Preliminary Analysis on the Regionalization of Landscape Pattern Indices using Multivariate Cluster Analysis – MBPI Working Paper.

Jed A. Long[1], Trisalyn A. Nelson[1], and Michael A. Wulder[2]

[1] Spatial Pattern Analysis & Research (SPAR) Laboratory, Department of Geography, University of Victoria, PO Box 3060, Victoria, BC V8W 3R4, Canada.

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

MPBI Project # 7.32

forest. forward.
moving beyond the pine beetle
Natural Resources Canada Ressources naturelles Canada
Canada

Abstract

Regionalization, or the grouping of objects in space, provides a useful tool for organizing, visualizing, and synthesizing the information contained in multivariate spatial data. Landscape pattern indices can be used to quantify the spatial pattern (composition and configuration) of land cover features. Observable patterns can be linked to underlying processes affecting the generation of landscape patterns (e.g., forest harvesting). The objective of this research is to develop an approach for investigating spatial distribution of forest pattern across a large region where the occurrence and causes of forest pattern is variable. We generate spatial pattern regions (SPR) that describe forest pattern with a regionalization approach. Analysis is performed using a 2006 land cover dataset covering the Prince George and Quesnel forest districts, 5.5 million ha of primarily forested land base situated within the interior plateau of British Columbia, Canada. Currently, forest land cover in this region is being altered by increased forest harvesting related to insect salvage and mitigation activities. Multivariate cluster analysis using the CLARA (Clustering for LARge Applications) algorithm is used to group landscape objects, using a 1 km^2 analysis unit containing forest pattern information, into SPR. Evaluative criteria are used to determine an optimal clustering level of 6 clusters. Output clusters are labelled to represent levels of forest fragmentation. Of the six generated SPR, SPR2 is the most prevalent covering 22% of the study area. On average landscapes in SPR2 are comprised of 55.5% forest cover, and contain the highest number of patches, and forest/non-forest joins of any of the SPR, indicating highly fragmented landscapes. Underlying processes generating forest patterns can be investigated. The influence of topography on land cover classes is linked to fragmented landscapes located in the Eastern part of our study area near the Rocky Mountains. In the Central and Western portions, anthropogenic disturbances are identified as shaping the forest patterns observed. Our method can be applied across a large range of applications and spatial scales. Regionalization of landscape pattern indices provides a useful approach for examining the spatial distribution of forest pattern.

Keywords: forest fragmentation, regionalization, landscape pattern indices, multivariate cluster analysis, SPR, CLARA

forest. forward.
moving beyond the pine beetle
Natural Resources Canada Ressources naturelles Canada Canada

Table of Contents

List of Tables

List of Figures

1 Introduction

Researchers in a wide variety of disciplines are often concerned with exploring patterns and trends in spatially referenced data. In practice this has led to a number of quantitative methods aimed at exploring and measuring the spatial structure of spatial data including: measures of spatial autocorrelation (Cliff & Ord 1981); geostatistics (Cressie 1993); geographically weighted regression (Fotheringham et al. 2002); and local measures of spatial association (Boots 2002). As well, a number of qualitative techniques for mapping and visualizing data have evolved, serving as valuable tools for viewing patterns and properties of spatial data (DiBiase 1990). When researchers are investigating spatial trends in one (or many) variables, it is often advantageous to group data. Typically this grouping process is termed classification (or clustering), referring to how objects are assigned to classes or clusters (Johnson & Wichern 1982).

Classification broadly refers to the grouping of entities based on common properties or relationships (Sokal 1974). Regionalization, or spatial classification, is a specialized form of classification that deals with geographic data (Chorley & Haggett 1967, Johnston 1968). The geographic entities created through the regionalization process are often used for multiple monitoring, mapping, and management activities. However, often classifications and related maps are best for a single purpose (Grigg 1965, Wood 1992). Moreover, there is considerable disagreement on how to delineate regions, which has led to a number of studies attempting characterize the same features with different regionalization processes (see Omernik 2004).

The spatial datasets used in regionalization are often large and contain several levels of detail making them difficult to view and interpret (Ng & Han 2002). The advantage of regionalization is that it allows large and detailed spatial datasets to be viewed and analyzed in a more manageable way. Cluster analysis has been used extensively with aspatial data and represents a useful tool for exploring groups in spatial data (Ng & Han 2002).

We implement multivariate cluster analysis as a quantitative approach to regionalizing landscape spatial pattern. Multivariate cluster analysis provides a novel approach to the regionalization of landscape spatial pattern. Landscape pattern indices are calculated and cluster analysis is performed on these metrics to generate Spatial Pattern Regions (SPR). SPR represent landscapes that exhibit similar spatial pattern characteristics. By mapping SPR we can explore the spatial distribution of forest pattern across our study area. In a region of British Columbia, Canada where increased forest harvesting is occurring due to insect salvage and mitigation activities SPR are used to identify the spatial distribution of forest pattern and its underlying cause. Furthermore, this is an example of how ecological management can be enhanced using regionalization and landscape pattern information.

1.1 Background

In ecology, the practice of regionalization has developed through the creation of ecological zonations (sometimes referred to as ecoregions). Ecological zonations represent a holistic framework, integrating the significant or enduring environmental characteristics of the landscape into regions with similar properties and potentials that serves as a flexible, multipurpose spatial framework for a wide range of applications (Loveland & Merchant 2004). Table 1 lists regionalization examples from ecology and from a number of other disciplines, and illustrates the wide range of contexts for the creation and use of regions.

A common reason for generating ecological zonations is as a first step in studies to improve our understanding of ecological processes. Ecologists have demonstrated the important linkages between landscape spatial pattern and ecological process for a number of processes (e.g., nutrient sediment loadings to streams, Jones et al. 2001; habitat occupancy by grassland birds, Helzer & Jelinski 1999; organism dispersal, Wiens et al. 1997; and the spread of natural disturbances, Turner et al. 1989). Humans impact landscape spatial pattern through urbanization (Luck & Wu 2002), agriculture (Agger & Brandt 1988), road-development (Riitters & Wickham 2003), and forestry (Franklin & Forman 1987). In forested landscapes the spatial pattern of forest affects the occurrence and spread of natural disturbances, such as fire (Romme 1982, Agee 1998) and insect outbreaks (Radeloff et al. 2000, Barclay et al. 2005), which has a number of important implications for forest management and monitoring. Given important linkages between pattern and process, there are benefits to including a landscape pattern component when performing regionalization.

An alternative goal of regionalization is to provide a spatial context for management and decision making. It has been suggested that the underlying theories of landscape ecology and spatial pattern analysis should be included in ecological planning and management (Haber 1990, Ruzicka & Miklos 1990). When regionalizations used for management and monitoring of ecological processes consider landscape spatial pattern, generated management units will have the benefit of incorporating the fundamental premise of landscape ecology—the link between pattern and ecological process (Turner 1989).

There are only a small number of examples in peer-reviewed literature that explore how landscape pattern can be used in regionalization. The first example, MacPhail (1971), used aerial photography to map landscape spatial patterns and relate them to fabric patterns and textures to aid the visual interpretation of different pattern regions. Similarly, Wickham and Norton (1994), created landscape pattern types defined as a kilometres-wide geographical area throughout which a limited number of land cover categories form a consistent pattern. Wickham and Norton (1994) employ visual interpretation of Landsat Thematic Mapper (TM) imagery in order to derive landscape pattern types. These studies took a qualitative approach using human interpretation and subjectivity for the regionalization process. A quantitative approach may be advantageous as it is more explicit, repeatable, transferable, and defensible (Hargrove & Hoffman 2004). Examples of quantitative approaches to mapping landscape spatial pattern also exist. Riitters et al. (2000) developed a classification of forest fragmentation using two indices of spatial pattern. Classified objects are mapped to examine the spatial distribution of forest fragmentation globally (Riitters et al. 2000) and in the United States (Riitters et al. 2002). Morphological image processing (Soille 2003) has also been utilized for mapping forest components characterized as core, edge, or patch (Vogt et al. 2007).

