^* \sqrt{\frac{(nS_{1j}^*) - W_i^{*2}}{n-1}}} \quad (9)$$

where  $i$  is the centre of the local neighbourhood,  $d$  is the lag distance (bandwidth of the sample window),  $w_{ij}$  is the weight for neighbour  $j$  from location  $i$ ,  $n$  is the number of samples in the dataset,  $W_i^*$  represents the summation of the weights,  $S_{1j}^*$  is the number of samples within  $d$  of the central location (=  $W_i^*$  for a binary weights case),  $\bar{x}^*$  is the mean of the entire dataset, and  $s^*$  is the standard deviation of the entire dataset (Getis & Ord, 1996; Laffan, 2006). For this analysis, the results of the Getis-Ord  $G_i^*$  testing may be best visualized in a cartographic output format to easily identify local variation within the data (Laffan, 2006).

### 3. Results and Discussion

#### 3.1 Exploratory Spatial Data Analysis

The average nearest neighbour indices derived for both the 19.02 cm and 76.09 cm spatial resolution data showed a significant clustered pattern distribution indicating a mainly local spread of attack in each of the analyzed years, with the 19.02 cm data for the 2000 attack displaying the strongest clustering with a  $z$ -score value of -8.38 (Table 1). A 99% confidence interval was maintained for each of the datasets  $z$ -score. There is limited difference between the two spatial resolutions with similar measurements of absolute nearest neighbour ratios and accompanied  $z$ -score values obtained across the three years examined. However, the exception to this is observed in the 76.09 cm data for the 2001 attack site, where a significant change in both the absolute nearest neighbour ratio and  $z$ -score values distinguishes this dataset from the others. Nevertheless, based on the results of this analysis, the null hypothesis ( $H_0$ ) that no difference exists between observed and random nearest neighbour values can be rejected as the datasets are statistically significant. Furthermore, it is concluded that the pattern of mountain pine beetle attacks in the study area shows a significant clustering pattern with a 1% or less likelihood that the pattern was a result of random chance.

Table 1. Detailed results from the average nearest neighbor analysis

Attack Year	Resolution (cm)	Observed Mean Distance (m)	Expected Mean Distance (m)	Absolute Index Value	Z-Score	Pattern
2000	19.02	2.50	4.25	0.59	-8.38	Clustered
2001	19.02	3.28	5.13	0.64	-7.27	Clustered
2002	19.02	2.74	3.19	0.86	-5.62	Clustered
2000	76.09	2.86	4.63	0.62	-7.22	Clustered
2001	76.09	4.13	4.92	0.84	-2.71	Clustered
2002	76.09	2.97	3.44	0.86	-5.12	Clustered

The quadrat analysis confirmed the results of the average nearest neighbour measurements displaying clustering between the examined point distributions at both scales of investigation (Table 2). The low  $p$ -values derived indicate statistically significant clustering of points based on the 10 m x 10 m grid considered (McGrew & Monroe, 2000). A substantial difference between the sets of data can be visualized through the mean cell frequency ( $MEAN$ ) measure, as well as the variance of cell frequencies ( $VAR$ ) as the attack sites in 2002 exhibit considerable difference in comparison to the attacks in 2000 and 2001. As a result of the fact that this measurement is designed to analyze point distributions and the point distribution was determined using the centroid of the polygons, not all attacked trees were represented in this analysis because of the varying sizes of the polygons with some containing more than one killed tree (Lloyd, 2007). This measurement was, therefore, useful in examining the spatial dependency of the absolute point location of the attack sites. However, it must be assessed in combination with the other measurements explored in this study to retain the spatial information of attack patterns and the associated quantities of trees killed (Griffith, 2003; Lloyd, 2007).

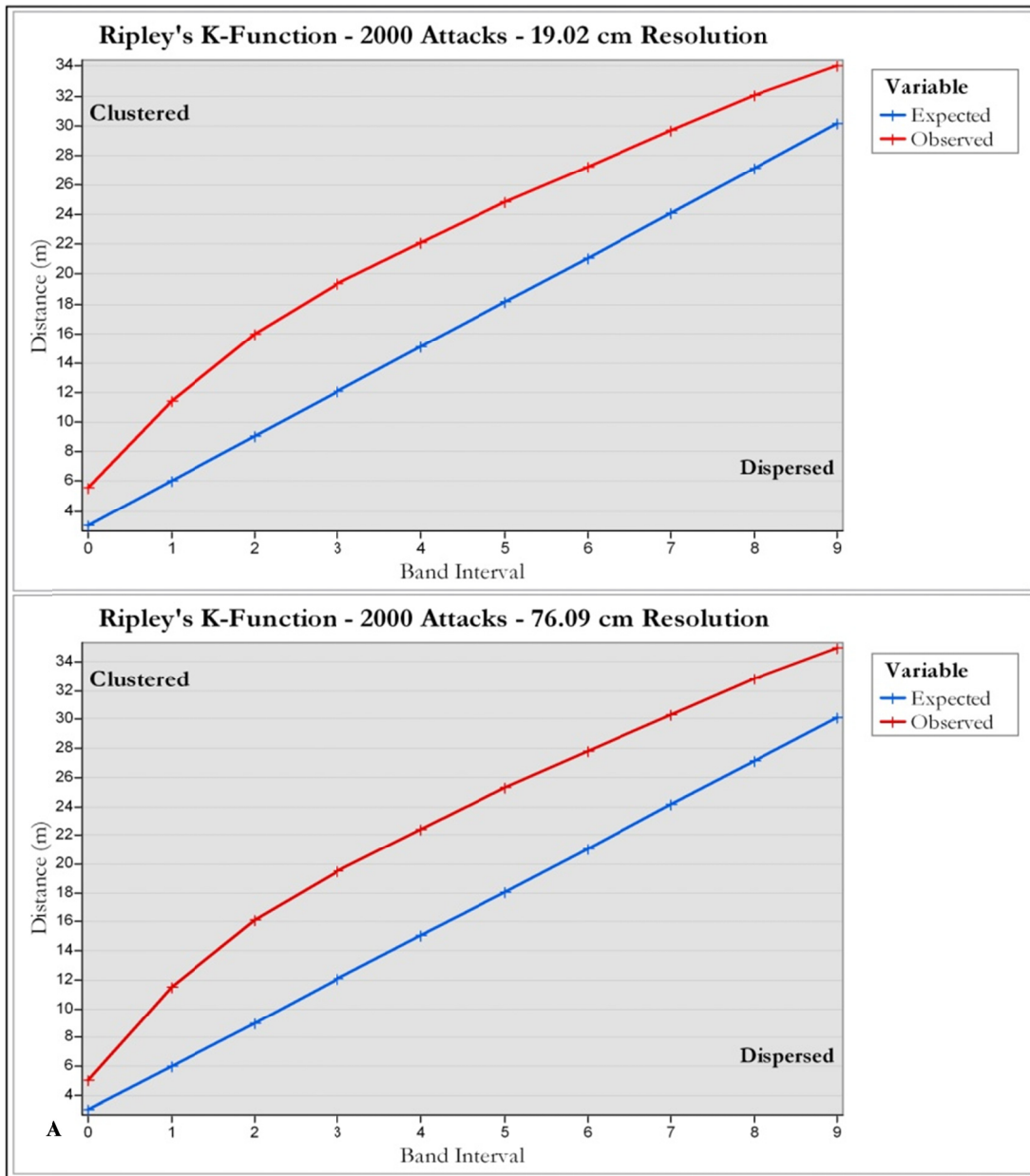
Table 2. Detailed results from the quadrat analysis with a 10 m x 10 m quadrat cell size

