Can. J. Remote Sensing, Vol. 37, No. 4, pp. 387–402, 2011

Forest complexity modelling and mapping with remote sensing and topographic data: a comparison of three methods

Valerie Torontow and Douglas King

Abstract. The concept of forest complexity has been recently adopted to represent the multiple horizontal and vertical forest structure and composition attributes that support ecological functions in a single measure or index. The index variables are often selected based on known associations with types of habitat and biodiversity potential. In modelling and mapping of forest complexity using geospatial data, three multivariate methods have been evaluated in previous studies: (*i*) defining an additive complexity index by manual a priori selection and combination of a set of field variables followed by regression-based modelling of the index against geospatial data; (*ii*) as the previous method, but where a complexity index is defined a priori using principal components analysis (PCA) of the field data; and (*iii*) direct modelling of a set of field variables against a set of geospatial variables using techniques such as redundancy analysis (RDA). The objective of this study was to compare these methods through an assessment of model quality and their relative merits and limitations in implementation. In the rural municipality of Chelsea, Quebec, 70 field plots were established, and 24 forest structure and composition variables were measured that had between-variable correlations of less than 0.8. Spectral and spatial Quickbird imagery information and topographic data were used to derive complexity models using the three methods. The manual additive index and the RDA-derived index had similar validation errors relative to their index means (24.3% and 22.3%, respectively). However, the RDA index was based on the full set of 24 field variables, which had a total structure-composition variance greater than the manual additive index comprised of a subset of 10 of those variables. Thus, the RDA index was deemed to more comprehensively represent forest structure and composition than the additive index. RDA also produced outputs that were richer in information content, showing associations between individual variables and plots, as well as forest–environmental gradients. The PCA method was useful for evaluating environmental gradients in the field data, but dimensionality was too high to provide a single useful complexity index for modelling with geospatial data. The additive and RDA index models were used to produce maps of predicted forest complexity to aid in biodiversity survey planning and habitat conservation efforts within the municipality.

Résumé. Le concept de complexité de la forêt a été adopté récemment pour représenter la structure horizontale et verticale multiple de la forêt et les attributs de la composition qui supportent les fonctions écologiques dans une mesure ou un indice unique. Les variables de l'indice sont souvent sélectionnées sur la base d'associations connues avec les types d'habitats et le potentiel de biodiversité. Dans la modélisation et la cartographie de la complexité forestière utilisant des données géospatiales, trois méthodes multivariées ont été évaluées dans des études antérieures : (*i*) en définissant un indice additif de complexité par la sélection manuelle a priori et la combinaison d'un ensemble de variables de terrain, suivi de la modélisation basée sur une régression de l'indice par rapport aux données géospatiales, (*ii*) comme ci-dessus, mais où un indice de complexité est défini a priori en utilisant une analyse en composantes principales (ACP) des données de terrain et (*iii*) la modélisation directe d'un ensemble de variables de terrain par rapport à un ensemble de variables géospatiales en utilisant des techniques comme l'analyse de redondance. L'objectif de cette étude était de comparer ces trois méthodes par le biais d'une évaluation de la qualité des modèles et de leur potentiel et limites dans leur implémentation. Dans la municipalité rurale de Chelsea, au Québec, 70 parcelles ont été établies et 24 variables de la structure et de la composition de la forêt qui affichaient des corrélations inter-variables de moins de 0,8 ont été mesurées. L'information spectrale et spatiale dérivée des images de Quickbird ainsi que des données topographiques ont été utilisées pour dériver des modèles de complexité à l'aide des trois méthodes. L'indice additif manuel et l'indice dérivé de l'analyse de redondance montraient des erreurs de validation similaires par rapport à leur indice (respectivement de 24,3 % et de 22,3 %). Cependant, l'indice dérivé de l'analyse de redondance était basé sur l'ensemble complet des 24 variables de terrain, qui avait une variance totale structure-composition supérieure à celle de l'indice additif manuel composé d'un ensemble de 10 de ces variables. Ainsi, l'indice dérivé de l'analyse de redondance a été jugé comme représentant de façon plus détaillée la structure et la composition de la forêt par rapport à l'indice additif. L'indice dérivé de l'analyse de redondance a aussi donné des

Received 3 March 2011. Accepted 18 November 2011. Published on the Web at http://pubs.casi.ca/journal/cjrs on 22 February 2012.

Valerie Torontow, and Douglas King.[1] Geography and Environmental Studies, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, Canada, K1S 5B6.

[1]Corresponding author (e-mail: doug_king@carleton.ca).

résultats plus riches en contenu d'information, mettant en relief des associations entre des variables et des parcelles individuelles de même que des gradients forestiers et environnementaux. La méthode ACP était utile pour évaluer les gradients environnementaux dans les données de terrain, mais la dimensionnalité était trop grande pour fournir un indice de complexité unique utile pour la modélisation avec des données géospatiales. Le modèle d'indice additif et le modèle dérivé de l'analyse de redondance ont été utilisés pour produire des cartes de complexité prédite de la forêt en soutien à la planification des relevés de la biodiversité et aux efforts de conservation des habitats dans la municipalité.
[Traduit par la Rédaction]

Introduction

Sustainable forest management requires monitoring, measurement, and assessment of forest habitat. Habitat loss typically results in declining biodiversity (Lawton et al., 1998; Fahrig, 2003; Gaston et al., 2003; Steffan-Dewenter et al., 2007) as disturbances from anthropogenic and natural causes result in altered vegetation growth, structure, and composition (Linke et al., 2007; Pisaric et al., 2008) and may reduce the provision of habitat resources (Rapport and Whitford, 1999; Danchuk and Willson, 2010). Field-based biodiversity assessment and monitoring are time consuming, even for simple metrics such as species richness, so surrogate indicators have been developed. Vegetation composition and structure have been shown to be strong indicators of the capacity for provision of habitat and biodiversity, amongst other ecological functions (Noss, 1999; Lindenmayer et al., 2000; Tews et al., 2004; Smith et al., 2008; McDermid et al., 2009), and heterogeneous habitat is known to generally support a greater number of species than homogeneous habitat (MacArthur and MacArthur, 1961; Newsome and Catling, 1979; Williams et al., 2002; Lassau et al., 2005). Recent research has focused on integration of forest attributes in indices of forest "complexity" (Neumann and Starlinger, 2001; Staudhammer and Lemay, 2001; Zenner, 2004; McElhinny et al., 2006; Lemay and Newton, 2007). Complexity indices typically include attributes representing the abundance and variability of horizontal and vertical structure (McElhinny et al., 2005), where horizontal structure describes patterns of canopy openness, tree size, and spacing; and vertical structure describes the number, spatial density, and arrangement of vegetation layers. Compositional information can also be included in such indices. For example, the U.S. Forest Service (2006) defines a complex forest as including: (*i*) multiple tree species, (*ii*) trees of different ages, (*iii*) a wide range of tree sizes, (*iv*) abundant and sometimes diverse understory and ground layer plant communities, and (*v*) abundant amounts of standing and fallen dead wood. Hereafter, unless specified, "forest complexity" refers to combined structure and composition diversity.

Derivation of a field-based complexity index for measurement in sample plots (e.g., McElhinny et al. (2006) for structural complexity) is a direct means to assess habitat heterogeneity. It generally involves selection of a set of forest attributes known to provide desired ecological values and services. The attributes may be weighted and additively combined, where weights may be selected subjectively based on the relative importance of each variable, or through objective means such as statistical analysis. Such a field-based index allows for comparison of forest complexity between measurement locations, but it is not amenable to spatial analysis and mapping unless many sample plots can be established throughout an area and spatial gradients are smooth enough to allow for interpolation between plots. Remote sensing and other geospatial data have strong potential for mapping forest complexity because of their explicit representation of space and scale, often at given temporal intervals. Image brightness represents radiance that is reflected from multiple elements of the forest canopy, including leaves, twigs, and branches. The abundance and three-dimensional spatial arrangement of these elements are determined by species composition, crown structure, tree height, and tree spacing and are manifested as crown image brightness and shadow variations (Franklin and Strahler, 1988). These image properties are also indirectly associated with vertical structure as light penetration into a more open upper canopy generally results in more vertically distributed vegetation (Adams et al., 1990), which can in turn affect shadow brightness (Seed and King, 2003; Pasher and King, 2010). Thus, these associations of image brightness variations with forest structure and composition variations provide strong potential for modelling and mapping of an integrated combination of structure and composition parameters in a complexity index.