2 Material and Methods

2.1 Study Area

Two adjacent forest districts within British Columbia's interior plateau were chosen as the study area (Figure 1). The Prince George and Quesnel forest districts cover 5.5 million hectares of primarily forested land base. The climate in the Prince George and Quesnel forest districts is characterized by long, cold winters interspersed with hot, humid summers (Meidinger & Pojar 1991). Forests here are comprised primarily of lodgepole pine (*Pinus contorta*), white spruce (*Picea glauca*) and sub-alpine fir (*Abies lasiocarpa*).

Currently, the largest ever recorded mountain pine beetle (*Dendroctonus ponderosae*) infestation is occurring in this region, causing extensive mortality in lodgepole pine stands. The range of this infestation in British Columbia is estimated to have increased from 166,000 ha in 1999 to 10.1 million ha in 2007 (Westfall & Ebata 2008). Short-term increases to the provincial allowable annual cut (with concessions to come in the future) have been prescribed in the Prince George and Quesnel forest districts as a means to recover economic value from infested timber resources (British Columbia Ministry of Forests and Range 2007). Under previous management scenarios, harvesting practices generally consisted of a series of smaller (<60 ha) forest openings (Eng 2004). Salvage and mitigation activities in response to the mountain pine beetle have the potential to generate larger (>1000 ha) forest openings (Eng 2004).

2.2 Data

A 2006 land cover dataset was created to facilitate the generation of landscape pattern metrics across our study area. Using a change detection method based on Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) data (Han et al. 2007), we update forest conditions of an existing land cover dataset produced by the Earth Observation for Sustainable Development of Forests (EOSD) program (Wulder et al. 2003, 2008a). Land cover is represented at a spatial resolution of 25 m, with up to 23 classes of categorical detail which can be aggregated to: forest, non-forest and other classes (Wulder & Nelson 2003). The forest, non-forest, and other categories provide a useful set of land cover classes for examining the spatial pattern of forests in this region, and is comparable to land cover schemes used for forest fragmentation studies in Canada (Wulder et al. 2008b) and the United States (Riitters et al. 2002).

A regular squared partition (fishnet) was used to generate an encompassing set of smaller analysis units (landscapes) within the study region. A 1 km landscape was chosen to capture the impacts of forest harvesting and insect salvage and mitigation activities. Larger landscape sizes exhibit varying levels of spatial pattern, while smaller landscape sizes tend towards a bifurcation of forest patch or no patch. A 1 km landscape has been demonstrated as an effective analysis unit for monitoring forest fragmentation across Canada (Wulder et al. 2008b). Thus, 1 km landscapes provide the spatial scale for both generating and reporting SPR.

2.3 Analysis

2.3.1 Landscape Pattern Variables

A large number of metrics exist for quantifying the spatial pattern of land cover features. It is typically appropriate to choose a subset of metrics relevant to a specific application (Gergel 2007). Previous work has used correlation analysis and theoretical knowledge to identify key components of landscape spatial pattern (e.g., Riitters et al. 1995, Hargis et al. 1997, Boots 2006), and for choosing relevant metrics. Often four to six primary components of spatial pattern are considered (e.g., Hargis et al. 1999, O'Neill et al. 1999, Lofman & Kouki 2001).

We select five indices of landscape pattern (Table 2) for regionalization in this context. Class proportion effectively defines the composition of the landscape in two class landscapes (Boots 2006), and researchers have demonstrated that class proportion is the driving factor of landscape spatial pattern (Boots 2006, Remmel et al. 2002). Join counts are useful in quantifying the level of spatial clustering (sometimes calculated as contagion; see Li & Reynolds 1993) in landscape components (Boots 2006) and can be related to edge amount, an important habitat component (Ranney et al. 1981). Number of patches is an important factor in monitoring the fragmentation of landscape components (Haines-Young & Chopping 1996). When quantifying patch area Boots (2006) suggests that the sum of squared area of patches provides more information than average patch area as it is more sensitive to patch size distribution (e.g., the difference between one large and one small patch, and two medium patches). We employ an area squared measure to quantify the areal properties of patches; a feature commonly used when monitoring habitat fragmentation (Fahrig 2003). Lastly we calculate the mean patch perimeter-area ratio, which is useful in monitoring the regularity/complexity of patch shapes. Natural landscapes frequently exhibit complex, irregular shapes (Forman 1995), while anthropogenic landscapes generally contain regular shapes and straight edges (Hammett 1992, Forman 1995).

2.3.2 Multivariate Cluster Analysis

Cluster analysis has been referred to as the art of finding groups in data (Kaufman & Rousseeuw 1990). More specifically, cluster analysis is a quantitative statistical method that uses unsupervised learning to explore, find and categorize features and to gain insight on the nature or structure of data (Duda et al. 2001). Clustering algorithms fall into two broad categories: hierarchical or flat-partition (Kaufman & Rousseeuw 1990). Hierarchical methods are advantageous when the initial number of clusters is unknown (Duda et al. 2001); however, hierarchical methods are most suited for the classification of variables rather then objects (Johnson & Wichern 1982) and are computationally constraining with large datasets. Flat-partition methods can better handle large datasets, but require that the user specify a pre-determined number of clusters (k).

The CLARA (Clustering for LARge Applications) algorithm (Kaufman & Rousseeuw 1990) was used to perform cluster analysis. CLARA is a flat-partition method that has been specifically designed for use with large datasets. User definition of the k parameter is required, and since k is unknown we implemented the algorithm for a range of k values (2-10). An optimal clustering level (k) can be chosen iteratively using evaluative criteria (Milligan & Cooper 1985, Halkidi et al. 2002). Cluster analysis is performed using the R statistical software package (R Development Core Team 2008).

2.3.3 Normalization and Weighting

When performing cluster analysis it is often suggested that data be normalized. Normalization (standardization) assigns equal spacing for parameters of varying ranges and units by assigning a zero mean and a unit standard deviation to all data variables (Kaufman & Rousseeuw 1990). We normalize our data following Kaufman & Rousseeuw (1990):

$$z_{if} = \frac{x_{if} - m_f}{s_f} \qquad [1]$$

Where z_{if} is the normalized value for observation i of variable f, x_{if} is the original value for observation i of variable f, m_f is the mean of variable f and s_f is a measure of dispersion for variable f. We use the mean absolute deviation as the measure of dispersion which is defined as:

$$s_f = \frac{1}{n}\left\{ \left| x_{1f} - m_f \right| + \left| x_{2f} - m_f \right| + \cdots + \left| x_{nf} - m_f \right| \right\} \qquad [2]$$

Where n is the number of observations and other variables as in [1]. This dispersion measure is more robust then the standard deviation and is therefore recommended (Kaufman & Rousseeuw 1990).

In many cases, particular attributes may have added meaning or importance for a given analysis. *A priori* knowledge can be a useful tool that can improve a clustering by adding weight to given attributes (e.g., Abrahamowicz 1985). In landscape pattern analysis several studies have identified the importance of land cover composition for the generation of spatial pattern (Gustafson & Parker 1992, Remmel et al. 2002, Boots 2006), and for ecological function (Fahrig 1997). In the absence of quantitative information expert opinion was used to assign weights to input variables on a case specific basis. Considering the relative importance of land cover composition over configuration metrics we increase the weighting of the class proportion metric by a factor of two over the other metrics in our study.

With geographic data we also have the potential to include spatial weighting in subsequent analysis. Spatial weighting refers to increasing similarity, in attribute space, to objects that are spatially proximal to one another. In practice this is often incorporated by including the x and y coordinate values as attributes in the analysis; although other methods exist for creating spatial weights (see Oliver & Webster 1989 for a discussion). Often however, regions tend to be geographically cohesive without spatial weighting because of the spatial autocorrelation present in the data (Hargrove & Hoffman 2004).

Sometimes it is necessary to identify new/different locations that exhibit similar ecological (or other) characteristics of a chosen region for the placement of parks or industrial activities. Objective, statistical methods for regionalization that do not include spatial weighting offer the opportunity for locating spatially disjoint regions with similar characteristics (Coops et al. 2009). Moreover, when it is expected that regions with similar characteristics are spatially distanced (e.g., separated by natural or anthropogenic features), the use of spatial weighting may be unwarranted. We expect forest patterns to portray similar characteristics in spatially distanced regions separated by developed areas or mountain ridges, and because of this expectation we did not include spatial weighting in the generation of our SPR.