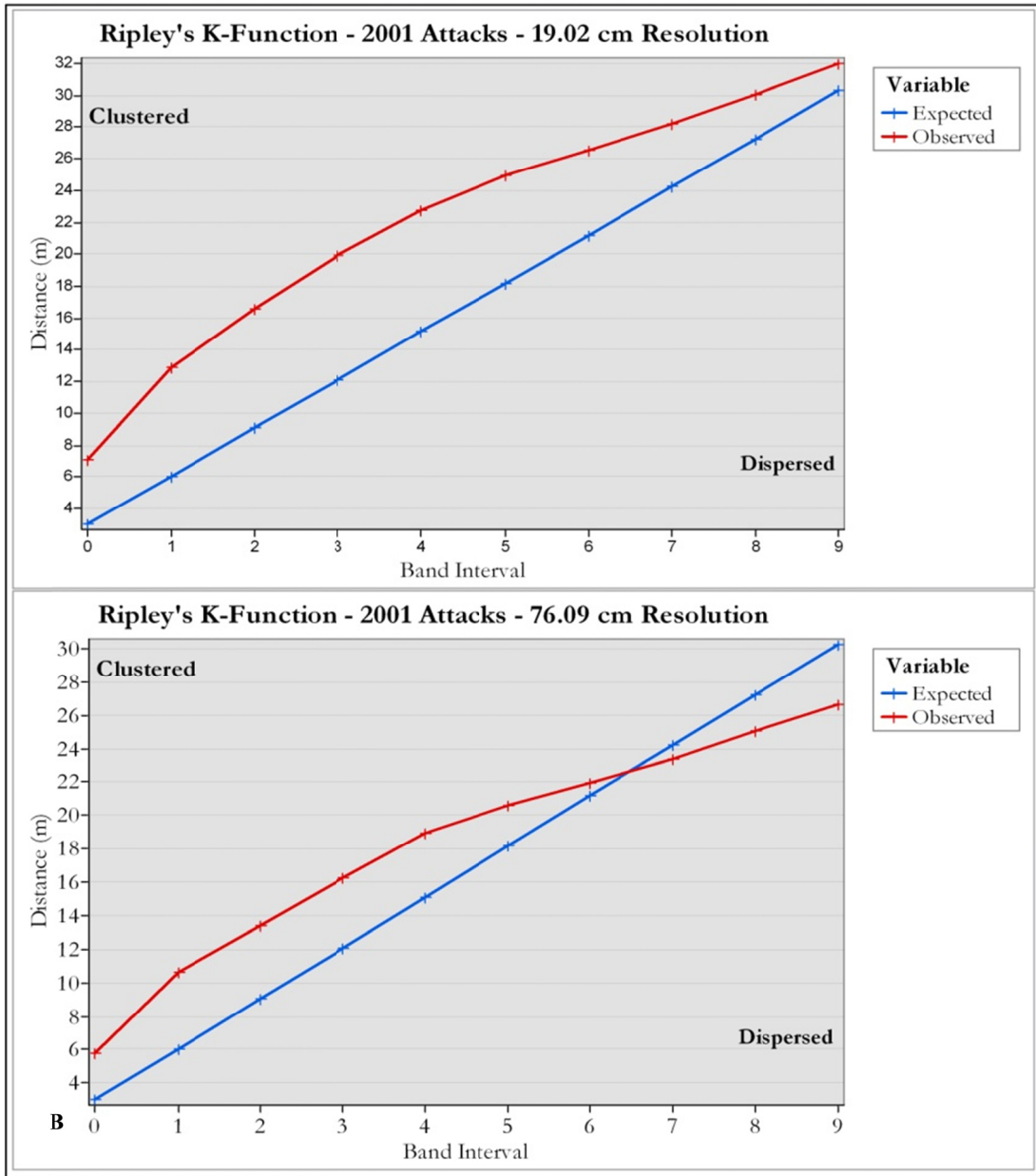
Attack Year	Resolution (cm)	Mean Cell Frequency ( $MEAN$ )	Variance of Cell Frequencies ( $VAR$ )	Variance-Mean Ratio ( $VMR$ )	$p$ -value	Pattern
2000	19.02	0.58	3.07	5.32	$4.46 \times 10^{-115}$	Clustered
2001	19.02	0.56	2.61	4.64	$5.66 \times 10^{-92}$	Clustered
2002	19.02	2.23	8.02	3.59	$5.33 \times 10^{-58}$	Clustered
2000	76.09	0.49	2.04	4.16	$3.16 \times 10^{-76}$	Clustered
2001	76.09	0.39	1.36	3.51	$1.02 \times 10^{-55}$	Clustered
2002	76.09	1.92	6.60	3.43	$2.50 \times 10^{-53}$	Clustered

### 3.2 Ripley's $K$ -function

Similar to the results derived from the average nearest neighbour analysis, a clustered pattern dominated the majority of the measurements derived from the Ripley's  $K$ -function index taken at each of the 10 band intervals (Vadrevu & Badarinath, 2009). The  $K$ -function measure for the 2000 attacked trees indicated a clustered pattern stood out at each of the ten band distances measured with little change between the 19.02 cm and 76.09 cm spatial resolutions (Figure 4). The observed value was closest to the expected distance value in the first band interval, but then remained as a clustered pattern arrangement for each of the following band intervals with little change in distance from the expected value. The 2001  $K$ -function measurement provided the only significant change in patterns across the band intervals. The 2001 attack at a 19.02 cm resolution data showed a significant movement from a strong clustered pattern towards the expected value near the final band interval measurements and the 76.09 cm resolution identifies a similar pattern, but moved from a clustered to a slightly dispersed pattern during these final band intervals. This was the only point in this study throughout each of the analyses

and at both spatial resolution scales where a region of measureable uniform dispersion was recognized. The results for the 2002 attacked trees yielded the observations closest to the expected value with a similar straight line, slightly clustered results for both the 19.02 cm and 76.09 cm resolution sites.