Three methods were most commonly used in previous studies where the goal was to develop a type of ecological (or complexity) index combining multiple forest attributes for mapping with geospatial data.

i. An additive index defined by manual selection of field variables, as described previously, was modelled against remotely sensed variables, typically using regression (Cohen and Spies, 1992; Coops and Catling, 1997; Estes et al., 2010).
ii. Statistical techniques such as principal component analysis (PCA) were applied to field data to extract components (PCs) that represent ecological gradients. The PCs were then used as indices and regressed against remotely sensed data (e.g., Wunderle et al., 2007; Estes et al., 2010). This method differs from the additive method only in terms of how field variables are selected for inclusion in the complexity index.
iii. Direct multivariate statistical modelling of geospatial data against field data has been conducted using techniques such as canonical correlation analysis (CCA) or redundancy analysis (RDA) to derive a complexity index from the resulting relationships

(e.g., Cosmopoulos and King, 2004; de la Cueva, 2008; Pasher and King, 2010, 2011). This method differs conceptually from the other two because the complexity index is defined by the relationships between the field and geospatial data and not a priori from the field data alone.

For all three methods, the resulting complexity index model can then be applied across the study area to map predicted complexity (e.g., Coops et al., 1998; Pasher and King, 2010).

The objective of this study was to evaluate and compare these methods using high-resolution satellite and topographic data. The comparison was based on (*i*) model quality (variance accounted for and validation error) and (*ii*) a qualitative assessment of implementation differences and how well each method represents overall forest complexity. Given that the additive and PCA methods include initial variable reduction steps in a priori definition of the complexity index, the RDA method was expected to provide a more comprehensive index based on more elements and the variance of forest structure and composition. In particular, the PCA method produces one or more components that must be interpreted and each component represents only a portion of the total data variance. Individual components, therefore, were not expected to represent forest complexity as comprehensively as the additive or RDA indices.

Study area

The study area was Chelsea, Quebec, a 111 km^2 rural municipality located on the Canadian Shield, directly north of Gatineau, Quebec. Sixty per cent of the municipality is within the southeastern part of Gatineau Park (**Figure 1**) and was excluded from this study. The remaining area is comprised primarily of forests and wetlands, with some abandoned and working agriculture land. Much of the forested area includes low-density housing or other low-intensity land uses such as selective cutting, whereas agricultural lands in the southern part of the municipality are generally zoned for more intensive community development. The municipality is mostly concerned with conservation management in the forested areas; therefore, this research was designed to help identify areas of high and low forest complexity.

The topography is generally rolling with elevation varying between 100 and 230 m above sea level (Plan Environnemental de la Municipalité de Chelsea, 1991). The region is one of the most biodiverse within the vicinity of Ottawa (Municipality of Chelsea, 2007) and is composed mostly of deciduous forest with smaller areas of coniferous and mixed forest as well as wetlands. Sugar maple (*Acer saccharum* Marsh.) is the dominant deciduous tree species, but others include red maple (*Acer rubrum* L.), ironwood (*Ostrya virginiana* (Mill.) K. Koch), white oak (*Quercus alba* L.), American beech (*Fagus grandifolia* Ehrh.), white ash (*Fraxinus americana* L.), and paper birch (*Betula papyrifera* Marsh.). Coniferous species include eastern white pine (*Pinus strobus* L.), red pine (*Pinus resinosa* Sol.), eastern white cedar (*Thuja occidentalis* L.), white spruce (*Picea glauca* (Moench) Voss), balsam fir (*Abies balsamea* (L.) Mill), and eastern hemlock (*Tsuga canadensis* L.), with the latter being the most abundant (Municipality of Chelsea, 2007).

Field plots and forest measurements

Plot size was selected based on previous geostatistical analysis in similar forests of the adjacent Gatineau Park (Butson and King, 2006) that showed the dominant image pattern scale (from lacunarity analysis) and the semivariogram range of field measured crown size to be about 14 m. To allow for sufficient sample tree numbers in less dense plots, a plot size of 20 m $\times$ 20 m was adopted. Before selecting plot locations, the study area was traversed several times in the late spring of 2009 to develop a strong understanding of the range of forest structure and composition. Three broad forest complexity conditions were visually identified: (*i*) low complexity – homogeneous closed overstory canopy, no to little understory vegetation, and no to little coarse woody debris (cwd); (*ii*) high complexity – high numbers of overstory species, variable tree height and gap size, abundant understory vegetation and dead wood (snags and cwd); and (*iii*) medium complexity – the most common condition, with moderate levels of these attributes. Seventy plots were then established at locations broadly matching the observed distribution of structure and composition. Of these, 15 (21%) were visually assessed as low complexity, 34 (49%) as medium complexity, and 21 (30%) as high complexity. All plots were established within areas that were uniform for the given complexity class over an extent at least two to three times the plot dimensions (40 m $\times$ 40 m to 60 m $\times$ 60 m), the former being required for a few high-complexity plots in localized areas. The plot locations are shown in **Figure 1**. They were purposely spatially grouped to maximize fieldwork efficiency and the range of forest conditions sampled, from large tracts of undisturbed forest to residential areas. This allowed a greater number of plots to be used compared with a random sample that would have required more travel time, placed many plots in less accessible areas, and may have not provided a representative sample of high- and low-complexity conditions, as they were generally smaller in extent than the more common medium-complexity condition. Spatial autocorrelation in either the response or predictor variables was not expected to be an issue as the minimum distance between plot edges (65 m) was much greater than the semivariogram range of structure variables such as crown size, diameter at breast height (DBH), and height measured in a similar nearby forest (Butson and King, 1999), and the semivariogram range of image brightness in spectral bands of resolutions between 50 cm and 3 m (Butson and King,