2.3.4 Measure of Separation

In cluster analysis it is necessary to calculate a measure of separation between objects. The use of the Euclidean distance measure is common in the ecological literature (e.g., Fovell & Fovell 1993, Gong & Richman 1995), and is easily computed on standardized variables in attribute space. Euclidean distance was implemented as the measure of separation between objects and is calculated using [3].

$$d(i,j) = \sqrt{\left(x_{i1} - x_{j1}\right)^2 + \left(x_{i2} - x_{j2}\right)^2 + \cdots + \left(x_{ip} - x_{jp}\right)^2} \qquad [3]$$

Where d is the distance in attribute space between the i^{th} and j^{th} objects and x is the value of p^{th} attribute. In this example we employ only interval-scaled variables; when using binary, ordinal, nominal or some mixture of variable types, other separation measures are more appropriate (Kaufman & Rousseeuw 1990).

2.3.5 Cluster Evaluative Criteria

Flat partition clustering methods require user definition of the number of clusters (*k*). When *k* is unknown, cluster evaluative criteria provide information for determining an optimal *k*. Measures of cluster strength frequently suggested are often tested on datasets with clearly defined clusters (e.g. Milligan & Cooper 1985). These measures are known to work well with clusters that are compact, however novel approaches to testing cluster validity when data does not exude compact clusters (e.g., spatial data) are needed (Halkidi et al. 2002). Strongly defined clusters are not expected here, as landscapes vary continuously over the range of metrics tested. The Davies-Bouldin index (DB) (Davies & Bouldin 1979) and average silhouette width (ASW) (Kaufman & Rousseeuw 1990) were chosen as evaluative criteria for selecting the optimal *k*.

ASW is calculated using [4], which is a measure of how well clusters are separated from their closest neighbour (Kaufman & Rousseeuw 1990).

$$ASW = \frac{1}{k}\sum_{i=1}^{k}\frac{b_i - a_i}{\max\{b_i, a_i\}} \qquad [4]$$

Where; *a* is the average dissimilarity between objects in a cluster *i*, *b* is the average dissimilarity of the objects in *i* to those in its closest neighbour, and *k* is the number of clusters. The maximal ASW for all *k* is interpreted as the optimal or strongest cluster level (Kaufman & Rousseeuw 1990). Kaufman and Rousseeuw (1990, p. 88) suggest that ASW values of 0.71-1.00 indicate a *k* with well defined clusters, while ASW values of 0.26-0.50 indicate a *k* with weakly defined clusters, and cluster structure may be artificial.

DB uses a ratio of intra-cluster dispersion to inter-cluster separation divided by *k* to determine clustering strength for a given *k* [5] (Davies & Bouldin 1979).

$$DB = \frac{1}{k}\sum_{i=1}^{k}\frac{S_i + S_j}{m_{ij}} \qquad [5]$$

Where; S_i is the dispersion of cluster *i*, S_j is the dispersion of the next closest cluster *j*, m_{ij} is the distance between the cluster centres of *i* and *j*, and *k* is the number of clusters. DB is advantageous in that it does not require user definition of parameters, such as minimum acceptable cluster distance or minimum acceptable standard deviation, which are often unknown (Davies & Bouldin 1979). Optimal *k* is found at the minimum DB value; when intra-cluster dispersion is low and inter-cluster separation is high.

Multivariate cluster analysis allows users to examine many statistical and qualitative properties of each cluster. Descriptive statistics such as mean, median and coefficient of variation were computed for each SPR and landscape metric combination. We generate the relative frequency histogram for each SPR and landscape pattern metric combination to assess the distributional properties of each SPR. Examining the relative frequency histogram adds to interpretation of the landscape pattern properties of each SPR. Also, each cluster that is created by the CLARA algorithm has a *medoid* (or *centrotype*) which is a representative object for each cluster and has been used as a surrogate centre for each cluster during its creation (Kaufman & Rousseeuw 1987). Medoids are a good representation of cluster centres (the middle object in a cluster), that are less sensitive to outliers then other cluster profiles (Van Der Laan et al. 2003). With this algorithm we can extract the medoids and use them as a spatial representation of each SPR.

3 Results

Using cluster evaluative criteria (DB and ASW) an optimal cluster level was identified at $k = 6$ (Figure 2). In Figure 2 we see that DB is minimal at $k = 6$, and ASW is maximum at $k = 2$, but has a second peak when $k = 6$. We choose $k = 6$ as the optimal clustering over the case when $k = 2$ based on the evaluative criteria and because the case where $k = 2$ provides few unique insights on landscape processes (largely representing a bifurcation of landscapes with high and low forest composition).

The mean, median, and coefficient of variation for each of the 6 generated SPR are portrayed in Table 3. We can see that mean and median forest proportion increases going from SPR1 to SPR5, and then decreases slightly to SPR6. The variation in forest proportion is highest when forest proportion is low (i.e., in SPR1 and SPR2). Number of patches is highest in SPR2 and SPR4 and lowest in SPR6. Forest/non-forest joins are substantially higher in SPR2 (mean = 629.5 joins) than in the others, but are most variable in SPR1, SPR5, and SPR6. Patch areas based on the area-squared measure are highest in SPR5 and lowest in SPR2. Average perimeter-area ratio is highest in SPR4, but the most notable result from this is the marked difference between the lowest SPR6 (mean = 527.6 m/ha) and the next lowest SPR3 (mean = 752.3 m/ha). SPR4, SPR5, and SPR6 all have high forest composition (mean > 80%), but SPR4 and SPR5 have high perimeter area ratios (mean = 955.0 m/ha, 879.6 m/ha, respectively), relative to SPR6.

Figure 3 shows the relative frequency histogram for each SPR-landscape metric combination. There is a gradient in forest proportion with SPR1 being low and SPR5 and SPR6 being high. SPR2 and SPR3 have similar distributions for forest proportion. If we compare the configurational attributes between SPR2 and SPR3, we can see that they differ considerably in number of patches and forest/non-forest joins. Based on Table 3 we would expect SPR4, SPR5, and SPR6 to have similar forest proportions, but we can see a clear difference in the distribution of SPR4 from SPR5 and SPR6. SPR4's number of patches and forest/non-forest joins distributions are also noticeably different then that of SPR5 and SPR6. SPR5 and SPR6 have similar distributions for most of the spatial pattern metrics, but differ slightly in number of patches and substantially in average perimeter area ratio. The difference in average perimeter-area ratio between SPR6 and the rest of the SPR is clearly demonstrated by the relative frequency distributions.

Figure 4 contains a view of the medoid landscape of each SPR, as well as a summary of the expected spatial pattern properties for each. This information is useful in interpreting results, as it is a visual reference of the forest pattern expected in each SPR. SPR1 portrays landscapes with low forest composition, and high forest fragmentation. SPR2 and SPR3 contain similar forest composition, but in SPR3 forest is contiguous, while in SPR2 there are more numerous, smaller patches indicating increased forest fragmentation. SPR4 has patch shapes with very irregular edges. SPR5 is characterized by high forest composition with small perforations of non-forest. SPR6 is the least fragmented and has high forest composition and contiguous disturbances, indicative of patch harvests.

A map of the distribution of SPR is shown in Figure 5, and includes a digital elevation model (DEM) to aid in interpretation. SPR0 and SPR100 (0% and 100% forest respectively) are only a minor constituent across our study area representing 2% and 4% of the total area (Table 4). SPR2 is the most prevalent of the SPR (22%), while SPR1 (9%) has the lowest area of any of the generated SPR. Topography, especially along the eastern edge of our study area, plays an important role in landscape spatial pattern, as valleys contain predominantly SPR5 and SPR6, regions with high levels of forest composition, and low number of forest patches, while alpine areas are SPR0, SPR1, SPR2, and SPR3, those with low forest composition, high number of patches, high forest/non-forest joins. The topographic influence on land cover in alpine regions results in these regions being labelled the same as a naturally forested landscape that has similar pattern characteristics resulting from some form of disturbance (e.g., harvesting, natural disturbance). In reality, the patterns observed in alpine areas are natural and generally more static, and should be distinguished from low lying areas where landscape pattern originates from some other process. Anthropogenic activities are expected to be highest near the cities of Prince George and Quesnel (see study area, Figure 1). These areas appear as predominantly SPR1, SPR2, and SPR3, the SPR with the lowest forest proportion, and highest number of patches. In the western portion of the Quesnel forest district (the lower left portion of Figure 4), noticeable pockets of SPR100 (intact forest) and SPR5 and SPR6 (high forest proportion, low number of patches), are interspersed with pockets of SPR1, SPR2, and SPR3 (low forest proportion, high number of patches). This may be an indication of the types of forest harvesting occurring in this area. Here, where topography is less extreme and has less of an influence on the spatial pattern of the landscape, forest harvesting activities are expected to be the driving factor in shaping observed forest patterns.