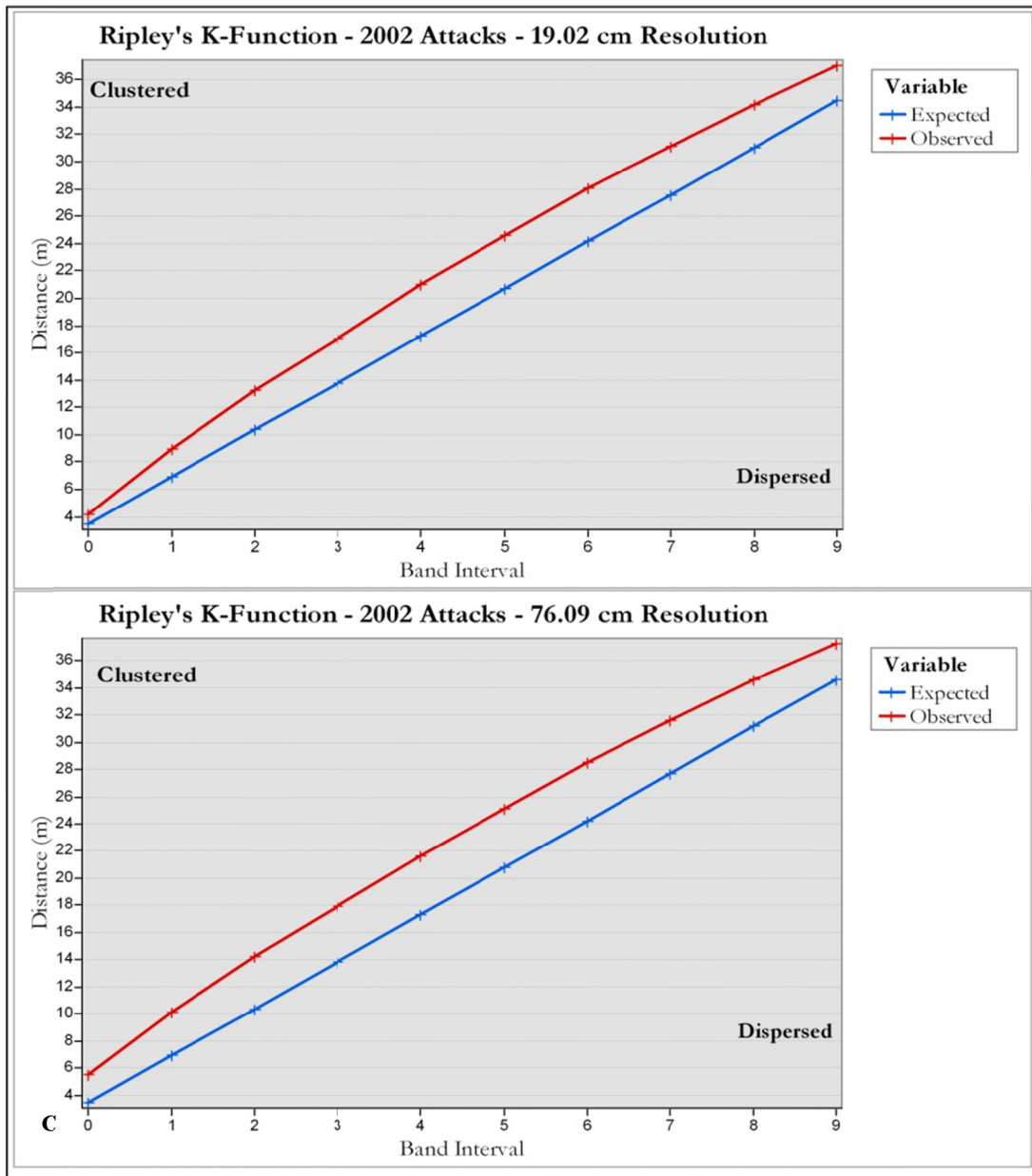


Figure 4a, b, c. Graphs showing the difference between the expected value and observed values in accordance with Ripley's K-function analysis

Despite the relatively unique results for the 2001 attack at the 76.09 cm resolution, the overall pattern displayed slightly clustered values and a linear path for each of the data sets examined throughout each of the band intervals. The Ripley's K-function value indicated there was no distinct change in structural pattern within the 2000 and 2002 attacked trees as is apparent within the 2001 data set and there was no change in pattern at the two scales analyzed.

### 3.3 Global and Local Moran's I Spatial Autocorrelation

The global Moran's I statistic was used to measure the nature of spatial autocorrelation between the mountain pine beetle attack periods at both the 19.02 cm and 76.09 cm spatial resolutions. Because this index required an associated attribute variable, the total amount of killed trees in each polygon in each of the examined attack years was assessed. The results of both spatial resolutions gave similar results with the year 2000 and 2001 displaying very similar spatial autocorrelation values while the dataset for the trees attacked in 2002 yielded different results (Table 3). Both the 2000 and 2001 attack sites at both scales of observation contained no spatial autocorrelation indicating no systematic structure of pattern formation with all Moran's I index values in close proximity to zero

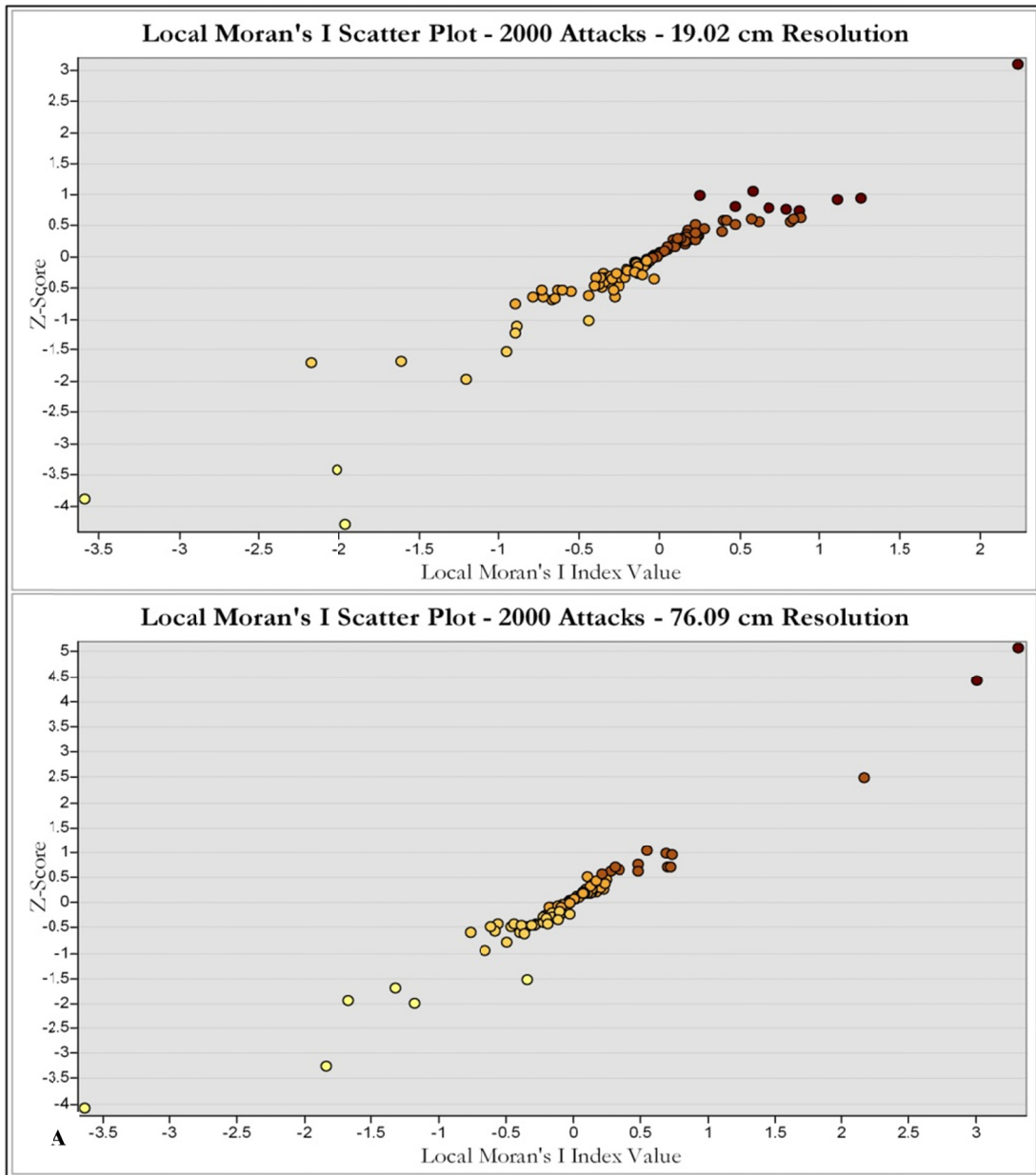
on the negative side (Griffith, 2003). Conversely, the trees killed in 2002 yielded a clustered pattern despite a relatively low positive Moran's  $I$  Index score. This indicates a positive spatial autocorrelation within the dataset based on the relationship between the geographic location and spatial scattering of the attack patches and the total number of killed trees within each. A  $z$ -score of 2.80 was derived for this site at the 19.02 cm scale falling outside of the 99% confidence interval level. The  $z$ -score value of -2.27 for the 76.09 cm spatial resolution produces a clustered pattern with a slightly lower 95% confidence interval indicating strong statistical significance, but a higher likelihood that the pattern could be a result of random chance.