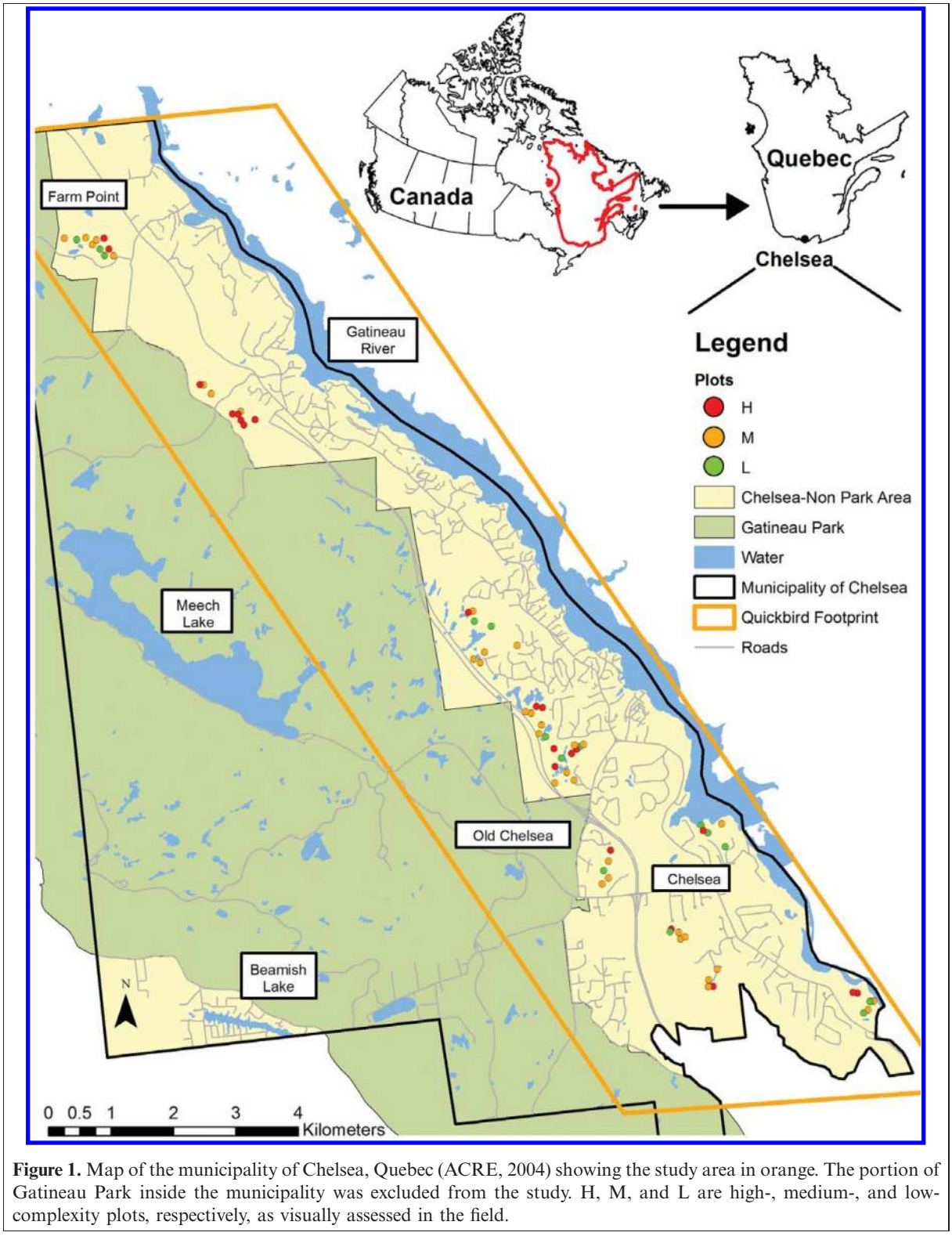

Figure 1. Map of the municipality of Chelsea, Quebec (ACRE, 2004) showing the study area in orange. The portion of Gatineau Park inside the municipality was excluded from the study. H, M, and L are high-, medium-, and low-complexity plots, respectively, as visually assessed in the field.

1999, 2006). **Figure 2** presents example plots for each complexity class. **Figure 3** shows skyward hemispherical photographs taken at 1 and 6 m heights, illustrating greater understory vegetation in the high-complexity plots.

Field variables were selected based on their (*i*) successful use in previous forest modelling of individual attributes (e.g., King et al., 2005) and in multivariate RDA structural complexity modelling (Pasher and King, 2010, 2011) in deciduous forests of the nearby Gatineau Park; (*ii*) importance as habitat indicators in an in-depth study of field-based structural complexity index development (McElhinny et al., 2006); and (*iii*) ease and efficiency of measurement.

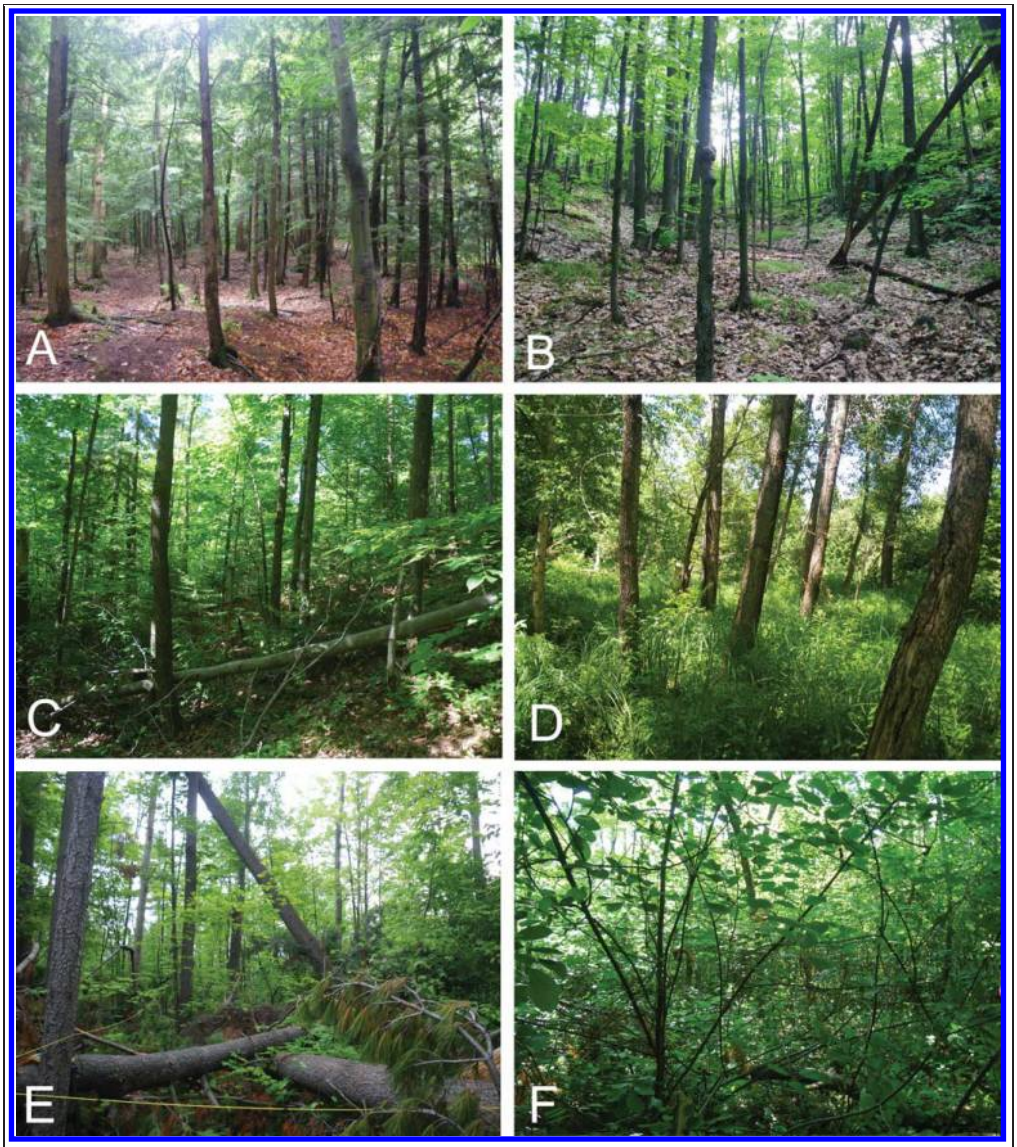

Figure 2. Typical conditions representing visually assessed low structural complexity ((A) mixed deciduous–coniferous and (B) deciduous), medium structural complexity ((C) mixed deciduous–coniferous and (D) deciduous), and high structural complexity ((E) mixed deciduous–coniferous, and (F) deciduous).

They included information on tree size distribution, snags and cwd, ground and understory vegetation structure, canopy openness and leaf area index (LAI), the number of nonwoody understory plant species, and the proportion of coniferous versus deciduous trees. Within the plots, 1837 trees with DBH greater than 10 cm were measured. The number of trees per plot ranged from 8 to 61, with higher numbers in mixed composition areas with abundant hemlock. Generally, coniferous trees were less abundant and their DBH was significantly smaller than deciduous trees; over all the plots, coniferous trees represented 24% of the live basal area. For the four understory vegetation height classes, average cover over all the plots ranged from 11% to 21%. In total, 26 variables were measured or derived from measured variables. Of these, 24 nonredundant variables were retained (**Table 1**) for which between-variable correlations were less than 0.80.