4 Discussion

Multivariate clustering provides a quantitative method for the regionalization of spatial data (Hargrove & Hoffman 2004). Using landscape pattern metrics, we create Spatial Pattern Regions (SPR) at a 1 km^2 spatial scale. Landscape pattern metrics are selected based on previous studies relating to spatial pattern characteristics (Riitters et al. 1995, Hargis et al. 1997, Boots 2006) and focused on examining forest fragmentation (Wulder et al. 2008b). By considering *a priori* information on the importance of landscape composition, we weight the forest proportion metric by a factor of two over the other four spatial pattern metrics. Generation of relative frequency histograms for each SPR-metric combination proved to be useful for interpretation of SPR properties. For example, based solely on tabulated results SPR4, SPR5, and SPR6 exhibit similar forest composition levels. The use of relative frequency histograms provides added information on SPR4 as it exudes a noticeably different distribution from SPR5 and SPR6. Similarly, extracting the medoid landscapes for each SPR provides a useful visualization tool. Medoid landscapes provide representative examples of the spatial pattern expected in each SPR.

A quantitative assessment to cluster similarity characterizes the relatedness of cluster centres. When applied to spatial datasets, maps of similarity can provide useful results (e.g., Hargrove and Hoffman 2004). We used a qualitative assessment of SPR similarity, labelled as a forest fragmentation gradient. Forest fragmentation can be broadly described as the breaking up of forest habitat into smaller and more numerous parcels (Forman 1995). Based on this concept of fragmentation, SPR1 contains the highest level of forest fragmentation and SPR6 the lowest, with the others falling in between. SPR0 (0% forest) and SPR100 (100% forest) provide external bounds on this scale that represent no-forest and all-forest. Alternative interpretations of forest fragmentation may have considered SPR2 to be more fragmented then SPR1 based on its high number of patches and small patch areas. Similarly, forest composition can be used as a proxy for forest fragmentation (e.g., Wickham et al. 2008) and SPR5 would be interpreted as less fragmented then SPR6. Likewise interpretations of perimeter area ratio can be used to describe the complexity of landscape patch shapes, and relate to the source of forest fragmentation. Anthropogenic disturbance often result in regular shapes, while natural fragmentation processes may lead to more complex shapes (Forman 1995). SPR4, SPR5, and SPR6 were all found to have high forest composition but the small openings in SPR6 have regular shapes (low perimeter area ratio) while those in SPR4 and SPR5 have more complex shapes (high perimeter area ratio). Thus, the source of forest openings in SPR6 are likely the result of anthropogenic disturbance (e.g., harvest), while in SPR4 and SPR5, openings may be of natural origin (e.g., topography, wind-throw), or represent areas where anthropogenic disturbances more appropriately emulate natural patterns.

Any multivariate cluster analysis is dependant on the data, input parameters, and methods applied. We provide an example of multivariate clustering using the CLARA algorithm and two tests for determining the optimal clustering (DB and ASW). Changing the clustering algorithm or the evaluative criteria will impact results. As it was specifically designed for large datasets, the CLARA algorithm is suited for large spatial datasets, where other methods (e.g., hierarchical) are computationally constrained. It is up to the user to explore combinations of algorithms and criteria that are useful and relevant for their research.

The landscape pattern metrics employed in this study represent only a small subset of the suite of metrics available to researchers. Choice of metrics should be related to the ecological questions being investigated (Gergel 2007). Here we investigate effects of large-area forest harvests on landscape spatial pattern, and are interested in monitoring forest fragmentation. We employ metrics useful at quantifying the key components of landscape pattern related to forest fragmentation (Haines-Young & Chopping 1996, Wulder et al. 2008b).

This study was conducted in a region where increased forest harvesting has been prescribed in response to insect infestation (British Columbia Ministry of Forests and Range 2007). Examining the spatial distribution of SPR suggests that in the Eastern parts of the study area topographical influence on land cover is the largest factor affecting forest patterns observed. In the areas with the most anthropogenic activity, located centrally within this study area, the highest levels of forest fragmentation are noticed. In the western portions, especially in the Quesnel forest district, parcels of fragmented landscapes (SPR1, SPR2, and SPR3) are interspersed with non-fragmented landscapes (SPR5, SPR6, and SPR100). Here, where mountain pine beetle infestation is widespread, salvage and mitigation harvesting activities may be the driving factor in shaping forest pattern, and this location may be most useful for investigating the potential impacts on various natural processes, such as wildlife habitat (Bunnell et al. 2004) and hydrologic regimes (Helie et al. 2005).

5 Conclusions

Multivariate cluster analysis provides a useful, quantitative approach for the regionalization of landscape spatial pattern. The CLARA clustering algorithm provides a suitable method for multivariate cluster analysis with large datasets (as is often the case with spatial data). When the desired number of clusters (*k*) is unknown, evaluative criteria can be used to determine an optimal *k*. Output clusters can be characterized by examining their statistical properties, and relative frequency histograms may provide added information. We found the medoid object (landscape) to be useful for representing each cluster visually. Mapping clusters creates a regionalization that provides information for management decision-making and can be related to ancillary datasets (e.g., topography). The use of quantitative methods in regionalization projects has been advocated (Hargrove & Hoffman 2004) and this study provides an example.

Regionalization provides an effective framework for viewing and understanding the characteristics of spatial data. With regionalization we can evaluate the spatial distribution of data groups relatable to forest fragmentation rather then data objects. In this study we focus on a relatively small extent and investigate the effect of increased harvesting from mountain pine beetle salvage and mitigation activities on forest pattern. Our approach is transferable across spatial scales and applications. Recent international forest management protocol has advocated for national forest management practices that include consideration of forest pattern (Montreal Process Liaison Office 2000). In Canada, the spatial extent of forest monitoring limits our ability to visualize and interpret forest pattern information. Similarly, the number of attributes required to effectively monitor forest pattern is not easily visualized with maps. Regionalization may provide an effective approach for meeting these monitoring directives.

6 Acknowledgements

This project was funded by the Government of Canada through the Mountain Pine Beetle Initiative, a six-year, $40 million program administered by Natural Resources Canada, Canadian Forest Service. Publication does not necessarily signify that the contents of this report reflect the views or policies of Natural Resources Canada – Canadian Forest Service. Chris Butson and Dr. Xiaoping Yuan of the British Columbia Ministry of Forests and Range and Joanne White of the Canadian Forest Service are thanked for insight and access to data critical to the success of this research.

7 Literature Cited

Abrahamowicz, M. 1985. The use of non-numerical *a priori* information for measuring dissimilarities. paper presented at the Fourth European Meeting of the Psychometric Society and the Classification Societies, July 2–5, Cambridge, UK.

Agee, J.K. 1998. The landscape ecology of western forest fire regimes. Northwest Science 72: 24–34.

Agger, P. & Brandt, J. 1988. Dynamics of small biotopes in Danish agricultural landscapes. Landscape Ecology 1:227-240.

Barclay, H., Li, C. Benson, L., Taylor, S.W., & Shore, T. 2005. Effects of fire return rates on traversability of lodgepole pine forests for mountain pine beetle (Coleoptera: Scolytidae) and the use of patch metrics to estimate traversability. Canadian Entomologist 137: 566–583.

British Columbia Ministry of Forests and Range. 2007. Timber supply and the mountain pine beetle infestation in British Columbia: 2007 update. British Columbia Ministry of Forests and Range, Forest Analysis and Inventory Branch, Victoria, BC. 38p.

Boots, B. 2002. Local measures of spatial association. Ecoscience 9: 168–176.

Boots, B. 2006. Local configuration measures for categorical spatial data: binary regular lattices. Journal of Geographical Systems 8: 1–24.