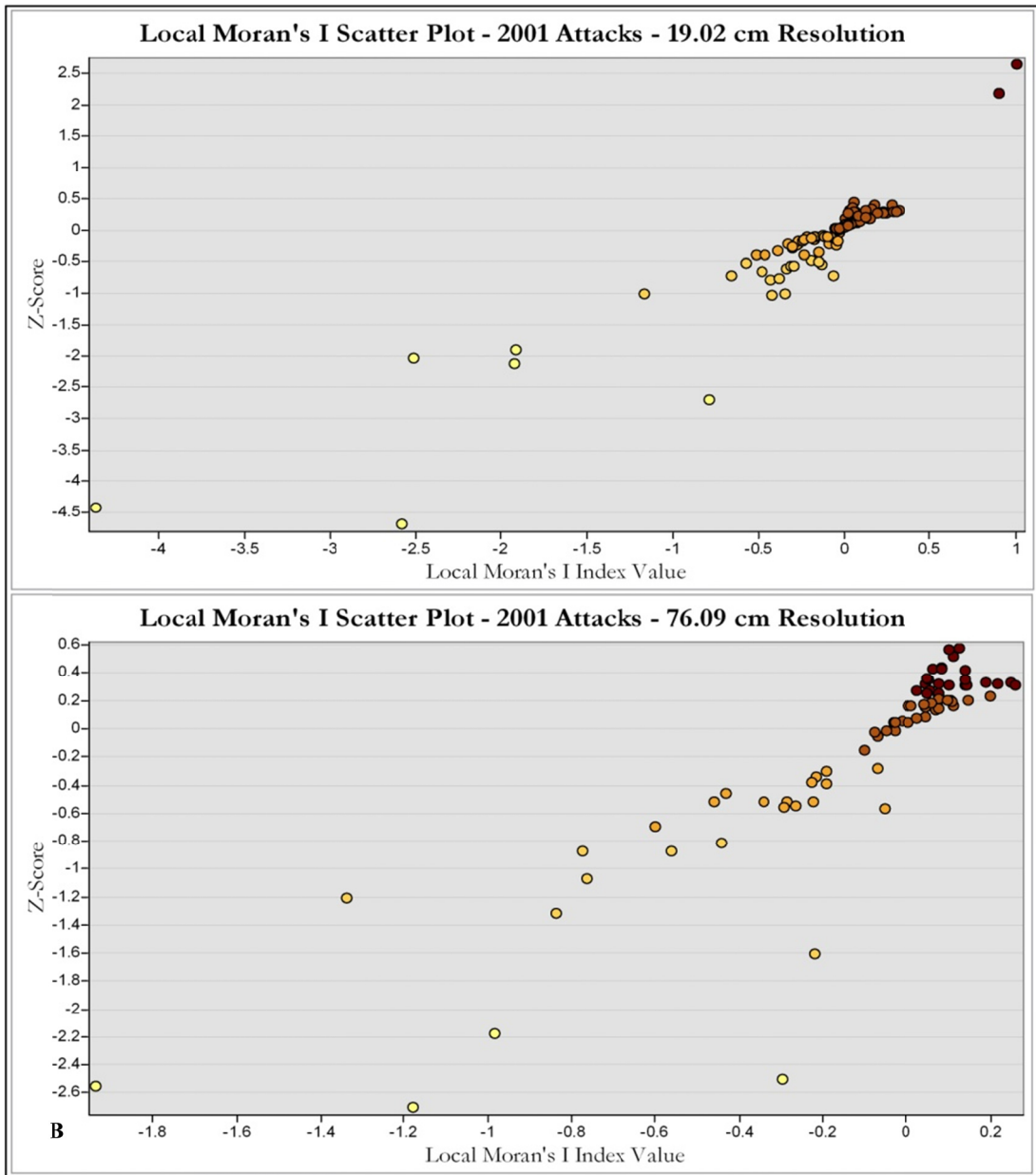
Table 3. Results obtained from the global Moran's  $I$  spatial autocorrelation measure

Attack Year	Resolution (cm)	Moran's $I$ Index	Z-Score	Confidence Interval	Pattern	Spatial Autocorrelation
2000	19.02	-0.03	-0.70	99%	Random	Not Apparent
2001	19.02	-0.06	-1.30	99%	Random	Not Apparent
2002	19.02	0.06	2.80	99%	Clustered	Positive
2000	76.09	-0.01	0.09	99%	Random	Not Apparent
2001	76.09	-0.05	-0.94	99%	Random	Not Apparent
2002	76.09	0.05	2.27	95%	Clustered	Positive

The local form of the Moran's  $I$  spatial autocorrelation allows for the assessment of local variation of spatial autocorrelation and is useful in detecting spatial clustering as well as outliers (Lloyd, 2007). The results derived from the local Moran's  $I$  index are visualized using a scatter plot to graph the generated  $z$ -score and Moran's  $I$  index values in order to easily identify influential values within the dataset, as well as the direction of local autocorrelation (Figure 5) (Griffith, 2003). Despite the fact that several significant  $z$ -score values representing values several standard deviations away from the mean exist in each of the three years examined and at both spatial resolutions, there are no observable outliers present within the data. Based on the local Moran's  $I$  results, although there are some values close to the threshold limits, there are no areas of local neighbourhood negative autocorrelation indicating a lack of uniform dispersion.