Image and topographic data

High-resolution satellite imagery was selected based on Pasher and King (2010), where the scaling of 20 cm pixel low-coverage airborne imagery to 1 m pixels produced significant RDA models relating image and field variables. An 18 km $\times$ 5 km Quickbird scene was acquired on 15 July 2009 at 1439 local time with sun azimuth and zenith angles of 225.12° and 30.39°, respectively. The data were comprised of four spectral bands (blue, 0.45–0.52 μm; green, 0.52–0.60 μm; red (R), 0.63–0.69 μm; and near-infrared (NIR),

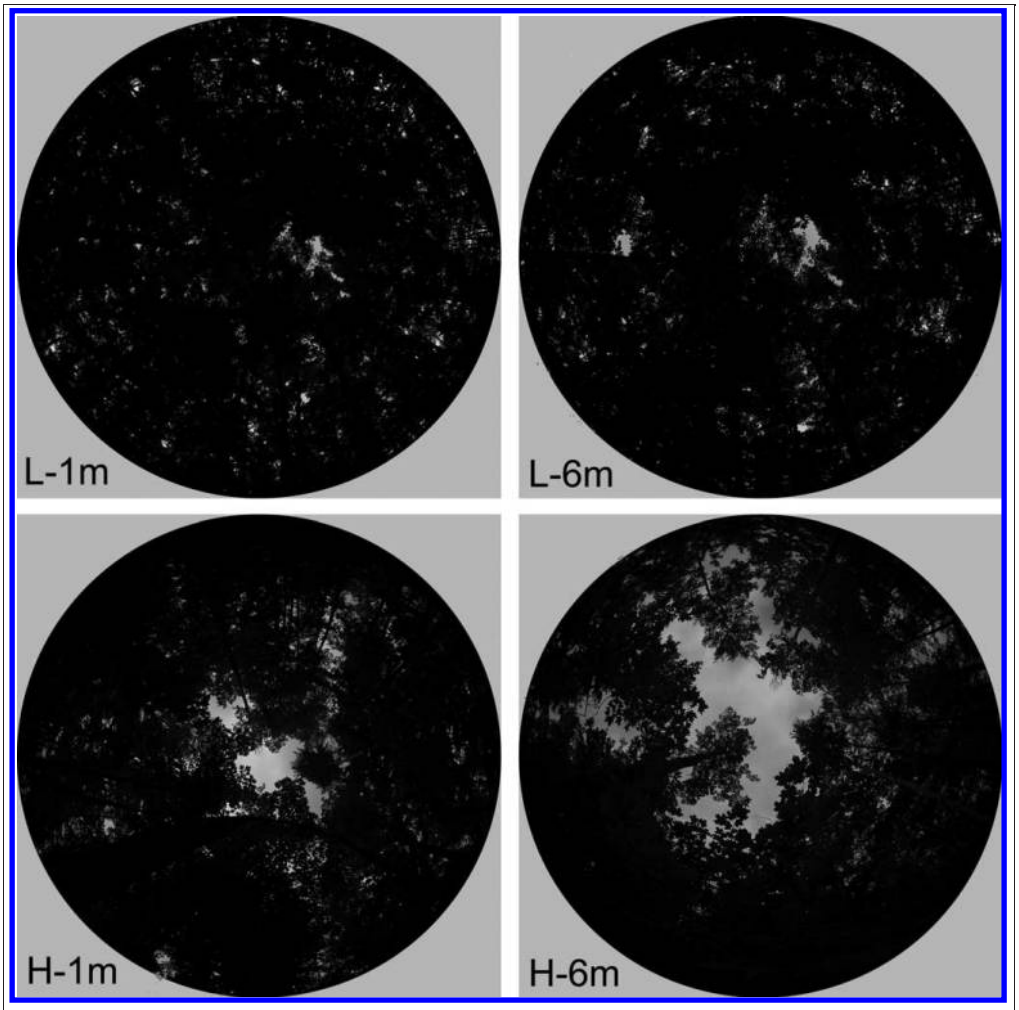

Figure 3. Hemispherical photographs taken at 1 and 6 m heights in plots visually assessed as low complexity (L) and high complexity (H), respectively.

0.76–0.90 μm), each with a nominal ground pixel size of 2.4 m × 2.4 m and a panchromatic image (0.45–0.90 μm) with a 0.6 m × 0.6 m nominal pixel size. All images were projected to the UTM Zone 18T NAD83 datum and georeferenced to existing 25 cm pixel orthophotos using a thin-plate spline model. The root mean square error (RMSE) of the georeferenced image was 3.8 m (x) and 2.7 m (y), which was deemed acceptable for extraction of image information from the 20 m × 20 m plots.

Sixty-eight spectral, texture, and topographic variables were extracted from the multispectral and panchromatic imagery. Of these, 45 nonredundant ($r < 0.80$) variables were retained (**Table 2**). Spectral variables included mean image brightness for all bands, as well as Normalized Difference Vegetation Index (NDVI) (NDVI = [NIR – R]/[NIR + R]). In addition, using Iterative Self Organizing Data Analysis (ISODATA) unsupervised clustering, the panchromatic image was classified into 16 clusters, the first 5 representing shadow, and the others representing vegetation (e.g., Seed and King, 2003; Pasher and King, 2010). From these, shadow and crown fractions were calculated for each plot.

The standard deviation of plot image brightness, a first-order texture measure, was extracted from all bands. Five second-order grey-level co-occurrence matrix (GLCM) texture measures (homogeneity, contrast, entropy, angular second moment (ASM), and correlation (Haralick et al., 1979)), selected based on previous studies (Franklin et al., 2000; King et al., 2005; Wunderle and Franklin, 2007; de la Cueva, 2008; and Pasher and King, 2010), were extracted for each plot from the NIR and panchromatic bands. The NIR band was used, as it displayed the most distinct texture among the spectral bands (as in Coops and Catling, 1997; Estes et al., 2010); the panchromatic band was used because of its higher spatial resolution and more detailed texture. Multiple window sizes were tested and the results were visually assessed for differentiation of high- and low-complexity plots. From these tests, moving windows of 3 × 3 (7.2 m × 7.2 m) and 9 × 9 (21.6 m × 21.6 m; i.e., one window covered the whole plot) pixels were applied to the NIR band and a 5 × 5 (3 m × 3 m) moving window was applied to the panchromatic image for texture metric calculation using omnidirectional pixel sampling within the window.

Average elevation, slope, and aspect were extracted for each plot from a 10 m cell digital elevation model (NCC, 2002). Northness (sine of aspect) and eastness (cosine of aspect) values were calculated from the aspect map (Guisan et al., 1999; Lassueur et al., 2006). A solar illumination image was also created, using the sun zenith and azimuth

Table 1. A summary of the 24 field variables, including the subset of 10 core attributes used in the additive index.

	Field variables	Additive core attributes	Plot average	+/− std
Overstory tree	Stem density	X	26.24	10.68
	No. of large trees (DBH > 35 cm)	X	2.47	2.08
	x tree DBH (cm)		21.90	4.46
	σ tree DBH (cm)		9.90	3.71
	σ tree height (m)	X	3.61	1.09
	x nearest neighbour distance (m)		1.86	0.55
	Live basal area (m^2/ha)	X	29.02	12.36
	Live coniferous basal area (m^2/ha)		23.77	29.18
	No. of tree species	X	4.64	1.99
Deadwood	No. of snags		4.47	3.30
	No. of of pieces of cwd	X	14.29	9.71
	xsnag DBH (cm)		19.52	9.44
	Dead basal area (m^2/ha)[1]	X	5.30	6.55
Ground and understory vegetation	Amount of rock (%)[2]		6.58	10.12
	Vegetation cover (0–10 cm) (%)	X	10.94	7.26
	Vegetation cover (10–50 cm) (%)[2]		14.89	10.77
	Vegetation cover (50 cm–1 m) (%)[1]	X	10.56	13.62
	Vegetation cover (1–2 m) (%)		21.03	15.75
	Vegetation cover (<2 m) (%)		57.63	25.20
	No. of different plant species	X	4.93	1.20
Canopy (from photographs)	1 m openness (0°–60°) (%)[1]		4.23	3.91
	6 m openness (0°–60°)[1]		12.15	12.31
	1 m LAI (57.5°)		7.10	2.30
	6 m LAI (57.5°)		8.81	2.57

Note: Non-normal variables that required logarithmic[1] or square root[2] transformations.

Table 2. A summary of the 45 image and topographic variables.