Bunnell, F.L., Squires, K.A. & Houde, I. 2004. Evaluating effects of large-scale salvage logging for mountain pine beetle on terrestrial and aquatic vertebrates. Pacific Forestry Centre, Canadian Forest Service, Natural Resources Canada, Victoria, BC. 57p.

Chorley, R.J., & Haggett, P. 1967. Models in Geography. Methuen, London. 816p.

Cliff, A.D., & Ord, J.K. 1981. Spatial Processes: Models and Applications. Pion, London. 266p.

Coops, N.C., Wulder, M.A. & Iwanicka, D. 2009. An environmental domain classification of Canada using earth observation data for biodiversity assessment. Ecological Informatics 4:8–22.

Cressie, N. 1993. Statistics for Spatial Data. John Wiley & Sons Inc., New York. 900p.

Davies, D.L., & Bouldin, D.W.. 1979. A cluster separation measure. IEEE Transactions on Pattern Analysis and Machine Intelligence 1: 224–227.

DiBiase, D. 1990. Visualization in the earth sciences. Earth and Mineral Sciences Bulletin 59: 13–18.

Duda, R.O., Hart, P.E. & Stork, D.G. 2001. Pattern Classification. John Wiley & Sons, New York. 654p.

Eng, M. 2004. Forest stewardship in the context of large-scale salvage operations: an interpretation paper. British Columbia Ministry of Forests, Forest Science Program, Technical Report 019, Victoria, BC. 10p.

Fahrig, L. 1997. Relative effects of habitat loss and fragmentation on population extinction. The Journal of Wildlife Management 61:603-610.

Fahrig, L. 2003. Effects of habitat fragmentation on biodiversity. Annual Review of Ecology and Systematics 34: 487–515.

Forman, R.T.T. 1995. Land Mosaics: The Ecology of Landscapes and Regions. Cambridge University Press, Cambridge, UK. 632p.

Fotheringham, A.S., Brunsdon, C. & Charlton, M. 2002. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. Wiley, Chichester. 282p.

Fovell, R.G. & Fovell, M.Y.C. 1993. Climate zones of the conterminous United States defined using cluster analysis. Journal of Climate 6: 2103–2135.

Franklin, J.F., & Forman, R.T.T. 1987. Creating landscape patterns by forest cutting: Ecological consequences and principles. Landscape Ecology 1: 5–18.

Gergel, S.E. 2007. New directions in landscape pattern analysis and linkages with remote sensing. In: Wulder, M.A. & Franklin, S.E. (Ed). Understanding Forest Disturbance and Spatial Pattern: Remote Sensing and GIS Approaches. Taylor & Francis Group, Boca Raton, FL. Pp 173–208.

Gong, X., & Richman, M.B. 1995. On the application of cluster analysis to growing season and precipitation data in North America East of the Rockies. Journal of Climate 8: 897–931.

Grigg, D. 1965. The logic of regional systems. Annals of the Association of American Geographers 55: 465–491.

Gustafson, E.J., & Parker, G.R. 1992. Relationships between land cover proportion and indices of landscape spatial pattern. Landscape Ecology 7: 101–110.

Haber, W. 1990. Using landscape ecology in planning and management. In: Zonneveld, I.S. & Forman, R.T.T. (Ed). Changing Landscapes: An Ecological Perspective. Springer-Verlag, New York. Pp: 217–232.

Haines-Young, R., & Chopping, M. 1996. Quantifying landscape structure: a review of landscape indices and their application to forested landscapes. Progress in Physical Geography 20: 418–445.

Halkidi, M., Batistakis, Y. & Vazirgiannis, M. 2002. Clustering validity checking methods: part II. SIGMOD Record 31: 19–27.

Hammett, J.E. 1992. The shapes of adaptation: Historical ecology of anthropogenic landscapes in the Southeastern United States. Landscape Ecology 7: 121–135.

Han, T., Wulder, M.A., White, J.C., Coops, N.C., Alvarez, M.F. & Butson, C. 2007. An efficient protocol to process Landsat images for change detection with tasselled cap transformation. IEEE Geoscience and Remote Sensing Letters 4: 147–151.

Harff, J., & Davis, J.C. 1990. Regionalization in geology by multivariate classification. Mathematical Geology 22: 573–588.

Hargis, C.D., Bissonette, J.A. & David, J.L. 1997. Understanding measures of landscape pattern. In: Bissonette, J.A. (Ed). Wildlife and Landscape Ecology: Effects of Pattern and Scale. Springer-Verlag, New York, NY. Pp: 231–261

Hargis, C.D., Bissonette, J.A. & Turner, D.L. 1999. The influence of forest fragmentation and landscape pattern on American martens. Journal of Applied Ecology 36: 157–172.

Hargrove, W.W. & Hoffman, F.M. 2004. Potential of multivariate quantitative methods for delineation and visualization of ecoregions. Environmental Management 34: S39–S60.

Helie, J.F., Peters, D.L., Tattrie, K.R. & Gibson, J.J. 2005. Review and synthesis of potential hydrologic impacts of mountain pine beetle and related harvesting activities in British Columbia. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Mountain Pine Beetle Initiative Working Paper 2005-23, Victoria, BC. 26p.

Helzer, C.J., & Jelinski, D.E. 1999. The relative importance of patch area and perimeter-area ratio to grassland breeding birds. Ecological Applications 9: 1448–1458.

Holdridge, L.R. 1967. Life Zone Ecology. Tropical Science Center, San Jose, Costa Rica. 206p.

Johnson, R.A., & Wichern, D.W. 1982. Applied Multivariate Statistical Analysis. Prentice-Hall Inc., Englewood Cliffs, NJ. 594p.

Johnston, R.J. 1968. Choice in classification: the subjectivity of objective methods. Annals of the Association of American Geographers 58: 575–589.

Jones, K.B., Neale, A.C., Nash, M.S., Van Remortel, R.D., Wickham, J.D., Riitters, K.H & O'Neill, R.V. 2001. Predicting nutrient and sediment loadings to streams from landscape metrics: A multiple watershed study from the United States Mid-Atlantic Region. Landscape Ecology 16: 301–312.

Kaufman, L. & Rousseeuw, P.J. 1987. Clustering by means of medoids. In: Dodge, Y. (Ed). Statistical Data Analysis Based on the L1 Norm and Related Methods. Elsevier, North Holland. Pp: 405–416.

Kaufman, L., & Rousseeuw, P.J. 1990. Finding Groups in Data. John Wiley & Sons, New York. 342p.

Lark, R.M. 1998. Forming spatially coherent regions by classification of multi-variate data: an example from the analysis of maps of crop yield. International Journal of Geographical Information Science 12: 83–98.

Leathwick, J.R., Overton, J.M., & McLeod, M. 2003. An environmental domain classification of New Zealand and its use as a tool for biodiversity management. Conservation Biology 17: 1612–1623.

Li, H., & Reynolds, J.F. 1993. A new contagion index to quantify spatial patterns of landscapes. Landscape Ecology 8: 155–162.

Lofman, S., & Kouki, J. 2001. Fifty years of landscape transformation in managed forests of Southern Finland. Scandinavian Journal of Forest Research 16: 44–53.

Loveland, T.R., & Merchant, J.W. 2004. Ecoregions and ecoregionalization: Geographical and ecological perspectives. Environmental Management 34: S1–S13.

Luck, M., & Wu, J. 2002. A gradient analysis of urban landscape pattern: a case study from the Phoenix metropolitan region, Arizona USA. Landscape Ecology 17: 327–339.

MacPhail, D.D. 1971. Photomorphic mapping in Chile. Photogrammetric Engineering 37: 1139–1148.

Meidinger, D. & Pojar, J. 1991. Ecosystems of British Columbia. Research Branch British Columbia Ministry of Forests. Special Report Series Vol. 06., Victoria, B.C. 330p.

Milligan, G.W. & Cooper, M.C. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50: 159–179.

Montreal Process Liaison Office. 2000. Montreal process year 2000 progress report: Progress and innovation in implementing criteria and indicators for the conservation and sustainable management of temperate and boreal forests. The Montreal Process Liaison Office, Canadian Forest Service, Ottawa, Canada.

Ng, R.T. & Han, J. 2002. CLARANS: A method for clustering objects for spatial data mining. IEEE Transactions on Knowledge and Data Engineering 14: 1003–1016.