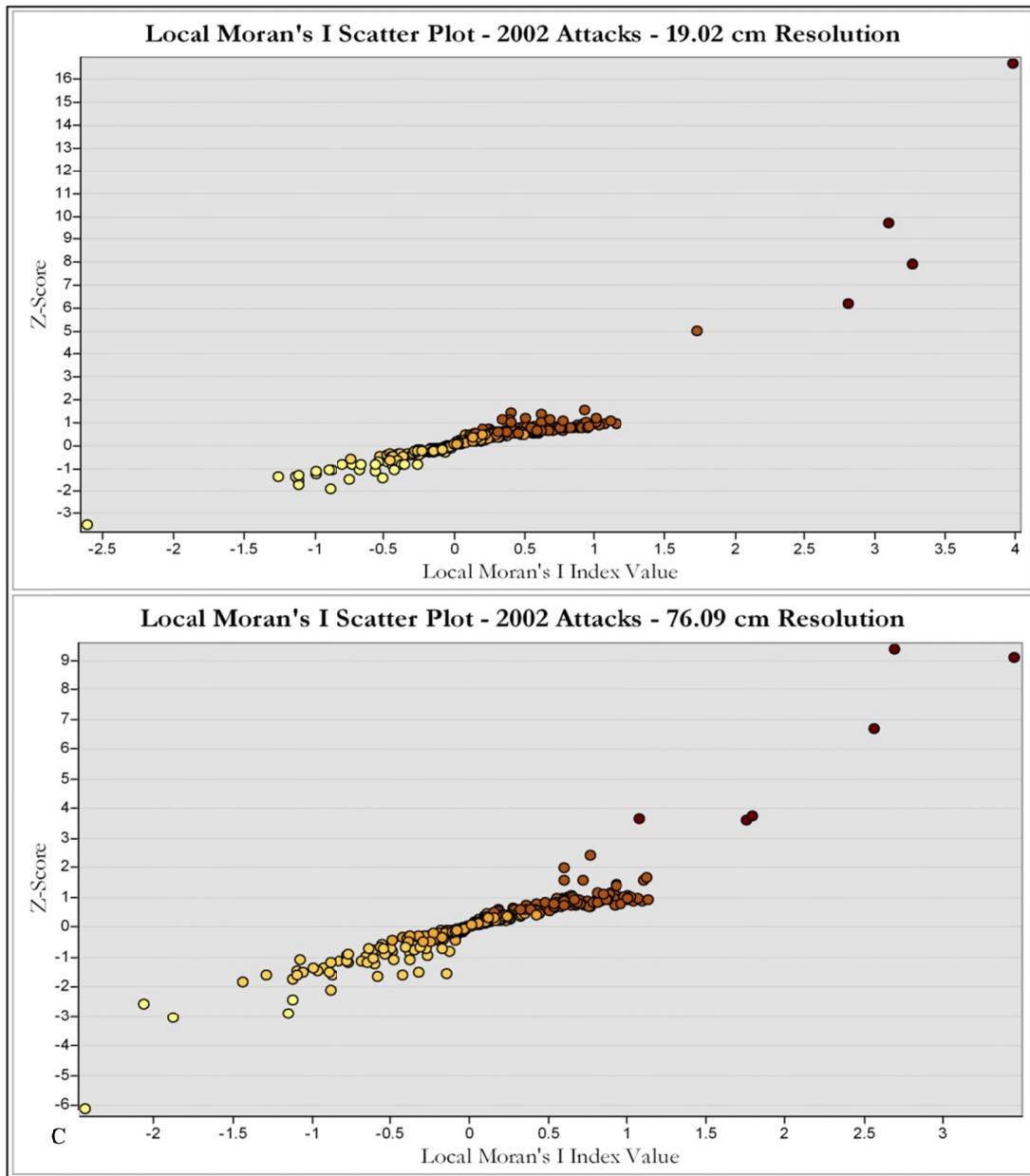


Figure 5a, b, c. Graphs showing results from the local Moran’s I spatial autocorrelation measure

### 3.4 Getis-Ord General G and Hot Spot Analysis

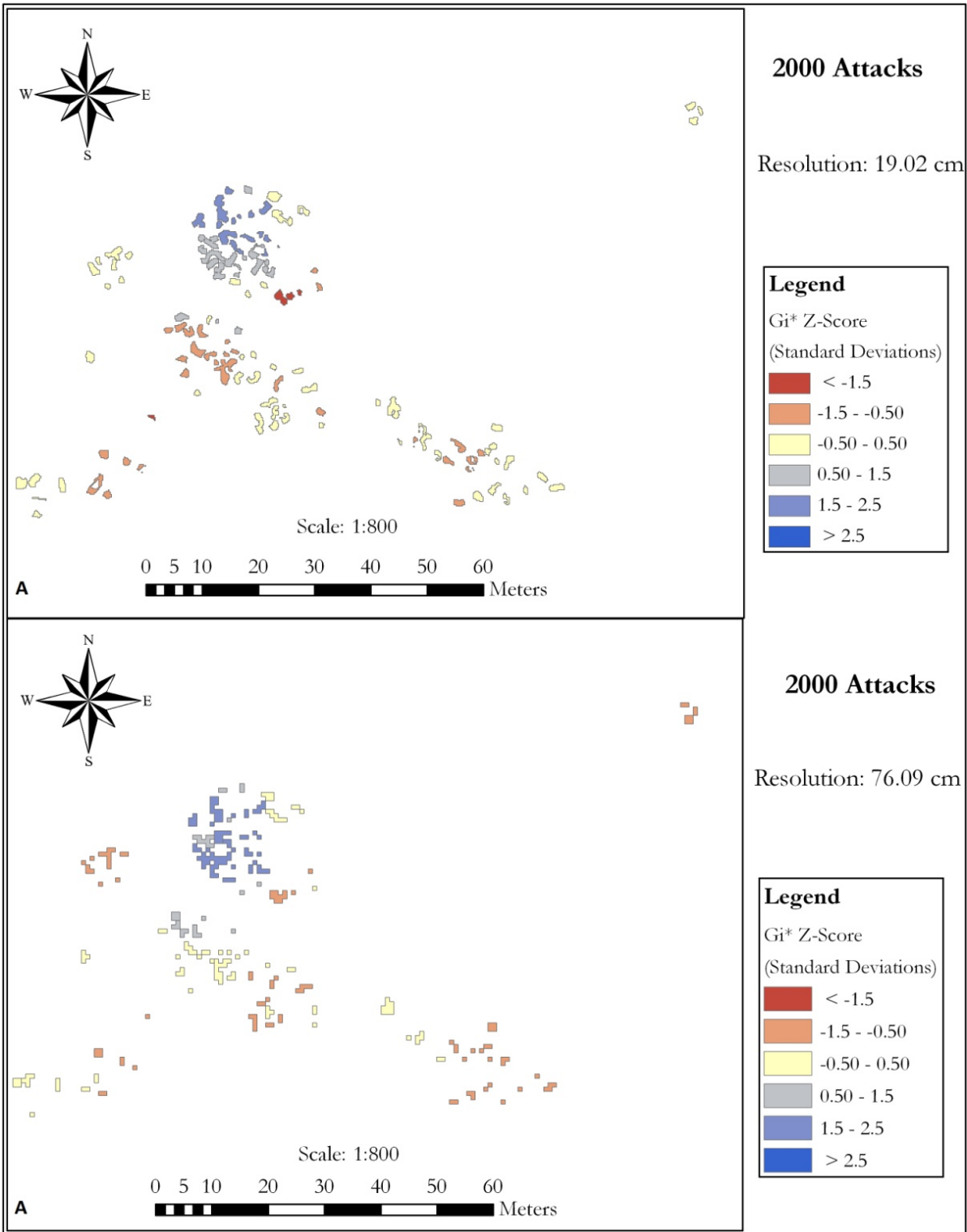
The Getis-Ord General G measurement aided in understanding the similarities between the 2000 and 2001 attack years and how these two datasets differ significantly from the 2002 attack year. As a global index measurement for the entire dataset, random patterns indicating no apparent clustering were determined for both the 2000 and 2001 attack sites, while the data collected on the 2002 attack sites indicated a significant level of low value clustering (Table 4). Limited variation exists between the two levels of spatial resolution. However, noticeable differences within the 2000 attack between the 19.02 cm resolution and the 76.09 cm resolution can be identified in the z-score value and represents the largest differentiation in z-score value across both scales. A General G index value of 0.03 was still maintained for both resolutions and a random pattern was assumed as the null hypothesis ( $H_0$ ) could not be rejected based on the significance test. The 2000 attack at the 19.02 cm resolution indicated the highest amount of low value clustering between the 2000 and 2001 datasets.