Variable type	Variable extracted	Spectral bands	Texture	Window size
Spectral	$\bar{x}$	Brightness: NIR, R, G, B, pan, and NDVI		Whole plot: 9 × 9 pixels
	% of plot classified as shadow[1] and crown			Whole plot: 9 × 9 pixels
Spatial	σ	NIR, R, G, B, pan and NDVI	Standard deviation as a first-order texture	Whole plot: 9 × 9 pixels
	$\bar{x}$	NIR	GLCM contrast, homogeneity*, ASM*, correlation*† and entropy	3 × 3[†] moving window; whole plot: 9 × 9 pixels[‡]
	σ	NIR	GLCM contrast, homogeneity, ASM*, correlation and entropy	3 × 3 moving window
	$\bar{x}$	Panchromatic	GLCM contrast, homogeneity*, ASM*, correlation and entropy	5 × 5 moving window
	σ	Panchromatic	GLCM contrast, homogeneity, ASM*, correlation and entropy	5 × 5 moving window
Topographic	Elevation, slope*, aspect Eastness, northness Solar illumination	DEM, 10 m pixels		

Note: Non-normal variables are identified that required logarithmic* transformation. Some variables were derived from a single 9 × 9 window representing the whole plot, while others represent plot means of values derived from 3 × 3 or 5 × 5 moving windows.

[†]Transformation of mean NIR Correlation was applied to the 3 × 3 window data only.

[‡]Single plot texture value; i.e., not a mean of multiple values.

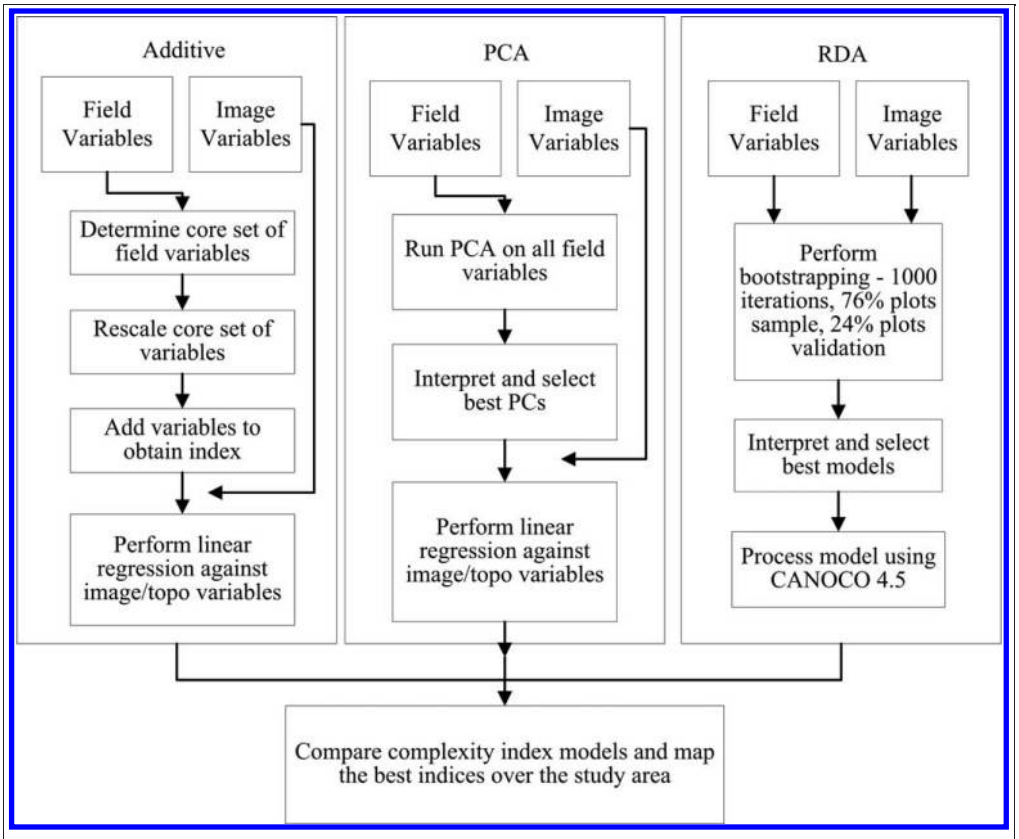

Figure 4. Process map showing the creation, selection, mapping, and validation of three complexity indices, including two based on a priori definition of a field-based index (additive and PCA methods) and an index derived from multivariate statistical modelling of geospatial and field data (RDA method).

corresponding to the date and time of image acquisition as a potential discriminator of forest composition and structure complexity. Among these, elevation was found to be the most significant in previous RDA studies (de la Cueva, 2008; Pasher and King, 2010). The image brightness variations owing to topography were not large enough to be able to conduct precise topographic brightness correction. Instead, it was assumed that the topographic variables above would contribute significantly to complexity index models if topography was indeed related to vegetation structure and composition.

Derivation of forest complexity indices

All variables were checked for normality; six field and nine image variables required logarithmic or square root transformations as shown in **Tables 1** and **2**, respectively. The forest complexity indices were then derived as described below and as shown in **Figure 4**.

Additive field-based complexity index created using manually selected variables

A simple additive index was created from the field variables using a process developed by McElhinny et al. (2006) for structural complexity and later adapted by Estes et al. (2010) to include composition. A set of 10 core attributes (**Table 1**) was selected where no variable pair was correlated with $r > 0.60$ (i.e., as in McElhinney et al. (2006) and more strict than the initial threshold ($r = 0.80$) of this study). For each pair of correlated variables, the variable that was more difficult to measure in the field was removed. These attributes represented the horizontal and vertical structure of the tree canopy, as well as understory vegetation, dead wood, and composition. Unlike the variables in McElhinny et al. (2006), which all increased with increasing complexity, live basal area and stem density decreased with increasing structural complexity so they were multiplied by –1. The 10 variables were then standardized (z-scores), rescaled to 0–10 using a simple linear function and added to create the index. The actual index values ranged between 0 and 47. Stepwise multiple regression of the index against the geospatial variables was conducted (probability of F-to-Enter = 0.05; F-to-Exit = 0.10), withholding 13 randomly selected plots for use in model validation. This process was manually repeated five times to evaluate the validation error variability owing to sample differences and each time predicted complexity to actual complexity, calculated directly from the field measurements, was compared.

Field-based complexity indices created using PCA

Standardized PCA was performed on the field variables (e.g., Wunderle et al., 2007). Individual PCs representing interpretable structure or composition gradients were then modelled against the geospatial data using stepwise regression as described previously. As interpretable PCs were expected to represent different gradients, a linear addition of PCs weighted by their respective proportion of explained variance was also modelled against the geospatial data. In all regressions for the additive and PCA indices, linearity between the predictor and response variables was verified, a variance inflator factor (VIF) of 5.3 was used as a cutoff threshold (Hair et al., 1998) to avoid multicollinearity, and normality and uniformity of the variance of the residuals were checked to ensure that these assumptions were not violated.

Complexity index created through RDA multivariate modelling of field and geospatial data

RDA was implemented to derive a complexity index based on information within the geospatial data. RDA was originally developed to examine effects of multiple environmental factors on several species simultaneously, and it is one of many multivariate techniques known as constrained ordination that order or arrange high-dimensionality data along one or more gradient axes (Pielou, 1984). In a simplified manner, RDA is carried out by calculating multiple linear regressions for each response variable (y) individually (in this case the field variables, which were standardized within the algorithm by dividing each by its standard deviation), followed by a single PCA of the fitted/predicted response variables (y'). The resultant components, which are uncorrelated and represent decreasing portions of the variance in the response variables, can then be interpreted in terms of the relationships between the response and predictor variables (in this case the geospatial variables) (Legendre and Legendre, 1998; Pasher, 2009). RDA was selected instead of CCA, which was used in previous research (Jakubauskas, 1996; Olthof and King, 2000; Cosmopoulos and King, 2004), because it is more suited to modelling the variance in a set of response variables explained by a set of predictor variables and because for large numbers of field and geospatial variables, CCA requires initial variable reduction using techniques such as PCA to avoid model overfitting. In doing this, data variance is lost as not all components are selected for input, and the output is an index comprised of field principal components modelled by image principal components. This assumes a straightforward interpretation is possible for all components and the resulting canonical variates, which is not always possible. In RDA, the original variables can be input, thus retaining all data variances and producing an output index that is comprised of the original field variables modelled by a subset of the original image variables. This renders the index model straightforward to interpret. RDA was implemented in Matlab using existing code from Pasher and King (2011). It included bootstrapping in two stages, one for variable selection and one for model validation. During the first stage, 1000 RDA iterations were run using forward stepwise selection, where 53 (76%) of the 70 plots were randomly selected in each iteration. This created 1000 potentially different models that were independently analyzed. In the second bootstrapping stage, the 1000 models were run through the RDA process but without using stepwise selection (i.e., entering only the variables that had been selected in the first stepwise regression stage of model creation). Again, 53 of the 70 plots were randomly selected for each iteration, but this time these plots were used for model validation. Summary statistics calculated for all models included: (*i*) model parameters (significance, adjusted R^2 (R^2_{adj})), and average Akaike's information criterion (AIC) (Akaike, 1974); (*ii*) the number of times variables were entered into models, which was evaluated to assess the stability and consistency of models; and (*iii*) validation parameters, including *r*-validate, which is the correlation between the predicted and observed index values, and the root mean square prediction error as a percentage of the mean of the derived index (% RMSE). In addition, models were evaluated subjectively based on their simplicity (fewer variables being considered better than more variables), ease of calculation of the geospatial variables entered into the model, and ease of interpretation of the resulting complexity index. The RDA model deemed to be best was processed with CANOCO 4.5 (ter Braak and Šmilauer, 2002), to generate triplots that were used to evaluate the relations between field and geospatial data and to aid in the interpretation of the RDA axes.