O'Neill, R.V., Riitters, K.H., Wickham, J.D. & Jones, K.B. 1999. Landscape pattern metrics and regional assessment. Ecosystem Health 5: 225–233.

Oliver, M.A & Webster, R. 1989. A geostatistical basis for spatial weighting in multivariate classification. Mathematical Geology 21: 15–35.

Omernik, J.M. 2004. Perspectives on the nature and definition of ecological regions. Environmental Management 34: S27–S38.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Radeloff, V.C., Mladenoff, D.J. & Boyce, M.S. 2000. The changing relation of landscape patterns and jack pine budworm populations during an outbreak. Oikos 90: 417–430.

Ranney, J.W., Bruner, M.C. & Levenson, J.B. 1981. The importance of edge in the structure and dynamics of forest islands. In: Burgess, R.L. & Sharpe, D.M. (Ed). Forest Island Dynamics in Man-Dominated Landscapes. Springer-Verlag, New York, NY. Pp: 67-96

Remmel, T.K., Csillag, F., Mitchell, S.W. & Boots, B. 2002. Empirical distributions of landscape pattern indices as functions of classified image composition and spatial structure. Joint Symposium on Geospatial Theory, Processing and Applications, Ottawa, ON. 10p.

Riitters, K.H. & Wickham, J.D. 2003. How far to the nearest road? Frontiers in Ecology and the Environment 1: 125–129.

Riitters, K.H., O'Neill, R.V., Hunsaker, C.T., Wickham, J.D., Yankee, D.H., Timmins, S.P., Jones, K.B. &. Jackson, B.L. 1995. A factor analysis of landscape pattern and structure metrics. Landscape Ecology 10: 23–29.

Riitters, K.H., Wickham, J.D., O'Neill, R.V., Jones, K.B. & Smith, E.R. 2000. Global-scale patterns of forest fragmentation. Ecology and Society (formerly Conservation Ecology) 4: Article 3.

Riitters, K.H., Wickham, J.D., O'Neill, R.V., Jones, K.B., Coulston, J.W., Wade, T.G. & Smith, J.H. 2002. Fragmentation of continental United States forests. Ecosystems 5: 815–822.

Romme, W.H. 1982. Fire and landscape diversity in sub-alpine forests of Yellowstone National Park. Ecological Monographs 52: 199–221.

Ruzicka, M. & Miklos, L. 1990. Basic premises and methods in landscape ecological planning and optimization. In: Zonneveld, I.S. & Forman, R.T.T. (Ed). Changing Landscapes: An Ecological Perspective. Springer-Verlag, New York. Pp: 233–260.

Schultz, J. 2005. The Ecozones of the World. Springer-Verlag, Berlin.

Soille, P. 2003. Morphological Image Analysis: Principles and Applications, 2nd ed. Springer-Verlag, Berlin. 252p.

Sokal, R.R. 1974. Classification: purposes, principles, progress, prospects. Science 185: 1115–1123.

Turner, M.G. 1989. Landscape ecology: The effect of pattern on process. Annual Review of Ecology and Systematics 20: 171–197.

Turner, M.G., Gardner, R.H., Dale, V.H., & O'Neill, R.V. 1989. Predicting the spread of disturbance across heterogeneous landscapes. Oikos 55: 121–129.

Van Der Laan, M.J., Pollard, K.S. & Bryan, J. 2003. A new partitioning around medoids algorithm. Journal of Statistical Computation and Simulation 73: 575–584.

Vogt, P., Riitters, K.H., Estreguil, C., Kozak, J., Wade, T.G. & Wickham, J.D. 2007. Mapping spatial patterns with morphological image processing. Landscape Ecology 22: 171–177.

Westfall, J., & Ebata, T. 2008. 2007 Summary of forest health conditions in British Columbia. Ministry of Forests and Range, Forest Practices Branch, Victoria, BC. 81p.

Wickham, J.D., & Norton, D.J. 1994. Mapping and analyzing landscape patterns. Landscape Ecology 9: 7–23.

Wickham, J.D., Riitters, K.H., Wade, T.G., & Homer, C. 2008. Temporal change in fragmentation of continental US forests. Landscape Ecology 23: 891–898.

Wiens, J.A., Schooley, R.L. & Weeks Jr, R.D. 1997. Patchy landscapes and animal movements: Do beetles percolate? Oikos 78: 257–264.

Wiken, E.B. 1986. Terrestrial ecozones of Canada. Ecological Land Classification Series No. 19, Environment Canada, Hull., Hull, PQ. 26p.

Wood, D. 1992. The Power of Maps. The Guilford Press, New York, NY. 248p.

Wulder, M.A., Dechka, J.A., Gillis, M.D., Luther, J.E., Hall, R.J., Beaudoin, A. & Franklin, S.E. 2003. Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program. The Forestry Chronicle 79: 1075–1083.

Wulder, M.A. & Nelson, T. 2003. EOSD land cover classification legend report: Version 2. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, BC. 81p.

Wulder, M.A., White, J.C., Cranny, M., Hall, R.J., Luther, J.E., Beaudoin, A., Goodenough, D.G. & Dechka, J.A. 2008a. Monitoring Canada's forests - Part 1: Completion of the EOSD land cover project. Canadian Journal of Remote Sensing 34: 549–562.

Wulder, M.A., White, J.C. Han, T., Coops, N.C., Cardille, J., Holland, T. & Grills, D. 2008b. Monitoring Canada's forests - Part 2: National forest fragmentation and pattern. Canadian Journal of Remote Sensing 34: 563–584.

Contacts:

Trisalyn Nelson
Spatial Pattern Analysis and Research (SPAR) Laboratory, Department of Geography, University of Victoria, P.O. Box 3060, Victoria, BC, V8W 3R4, Canada.
Email: Trisalyn@uvic.ca

Table 1: Examples using regionalization in ecology, and other disciplines.

Regionalization	Context	Reference
Life Zone	Classify vegetation regions using elevation, temperature, precipitation, and evaporation.	Holdridge (1967)
Ecozone	Global regionalization of a set of five physical characteristics: climate, relief & drainage, soils, vegetation and animals, and land use.	Schultz (2005)
Ecoprovinces/Ecozones/ Ecoregions	Nested system of ecological zones covering Canada using various input parameters.	Wiken (1986)
Environmental Domain* (New Zealand)	To identify unique ecosystems for protection. Used climate and landform variables.	Leathwick et al. (2003)
Environmental Domain* (Canada)	For monitoring biodiversity. Used land cover, productivity, and elevation data.	Coops et al. (2009)
Spatially coherent regions*	Derive spatially contagious regions of crop yield, using crop yield data.	Lark (1998)
Geologic Regions*	For oil and gas exploration. Used structural, petrographic, and petrophysical features.	Harff & Davis (1990)
Climate Zones*	Extract climate zones for conterminous United States. Used long-term monthly temperature and precipitation data.	Fovell & Fovell (1993)

*Use multivariate cluster analysis.

Table 2: Metrics chosen for multivariate cluster analysis, their formulation and selected reference.

Metric	Units	Formulation	Selected Reference
Class Proportion	%	$\frac{\sum a_i}{A} \rightarrow \{i = F\}$	Hargis et al. 1997
Join Counts	#	$\sum g_{jk} \rightarrow \{j = F, k = N\}$	Boots 2003
Number of Patches	#	$\sum n_i \rightarrow \{i = F, N, O\}$	Haines-Young & Chopping 1996
Patch Area Squared	ha^2	$\sum a_i^2 \rightarrow \{i = F, N, O\}$	Boots 2006
Mean Patch Perimeter-Area Ratio	m/ha	$\frac{1}{n}\sum \frac{p_i}{a_i} \rightarrow \{i = F, N, O\}$	Riitters et al. 1995

A – total area of landscape, a – area of patch, g – join between two neighbouring cells, n number of patches, p – perimeter of patch, F – forest, N – non-forest, O – other.

Table 3: Mean, median and coefficient of variation for each metric-SPR combination.