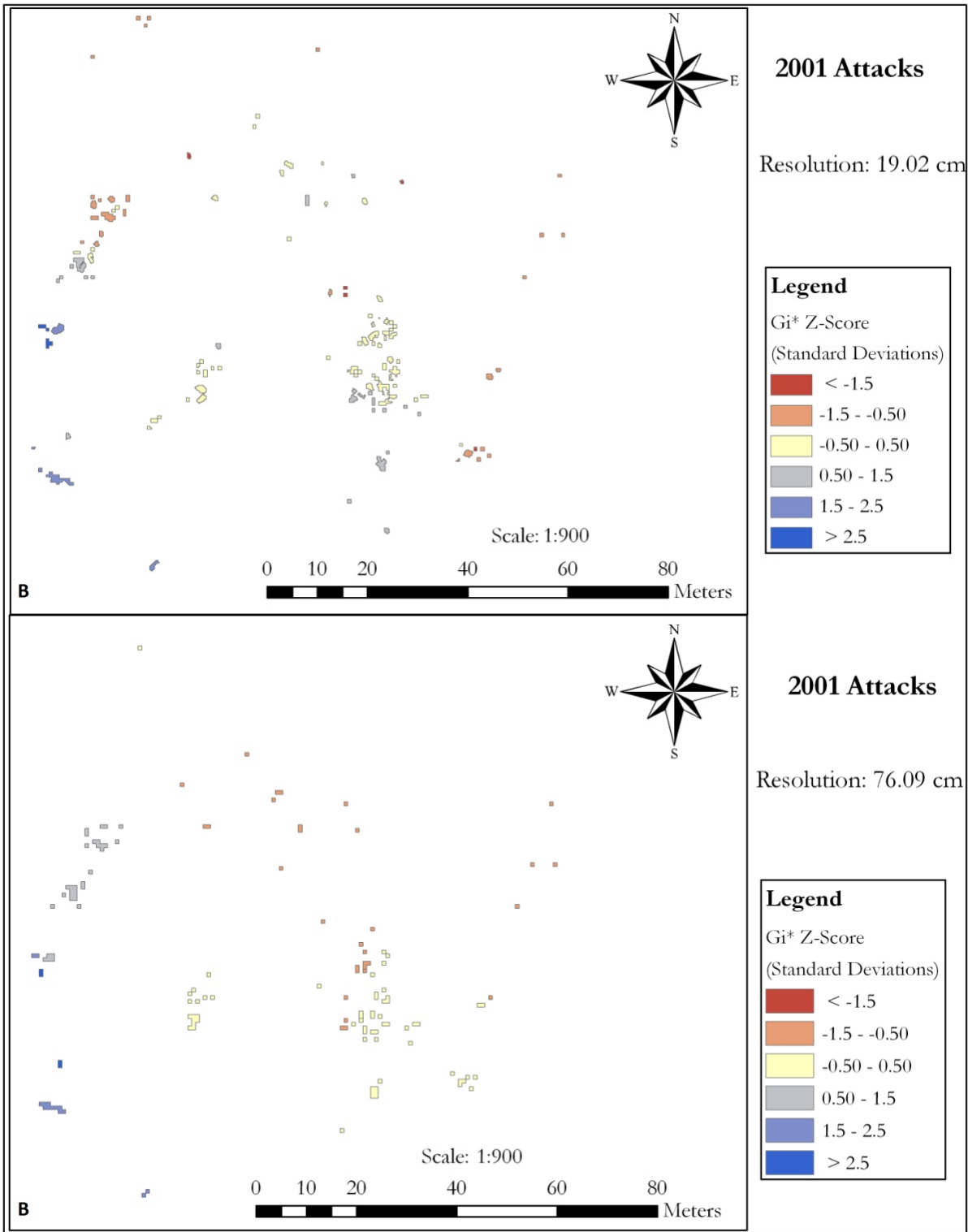
The Getis-Ord cluster analysis indices analyze data at both the global and local levels in order to determine overall as well as neighbourhood variation in spatial dependence (Fotheringham et al., 2002; Mueller-Warrant et

al., 2008). Figure 6 illustrates a choropleth map of the Getis-Ord  $G_i^*$  test statistic  $z$ -score values for each of the 2000, 2001 and 2002 years at both scales of analysis and utilizes a standard deviation classification structure. Based on the results from the General G analysis, there were a much larger proportion of the areas displaying negative  $z$ -scores indicating higher levels of low value clustering in comparison to high value clustering even though positive  $z$ -scores displayed a larger range with several locations with values more than 2.5 standard deviations from the mean. Using this local measurement and displayed spatially, it was easily recognized that significant areas of variation exist between the two scales especially within the 2002 dataset (Fotheringham et al., 2002). Within this dataset, variation can be observed, especially within positive  $z$ -scores between the two spatial resolutions in the northeast corner of the dataset as the data from the 19.02 cm resolution has a larger number of  $z$ -score values close to the mean compared to the larger range in values at the 76.09 cm resolution.