Mapping predicted forest complexity

The image and topographic variables included in the best models were then multiplied by their corresponding model coefficients for all pixels in the study area to produce maps of predicted complexity. These maps were compared visually to evaluate their differences and how well they represented the known spatial distribution of forest complexity across the study area.

Results

Additive field-based complexity index created using manually selected variables

A summary of the geospatial variables entered into the five independent stepwise regression models for the 10 variable additive index is given in **Table 3**. Standard deviation of NDVI (a first-order texture) was the first variable entered in all models and it accounted for most of the model variance.

Table 3. Results of complexity index modelling using the additive, principal component analysis, and redundancy analysis methods.

Index	Predictor variables	R^2	Adjusted R^2	Variable contribution to R^2		p
Additive index		0.53	0.49			
	STD plot NDVI			0.32		0.000
	Eastness			0.07		0.005
	5 × 5 pan STD homogeneity			0.05		0.015
	3 × 3 NIR entropy			0.04		0.034
	Mean slope			0.05		0.014
PC2		0.41	0.38			
	STD plot NDVI			0.31		0.000
	5 × 5 pan ASM			0.06		0.008
	Eastness			0.04		0.025
PC3		0.54	0.52			
	(−) Mean plot NDVI			0.41		0.000
	(−) 5 × 5 pan ASM			0.08		0.003
	(−) Northness			0.05		0.009
Weighted sum of PC 1–8		0.36	0.33			
	STD plot NDVI			0.28		0.000
	(−) Northness			0.05		0.019
	Eastness			0.03		0.031
		0.34	0.32	RDA1	RDA 2	
RDA index	Mean plot NDVI			−0.88	−0.06	≤ 0.050
	Plot shadow fraction			0.62	−0.01	
	3 × 3 NIR ASM			0.14	−0.04	
	9 × 9 NIR correlation			0.28	0.02	
	5 × 5 pan STD homogeneity			0.12	−0.28	
	Mean slope			−0.11	0.08	

Note: For the RDA index, the correlation (r) of each geospatial variable with each RDA axis is given. (−) indicates predictor variable decreased with increasing index value. STD = standard deviation over the whole plot; otherwise mean plot values are given.

Minor variables, each accounting for less than 8% of the variance, included GLCM texture and texture variation and two topographic variables. All regression models had a VIF at each model step significantly lower than the cutoff of 5.3; the maximum correlation between model predictor variable pairs was 0.35. In validation, the mean index error for the five regressions (each index error being the average error for the 13 plots randomly excluded from the respective models) was −0.05. This is very close to 0.0 and indicates negligible bias across multiple model derivations using different random subsets of the data. Four of the five regressions had errors between −0.5 and 3.6, while the error for the fifth was −6.3. The mean absolute error for the five regressions was 6.7 or 24.3% of the mean value of the index. The RMSE ranged between 22.4% and 50.0% of the mean index value, with an average of 33.1%.

Field-based complexity indices created using PCA

PCA of the 24 field variables resulted in eight PCs with eigenvalues greater than 1.0, a common metric among several in the literature used to evaluate significance of PCs (Kaiser, 1960). This illustrates the high dimensionality of the structure and composition of the field data. PC 9 accounted for only 3.5% of the variance, therefore, it and the subsequent PCs were discarded. To determine the suitability of PCs 1–8 as complexity indices, either on their own or in an integrated combination, efforts were made to interpret each based on factor loadings. PCs 1–4 were the most easily interpreted and together accounted for 55.1% of the variance. PC1 increased with increasing canopy openness and understory cover as well as decreasing LAI, stem density, and basal area. PC2 increased with increasing tree size and amount of dead wood. PC3 increased with increasing coniferous basal area, increasing numbers of tree species, and decreasing tree spacing. PC4 increased with decreasing canopy openness and understory above 1 m, and it was almost the inverse gradient of PC1. PCs 5–8 were difficult to interpret because of factor loadings that were generally lower and more evenly distributed.

Stepwise regression of each PC against the geospatial variables yielded eight significant models ($p < 0.05$), all with VIF at each step below the selected threshold; the maximum correlation between model predictor variable pairs was 0.36. The PC2 and PC3 models had the highest R^2 values and were stronger models than PCs 1–8 together (**Table 3**), demonstrating that the geospatial data are more related to some aspects of forest structure and composition than others. These models were not validated nor applied in mapping because they represented specific gradients that

Table 4. Percentage of the 1000 RDA models in which each geospatial variable type was entered.

Variable class	Frequency of variable class in models (%)	Variable	Frequency of variable in models (%)
Spectral	98	Mean NDVI	94
Topographic	27	Log mean slope	24
First-order spatial (texture)	27	Std dev. blue	22
Second-order spatial (texture)	23	Mean 3 × 3 NIR ASM	10
Radiometric fractions	16	Shadow fraction	23

accounted for only a small portion of the variance of the field data set (see Discussion).

Complexity index created through RDA multivariate modelling of field and geospatial data

Model selection

The 1000 models produced during the first stage of RDA bootstrapping were all highly significant ($p < 0.001$) and were comprised of different combinations of image variables; **Table 4** shows the frequency with which variable classes and the most commonly entered variable in each class were entered into models. The R^2_{adj} values ranged between 0.19 and 0.37 (mean = 0.23) and models contained between 1 and 12 variables (**Table 5**). From bootstrapping validation, RMSE was fairly high, ranging between 40.9% and 54.6% of the mean index value, whereas AIC ranged between 546.1 and 556.1.

In **Table 5**, it is evident that the model with the most variables (12) provides the best fit, but its capability to predict complexity using new data is lower than that for models with fewer variables; i.e., this model is "over fit". Using a 10:1 observation-to-variable rule (Babyak, 2004), only models with seven or fewer variables were evaluated. The best was a six-variable model with $R^2_{adj} = 0.32$ for the first two RDA axes ($R^2_{adj} = 0.23$ and 0.09; $p = 0.001$ and 0.009, respectively, for RDA1 and RDA2). Of all predictor variable pairs, NDVI and shadow fraction were moderately highly correlated ($r = 0.77$), indicating potential multicollinearity effects, whereas all other correlations were less than 0.35. The model had a mean error of 0.03, a mean absolute error of 5.2 or 22.3% of the mean index value, and RMSE ranged between 17.5% and 45.9% of the mean index value with an average of 31.7%.

Table 5. Summary statistics for the 1000 RDA models provided through stage two bootstrapping.

Model	Adjusted R^2	R validate	No. of variables	% RMSE	AIC
Minimum	0.19	0.61	1	40.9	546.1
Maximum	0.37	0.78	12	54.6	556.1
Average	0.23	0.75	2.283	48.1	548.3

Examination of the sign of the significant bivariate correlations (**Table 3**) and the triplot of **Figure 5** for RDA1 (x axis) and RDA2 (y axis) revealed how each field and geospatial variable contributed to the gradient of complexity. In **Figure 5**, red arrows represent the six geospatial variables and blue arrows represent the field variables. Correlation between any two variables or between a variable and an RDA axis increases as the angle between them decreases. Arrow length indicates the strength of variable contribution to the model. The field plots are also displayed in **Figure 5**, which allows them to be assessed in terms of their structure and composition.