	(units)	**Forest Proportion** (%)	**Number of Patches** (#)	**Forest/Non-Forest Joins** (#)	**Squared Area of Patches** (ha^2)	**Average Patch Perimeter Area Ratio** (m/ha)
SPR1	mean	18.9	14.9	277.3	5647	831.1
	median	19	15	284	5484	838.9
n = 5271	c.v.	0.60	0.40	0.49	0.24	0.14
SPR2	mean	55.5	24.0	629.5	3481	889.2
	median	57	23	616	3442	888.7
n = 12150	c.v.	0.25	0.26	0.21	0.21	0.08
SPR3	mean	65.7	12.7	367.6	4021	752.3
	median	68	13	362	3964	762.6
n = 11548	c.v.	0.19	0.29	0.28	0.22	0.14
SPR4	mean	84.7	21.4	371.1	5907	955.0
	median	85	20	365	5917	952.0
n = 8027	c.v.	0.07	0.28	0.27	0.13	0.08
SPR5	mean	93.5	9.7	153.4	7172	879.6
	median	95	10	148	7315	871.6
n = 10034	c.v.	0.05	0.35	0.50	0.09	0.11
SPR6	mean	88.3	5.2	137.6	6675	527.6
	median	92	5	128	6959	585.9
n = 5775	c.v.	0.15	0.49	0.62	0.18	0.34

Table 4: Area percentages of each SPR.

SPR	Area (%)
SPR0	2
SPR1	9
SPR2	22
SPR3	21
SPR4	14
SPR5	18
SPR6	10
SPR100	4

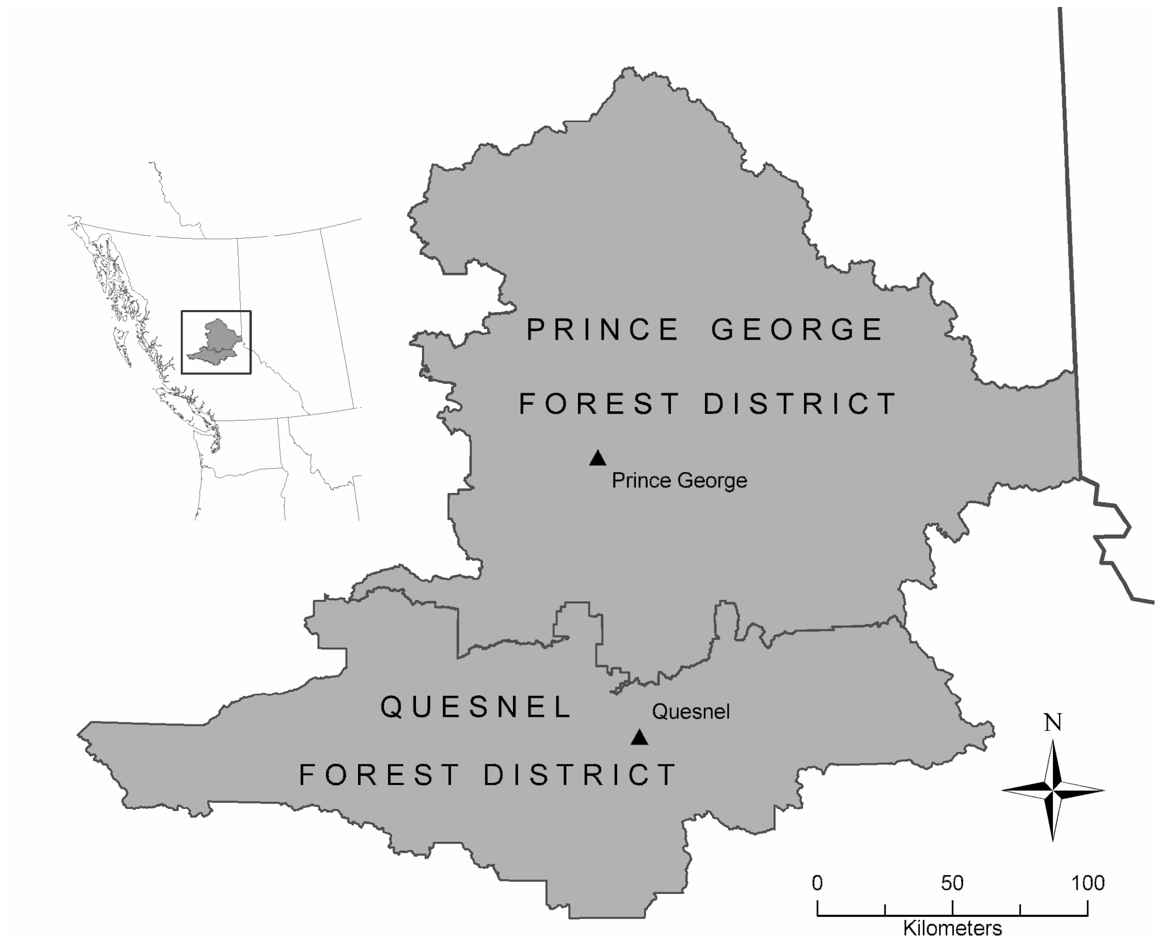


Figure 1: Study area, the Prince George and Quesnel forest districts located in British Columbia, Canada.

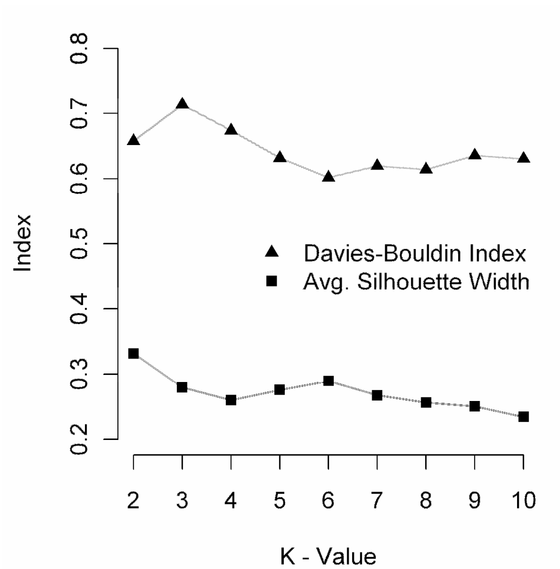


Figure 2: Davies-Bouldin Index (DB) and Average Silhouette Width (ASW) results for *k* values of 2 – 10. Optimal *k* is found at minimum DB and maximum ASW (in this case *k* = 6).

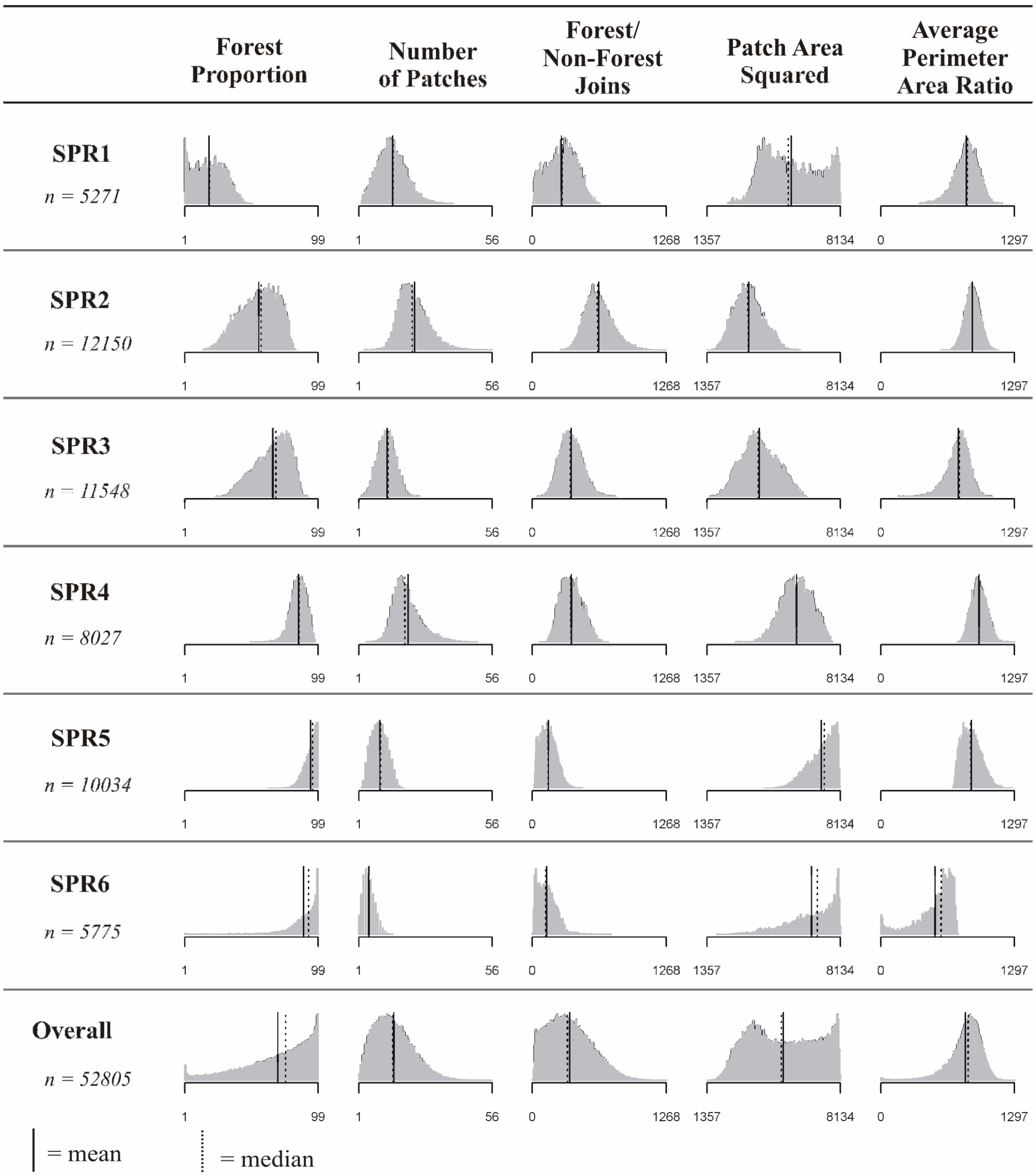
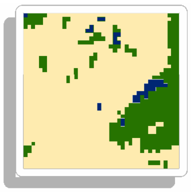
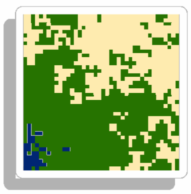