Table 4. Results obtained from the Getis-Ord General G spatial clustering measure

Attack Year	Resolution (cm)	General G Index	Z-Score	Confidence Interval	High/Low Values Cluster
2000	19.02	0.03	-0.82	99%	Random
2001	19.02	0.03	-0.17	99%	Random
2002	19.02	0.00	-3.73	99%	Lows Cluster
2000	76.09	0.03	0.05	99%	Random
2001	76.09	0.03	-0.35	99%	Random
2002	76.09	0.01	-3.17	99%	Lows Cluster







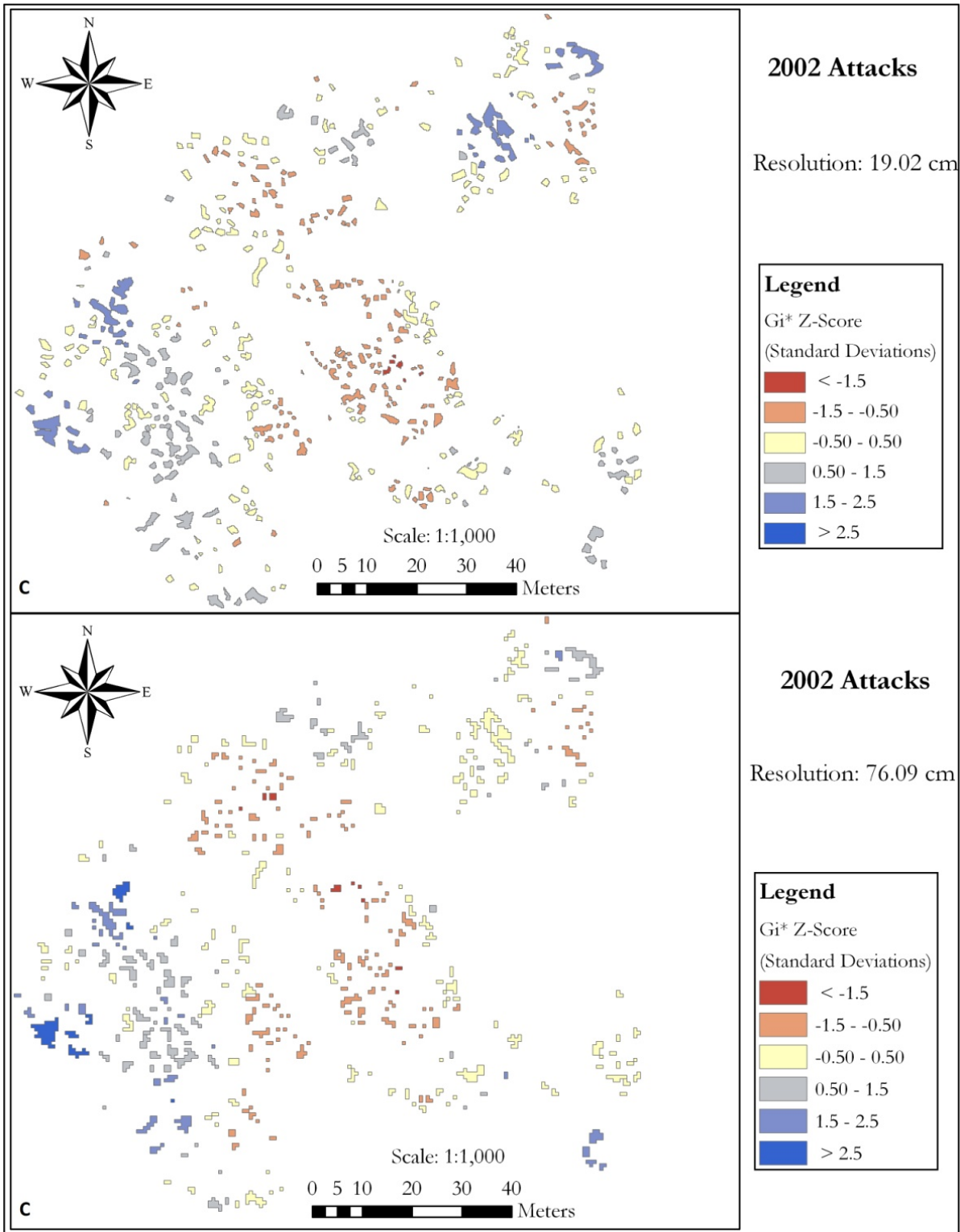


Figure 6a, b, c. Maps displaying the Getis-Ord  $G_i^*$  z-score values classified into standard deviation intervals used as a cluster analysis method to detect hot and cold spots

The results show a significant clustering pattern throughout the indices. This indicated a preference for some locations over others (Griffith, 2003; Tobin, 2004). With the exception of the one region within the 2001 dataset at the 76.09 cm spatial resolution, no dispersed patterns demonstrating a uniform distribution of attack locations were detected. This is logical in that both the mountain pine beetle and the lodgepole pine are biological species

that are not operating according to any systematic or structured pattern system and, thus, a dispersed pattern would not be expected from these objects of investigation (Safranyik & Wilson, 2006). Distinct differences were recognized between the wholly clustered patterns detected in the first two measurements utilized, with the mentioned exception of the 2001 attack at 76.09 cm resolution, and the remaining four measures that employed spatial location with the quantity of killed trees. These indices help uncover the fact that the 2000 and 2001 attack sites contain various levels of similarity in spatial dependency, while the mountain pine beetle attack changed dramatically in 2002 with significant increases in total attack area and killed trees influencing the detected patterns. This was confirmed in the spatial and exploratory data analysis procedures implemented, and even when a distinct pattern was shared by all three years, noticeable dissimilarities were observed between the 2000 and 2001 datasets and the 2002 dataset.

In conclusion, the results obtained from this study provided descriptive information about the spatial distribution of the 2000, 2001 and 2002 attack site years across two varying levels of spatial resolution in an attempt to advance existing knowledge on the geographic and ecological behaviour of mountain pine beetle. A significant change in spatial orientation of the 2002 attack site in comparison to the previous two years not only was demonstrated in the total attack area and, by extension, the amount of lodgepole trees killed, but also in the global and local level spatial autocorrelation as the extent of value clustering. A clustered spatial pattern with positive spatial autocorrelation dominated the organization of the mountain pine beetle at this study location.

This study demonstrated that aggregation of high resolution imagery produced slightly differing results when compared with the original spatial resolution. However, the main spatial organization and patterns of the individual datasets appeared to be preserved. The idea that the scale at which the analysis of a spatial phenomenon occurs has an influence on the output results are well documented in geography. However, multi-scale inspections must continue to be performed in order to quantitatively describe and predict the extent to which derived results change (Haining, 2003; Odland, 1988; Overmars et al., 2003). Nevertheless, the usefulness of performing spatial analysis at multiple scales was evident by the application and interpretation of various measurement indices, especially with regards to the Ripley's *K*-Function analysis.

Further research must be performed to analyze the spatial distribution and patterns of other sites affected by mountain pine beetle epidemics as well as to gather increased temporal resolution data in order to examine the spread of infestations over longer periods of time. Research is already underway by Coggins et al. (2008) and others to establish temporal models of attack sites over longer periods of time to improve predictions of future spreading as well as to derive accurate estimates of the amount of future hectares lost to attacks. Coggins et al. (2008) and Wulder et al. (2006) have demonstrated that remotely-sensed data has significant potential for mapping mountain pine beetle impacts and must continue to be used in combination with spatial analysis and other geographic information science techniques to succinctly describe and investigate the movement of the species. The importance of continued monitoring of mountain pine beetle infestations is vital for forest management to develop and improve mitigation efforts for affected areas and prevent stand depletion in non-affected regions.

### Acknowledgements

The aerial photographs were acquired under a research grant awarded to Arthur Roberts and Suzana Dragicevic. Christopher Bone and Liliana Perez assisted with some initial image preparation and processing. The Simon Fraser University Library Serials Division contributed the journal publication fee. The valuable feedback of the anonymous peer-reviewers improved the overall quality of the manuscript. All authors participated in the design and implementation of the research ideas, manuscript preparation, and final revisions.