Plots that are highly positive on RDA1 have a considerably greater coniferous basal area than plots on the negative side, which are comprised of almost all deciduous trees. This increasing coniferous component is manifested in the imagery by increasing shadow fraction and GLCM correlation texture as well as decreasing NDVI (the strongest of the image variables). These are expected relations owing to the generally lower reflectance of coniferous trees and their larger shadows (produced by their conical crown morphology) compared with deciduous trees in such forests with canopy openness typically less than 40%. Plots that are more positive on RDA2 have more trees (higher stem density), higher basal area, more uniform tree sizes (lower standard deviation DBH), less understory, and fewer plant species. Based on this, increasing RDA2 indicates reduced complexity. In the geospatial data, increasing complexity (more negative RDA2) is associated with greater image texture (decreasing ASM, a GLCM texture for which low values indicate heterogeneity), greater variation in GLCM homogeneity texture, and steeper slopes.

The most complex plots from both a structural and compositional perspective are located in the lower right quadrant of **Figure 5**. They are generally more open with more tree species (including more coniferous trees), canopy height variation, standing and fallen deadwood, and understory than plots in the upper left quadrant, which are the least complex. As an example, plots 62 and 65 (visually classified in the field as high complexity) are found in the lower right quadrant and have greater shadow fraction, texture, and texture variation in the imagery than simpler plots such as 27 and 37 (visually classified in the field as low complexity) in the upper left quadrant. Moderate-complexity plots vary between the upper right and lower left quadrants. Plot 29, an outlier in the top right of the triplot, was a coniferous plantation with a small amount of deciduous understory that was characterized by much greater stem density, live basal area, and much lower variation in DBH, numbers of plant species, understory, and deadwood than all other plots. It had very uniform image texture (the highest ASM texture value), the highest proportion of shadow and low NDVI. The differences between the RDA axes scores for plot 29 and the scores

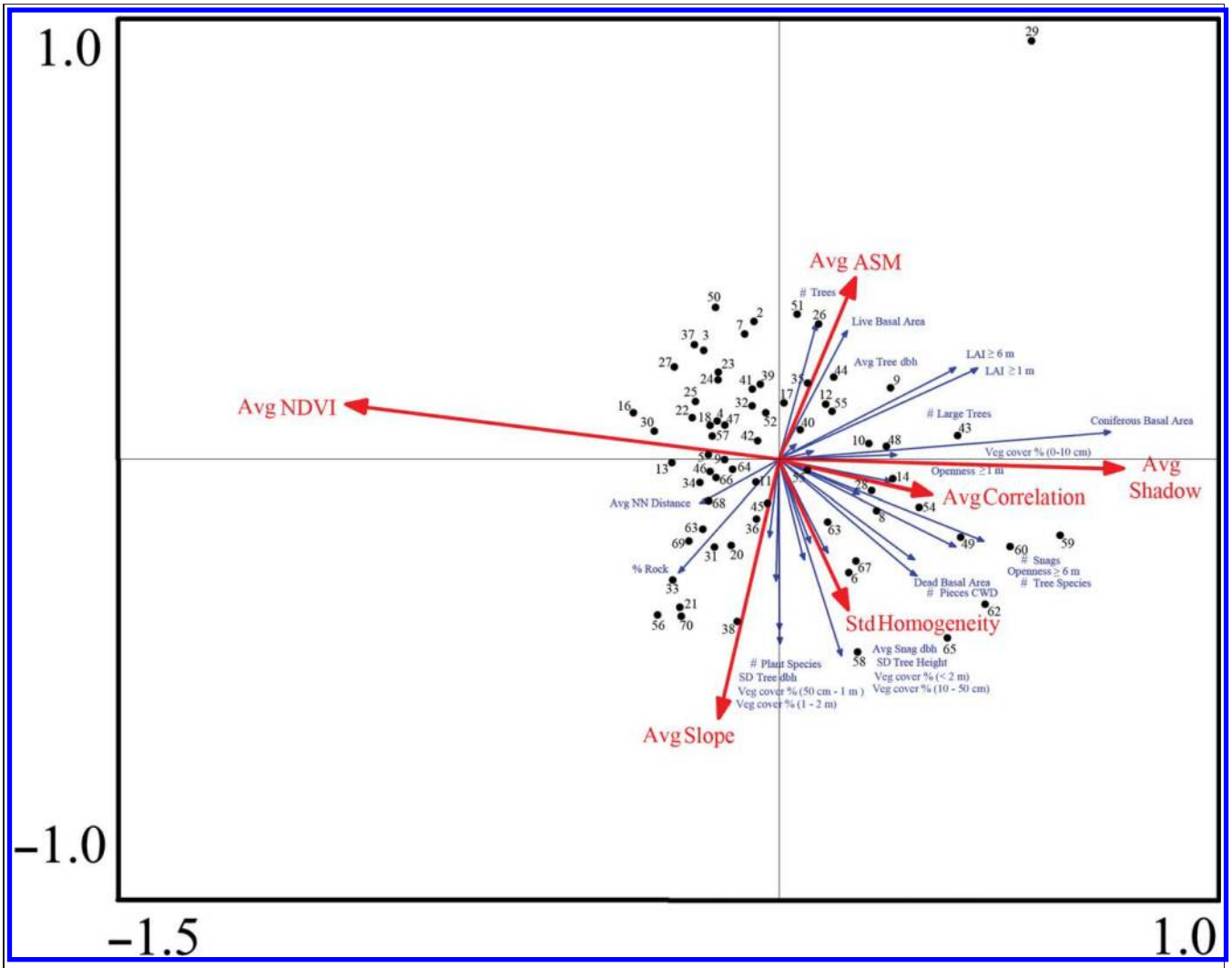

Figure 5. The triplot for the first and second RDA axes (*x* and *y* axes, respectively), showing a gradient of increasing forest structure– composition complexity from left to right for RDA1 and from top to bottom for RDA2. Axis values represent the ordination complexity scores. Red arrows are geospatial data variables, and blue arrows are field variables. Numbered black dots are field plot numbers.

calculated (predicted) from the geospatial variables in the index were 0.09 and –0.09 for RDA1 and RDA2, respectively. This illustrates that the RDA method correctly assigned this extremely simple structure–composition condition within the RDA1 and RDA2 gradients as interpreted above, and that the method and model could be useful for a wider range of conditions than the typical conditions of this study.

Complexity index mapping

The additive and RDA indices were mapped across Chelsea (**Figure 6**). For the RDA map, an additive combination of RDA1 and RDA2 and weighted by their respective proportion of explained variance, was produced as described earlier for the PC1–PC8 index. The additive and RDA maps were then normalized to the same attribute scale for comparison. Overall, both maps display similar patterns of complexity across the study area and at full resolution as shown in the insets of **Figure 6**. However, where predicted complexity is either higher (yellow-red) or lower (cyan-blue) than the more common moderate complexity (green), the RDA index map generally shows a larger proportional area with the given complexity level. For example, in the middle left of the maps the RDA index predicts a greater proportion of more complex forest than the additive index. Most of the study area forest was predicted to have complexity levels that were between the low and medium classes of the subjectively defined range used during field plot establishment. At that time, when the area was being traversed intensively, it was observed that high- and low-complexity areas occurred less frequently (i.e., they represented the upper and lower tails of the complexity distribution) and were smaller in extent than the more common conditions of medium complexity.

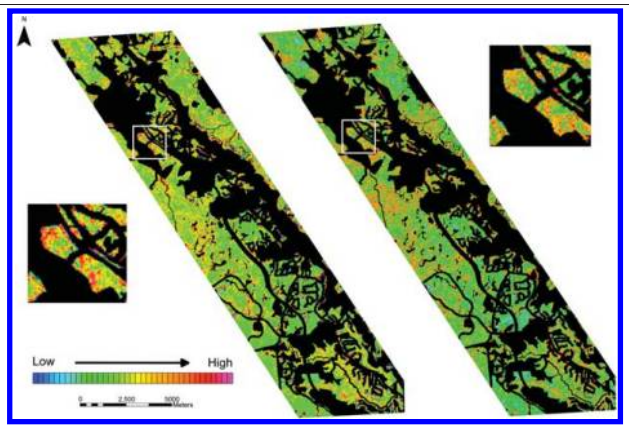

Figure 6. Map comparison with additive index on the left and RDA index of the right. Insets display similar areas of high complexity. The black areas in this image represent cloud shadow, roads, housing, and other nonforested areas.