Figure 3: Relative frequency histogram for each metric-SPR combination. Included are the mean and median values for each histogram.



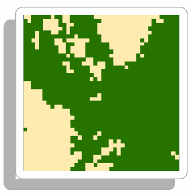
SPR1

SPR1 has the lowest level of forest proportion. It has moderate but variable number of patches and forest/non-forest join counts. It has high squared area of patches, but these are also quite variable. Perimeter-area ratio for SPR1 is very similar to the overall mean. This SPR has the highest level of forest fragmentation.



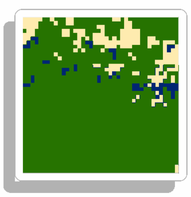
SPR2

SPR2 portrays medium levels of forest proportion, similar to that of SPR3. It has on average the most number of patches, which provide the lowest area squared value.The highest level of forest/non-forest joins was seen in this class indicating the least amount of spatial dependence in forest cover, and most forest edge.



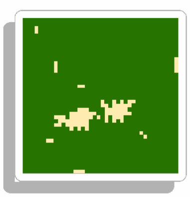
SPR3

SPR3 is comprised of similar forest proportion to SPR2. Conversely to SPR2 it has a low number of patches and lower number of forest/non-forest joins. patch areas are comparable to SPR2, but low compared to other SPRs.



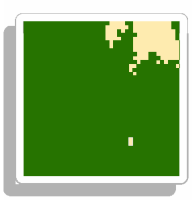
SPR4

SPR4 is a bit of an anomoly. It is characterized by relatively high forest composition, and high patch numbers but moderate forest/non-forest joins. Patch area squared measure for SPR4 is comparable to SPR1. SPR4 has the highest average perimeter area ratio of any of the SPRs.



SPR5

SPR5 has the highest mean forest proportion. It is characterized by a low number of patches and low forest/non-forest joins. SPR5 has high patch area squared measurements but similar perimeter area ratio to the overall mean.



SPR6

SPR6 has the second highest mean forest proportion, and is most comparable to SPR5. It has lower number of patches then SPR5. Forest/non-forest joins and patch area squared measure are highly variable for SPR6. It's distinguishing feature is that it contains the lowest average perimiter-ratio of all SPRs. It is the SPR with the lowest forest fragmentation.

Forest Non-Forest Other

Figure 4: Medoid landscapes for each SPR. Medoids are the central object in each cluster of the multivariate clustering. They are the representative landscape for each SPR. SPR0 and SPR100 are not shown but represent no forest and all forest respectively.

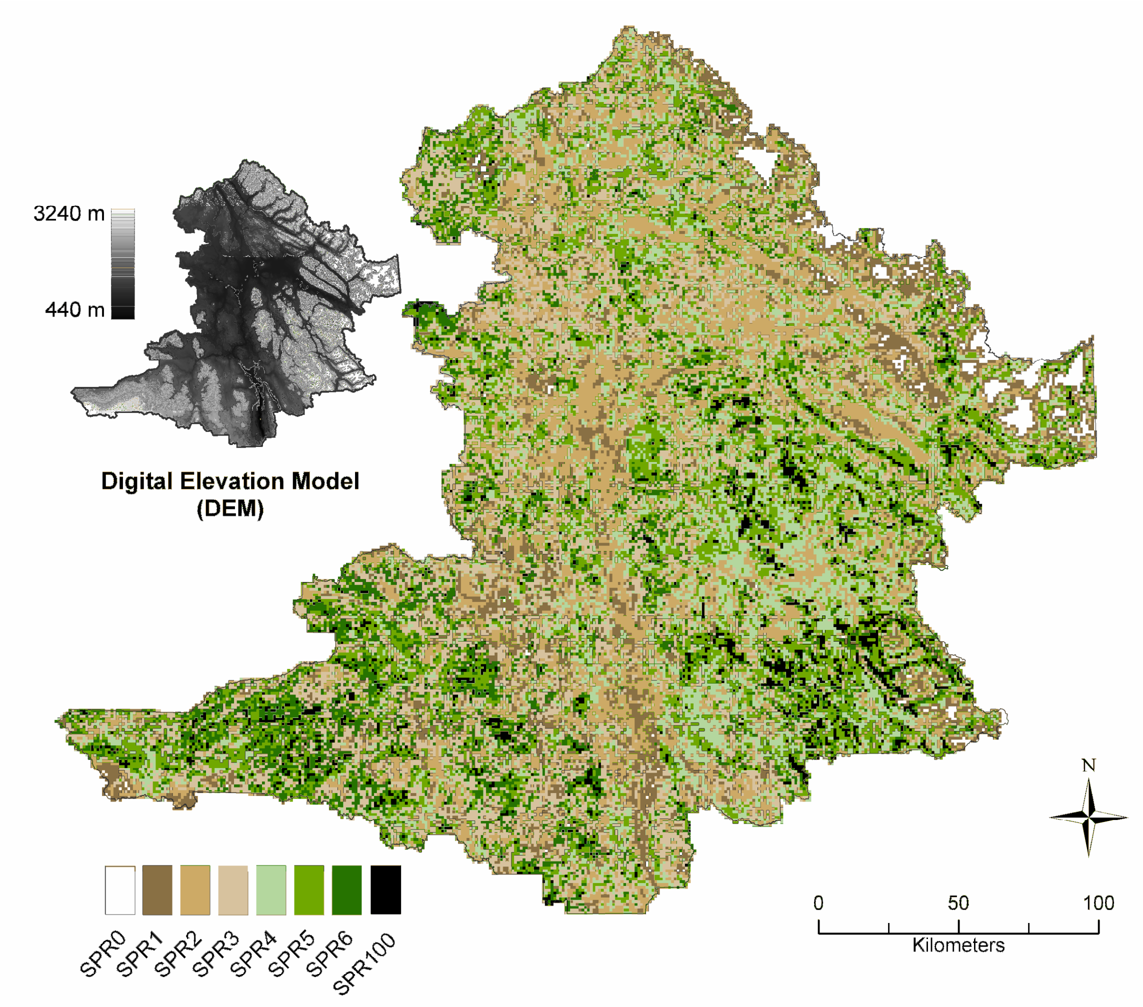


Figure 5: Map of SPR across the Prince George and Quesnel forest districts in British Columbia, Canada.