### References

- Almeida-Neto, M., & Lewinsohn, T. M. (2004). Small-scale spatial autocorrelation and the interpretation of relationships between phenological parameters. *Journal of Vegetation Science*, 15(4), 561-568. <http://dx.doi.org/10.1111/j.1654-1103.2004.tb02295.x>
- Bone, C., Dragicevic, S., & 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. <http://dx.doi.org/10.1080/01431160500239180>
- Bone, C., Dragicevic, S., & Roberts, A. (2006). A fuzzy-constrained cellular automata model of forest insect infestations. *Ecological Modelling*, 192(1-2), 107-125. <http://dx.doi.org/10.1016/j.ecolmodel.2005.09.013>
- Boots, B. N., & Getis, A. (1988). *Point pattern analysis*. Newbury Park, California: Sage Publications.
- Coggins, S., Coops, N. C., & Wulder, M. A. (2008). Initialization of an insect infestation spread model using tree

- structure and spatial characteristics derived from high spatial resolution digital aerial imagery. *Canadian Journal of Remote Sensing*, 34(6), 485-502. <http://dx.doi.org/10.5589/m08-073>
- Erfanifard, Y., Feghhi, J., Zobeiri, M., & Namiranian, M. (2009). Spatial pattern analysis in Persian oak (*Quercus brantii* var. *persica*) forests on B&W aerial photographs. *Environmental Monitoring and Assessment*, 150(1-4), 251-259. <http://dx.doi.org/10.1007/s10661-008-0227-4>
- Fotheringham, S. A., Brunson, C., & Charlton, M. (2002). *Geographically weighted regression: the analysis of spatially varying relationships*. Chichester, West Sussex, England: John Wiley & Sons Ltd.
- Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24(3), 189-206. <http://dx.doi.org/10.1111/j.1538-4632.1992.tb00261.x>
- Getis, A., & Ord, J. K. (1996). Local spatial statistics: an overview. In P. Longley & M. Batty (Eds.), *Spatial Analysis: Modelling in a GIS Environment* (pp. 261-277). New York: Wiley.
- Griffith, D. A. (2003). *Spatial autocorrelation and spatial filtering: gaining understanding through theory and scientific visualization*. Berlin, Germany: Springer-Verlag.
- Haining, R. (2003). *Spatial data analysis: theory and practice*. Cambridge, England: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511754944>
- Hering, A. S., Bell, C. L., & Genton, M. G. (2009). Modeling spatio-temporal wildfire ignition point patterns. *Environmental and Ecological Statistics*, 16(2), 225-250. <http://dx.doi.org/10.1007/s10651-007-0080-6>
- Laffan, S. W. (2006). Assessing regional scale weed distributions, with an Australian example using *Nassella trichotoma*. *Weed Research*, 46(3), 194-206. <http://dx.doi.org/10.1111/j.1365-3180.2006.00491.x>
- Lloyd, C. D. (2007). *Local models for spatial analysis*. Boca Raton, FL: CRC Press Taylor & Francis Group.
- Lynch, H. J., & Moorcroft, P. R. (2008). A spatiotemporal Ripley's K-function to analyze interactions between spruce budworm and fire in British Columbia, Canada. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, 38(12), 3112-3119. <http://dx.doi.org/10.1139/X08-143>
- McGrew, J. C., & Monroe, C. B. (2000). *An introduction to statistical problem solving in geography*. Boston, MA: McGraw-Hill.
- Mueller-Warrant, G. W., Whittaker, G. W., & Young, W. C. (2008). GIS analysis of spatial clustering and temporal change in weeds of grass seed crops. *Weed Science*, 56(5), 647-669. <http://dx.doi.org/10.1614/WS-07-032.1>
- Nelson, T. A., Boots, B., Wulder, M., & Carroll, A. (2007). Environmental characteristics of mountain pine beetle infestation hot spots. *BC Journal of Ecosystems and Management*, 8(1), 91-108.
- Odland, J. (1988). *Spatial Autocorrelation*. Newbury Park, CA: Sage Publications.
- Overmars, K. P., de Koning, G. H. J., & Veldkamp, A. (2003). Spatial autocorrelation in multi-scale land use models. *Ecological Modelling*, 164(2-3), 257-270. [http://dx.doi.org/10.1016/S0304-3800\(03\)00070-X](http://dx.doi.org/10.1016/S0304-3800(03)00070-X)
- Safranyik, L., & Wilson, B. (2006). *The mountain pine beetle: a synthesis of biology, management and impacts on lodgepole pine*. Victoria, BC: Natural Resources Canada, Pacific Forestry Centre.
- Shiode, S. (2008). Analysis of a distribution of point events using the network-based quadrat method. *Geographical Analysis*, 40(4), 380-400. <http://dx.doi.org/10.1111/j.0016-7363.2008.00735.x>
- Spooner, P. G., Lunt, I. D., Okabe, A., & Shiode, S. (2004). Spatial analysis of roadside *Acacia* populations on a road network using the network K-function. *Landscape Ecology*, 19(5), 491-499. <http://dx.doi.org/10.1023/B:LAND.0000036114.32418.d4>
- Suzuki, S. N., Kachi, N., & Suzuki, J. I. (2008). Development of a local size hierarchy causes regular spacing of trees in an even-aged *Abies* forest: Analyses using spatial autocorrelation and the mark correlation function. *Annals of Botany*, 102(3), 435-441. <http://dx.doi.org/10.1093/aob/mcn113>
- Tobin, P. C. (2004). Estimation of the spatial autocorrelation function: consequences of sampling dynamic populations in space and time. *Ecography*, 27(6), 767-775. <http://dx.doi.org/10.1111/j.0906-7590.2004.03977.x>
- Tobler, W. R. (1970). Computer movie simulating urban growth in Detroit region. *Economic Geography*, 46(2), 234-240. <http://dx.doi.org/10.2307/143141>
- Vadrevu, K. P., & Badarinath, K. V. S. (2009). Spatial pattern analysis of fire events in Central India - a case

- study. *Geocarto International*, 24(2), 115-131. <http://dx.doi.org/10.1080/10106040802230837>
- Vadrevu, K. P., Badarinath, K. V. S., & Anuradha, E. (2008). Spatial patterns in vegetation fires in the Indian region. *Environmental Monitoring and Assessment*, 147(1-3), 1-13. <http://dx.doi.org/10.1007/s10661-007-0092-6>
- Wulder, M. A., Dymond, C. C., White, J. C., Leckie, D. G., & Carroll, A. L. (2006). Surveying mountain pine beetle damage of forests: A review of remote sensing opportunities. *Forest Ecology and Management*, 221(1-3), 27-41. <http://dx.doi.org/10.1016/j.foreco.2005.09.021>
- Zhang, C. S., Luo, L., Xu, W. L., & Ledwith, V. (2008). Use of local Moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway, Ireland. *Science of the Total Environment*, 398(1-3), 212-221. <http://dx.doi.org/10.1016/j.scitotenv.2008.03.011>