Discussion

Comparison of the forest complexity index methods

Comparison of the three forest complexity index methods based on model quality parameters such as R^2_{adj} or validation error must be conducted with the understanding that such indices may be based on different field variables with different total variance. To illustrate, **Table 3** shows R^2_{adj} of the complexity models to be, from best to worst: PC3, additive index, PC2, PC1–PC8, RDA index. However, each PC represented only a small portion of the variance of the original 24 field variables (e.g., 14.9% and 10.8% for PC2 and PC3, respectively), and each PC also represented a specific environmental gradient. Thus, the PC2 and PC3 models in **Table 3** do not represent forest structure and composition complexity as comprehensively or with as much integration as the RDA or additive models. Even combining the first eight PCs produced a model R^2_{adj} of 0.33 while these PCs cumulatively only accounted for 75.72% of the field data variance. In a similar way, the additive index model was based on 10 "core" field variables that were manually selected to represent different aspects of forest complexity using the $r < 0.60$ criterion suggested by McElhinny et al. (2006). Although the R^2_{adj} of the model was 0.49, the 10 field variables in the model accounted for only 38.5% of the total variance of the original 24 variable set (based on a PCA analysis of each data set). Thus, the additive model is also not as comprehensive a representation of forest structure and composition complexity as the RDA model. This logic also applies to the validation errors for the additive and RDA models. The percentage errors in relation to the respective mean index values appear to be similar in magnitude, but the additive index is comprised of many fewer response variables and lower data variances than the RDA index.

The variables selected were similar for all methods. Mean plot NDVI or standard deviation of NDVI, as well as topographic variables, were entered into all five models of **Table 3**. GLCM textures were entered into four of the five models with ASM entered into three of the five. This indicates that the significant predictor variables are quite consistent when modelling a complexity index using a large data set such as the 24 field variables of this study, the reduced 10 core field variables, or the further reduced set of PCs. In implementation, the additive index and PCA-based methods are simpler and they produced models with fewer predictor variables than the best RDA model. Given that field measurement may be constrained by time or funding, or given a desire to produce a simpler model for mapping that may be more transportable to other areas, the additive index approach using fewer variables may be warranted. However, it is partly subjective to manually determine the most important core variables to include in an additive index, particularly when the goal is to model them with a variety of geospatial data types. It also does not provide as in-depth information on relationships between ecological and image data gradients or on relationships between variables and plots as does RDA.

In summary, the additive and RDA methods were deemed to be better overall than the PCA method but a decision of which to use should be based on the trade off between model simplicity and information content.

Comparison with previous studies

In comparison with previous studies, the same difficulty described previously arises; most studies report R^2 and sometimes prediction error, but studies vary in the number, type, and variance of the field variables used in modelling. Thus, the following is a brief comparison of the results of this study with others, but comparison of absolute model quality attributes should be avoided.

PCA has been used for a long time for dimensionality reduction and to analyze the vegetation structure and composition gradients (e.g., Cohen and Spies, 1992; Danson and Curran, 1993; Hansen et al., 2001). More recently, Wunderle et al. (2007) found PC1 to be the only significant component, representing 79% of the total variance in nine structural variables for classes of recently harvested, regenerating, and old growth forests. It was modelled using three spectral and textural variables with an R^2 of 0.74. This performance, that is seemingly superior to the results of the current study in which eight significant PCs were extracted, each modelled with much poorer fit (**Table 3**), may be explained by the greater distinction between their forest structure classes.

Estes et al. (2010) used PCA of 13 vegetation structure habitat features important to the mountain bongo antelope in Kenya. Geospatial data regressed against the PCs were derived from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) optical imagery, a shuttle

radar topography mission digital elevation model, and rainfall. PC1, which accounted for 28% of the field data variance, produced the best model ($R^2 = 0.51$). This is a similar general result to the current study; if such a model accounts for 51% of the variance in a PC that itself accounts for only a small portion of the total original field data variance (28%), then the model does not represent well the overall structural–compositional complexity, but may be useful for analysis of the specific gradient associated with the PC. Estes et al. (2010) also modelled an additive index adapted from McElhinny et al. (2006) using a subset of eight core variables, resulting in an R^2 of 0.46. The R^2 of 0.51 ($R^2_{adj} = 0.49$) achieved in this study with 10 core variables is similar, although the results of the two studies are dependent on the variance of their original data sets as discussed previously.

Similar to the studies of Pasher and King (2010, 2011), this study used spectral, spatial, and topographic data to derive an RDA-based forest complexity index. In Pasher and King, only the first RDA axis was significant, explaining 35% of the field data variance. Due primarily to the integration of composition with structure in the current study, two significant RDA axes were produced, providing a more complete representation of the diversity of forest conditions. In both studies, topographic, first- and second-order texture, and shadow fraction were statistically significant variables in the complexity indices. However, plot-level NDVI and NDVI variance were more significant in this study because of the larger pixel size and the reduced spatial information compared with their study, which produced a model including more within-crown spatial metrics. In contrast, the de la Cueva (2008) RDA study incorporated much coarser resolution Landsat spectral and DEM information to produce a model combining the first and second RDA axes that explained 24% of the field data variance. These three studies of varying resolution demonstrate the advantage of spatial information in higher resolution data.

Conclusion

Three methods for creating and mapping a multivariate forest structure and composition complexity index were compared using Quickbird imagery and topographic data. The results showed that an index derived through multivariate RDA analysis provided a more objective and complete representation of structure–composition complexity than defining an index a priori using a reduced set of variables (the additive index and PCA methods). RDA provides a very rich analysis of relations among and between complexity gradients, the field variables, and the geospatial variables. However, the additive index is easier to implement and understand. Both the RDA and additive indices identified the same locations of the study area that had been rated visually in the field as high or low complexity.

These empirical approaches to defining, modelling, and mapping forest complexity should be applicable in a wide variety of mixed wood forests. However, as the concept of complexity is relative to an ecoregion or biome, complexity in one region may be represented by some different variables than in another region. Future research will evaluate the spatial extent over which the models produced in this study can be effectively applied (e.g., to forests in an adjacent municipality or in eastern Ontario, about 100 km from this study area). In addition, while there is some literature on the relationships of forest complexity with biodiversity, both theoretically and empirically (as given in the Introduction), in collaboration with the Municipality of Chelsea research is being initiated to acquire biodiversity data for the study area to determine if the areas of mapped high and low complexity exhibit high and low biodiversity, respectively. These methods and results will then feed into conservation planning and management actions within the municipality.

Acknowledgements

This research was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant to D. King. Field support was provided by the Carleton University Geomatics and Landscape Ecology Research Laboratory through the Canada Foundation for Innovation and other funding. Carolyn Callaghan of Nature Chelsea provided access to existing Chelsea geospatial data. The field assistance of Chris Czerwinski and advice of Dr. Jon Pasher are greatly appreciated. Dr. Raphael Proulx of the Université du Québec à Trois Rivières provided the Matlab RDA code and RDA implementation advice.

References

Adams, J., Kapos, V., Smith, M.O., Almeida Filho, R., Gillespie, A.R., and Roberts, D.A. 1990. A new Landsat view of land use in Amazonia. *International Archives of Photogrammetry and Remote Sensing*, Vol. 28, pp. 177–185. doi: 10.1109/TAC.1974.1100705.

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, Vol. 19, No. 6, pp. 716–723.

Babyak, M.A. 2004. What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine*, Vol. 66, No. 3, pp. 411–421. doi: 10.1097/01.psy.0000127692.23278.a9.

Butson, C.R. and King, D.J. 1999. Semivariance analysis of forest structure and remote sensing data to determine an optimal sample plot size. *Proceedings 4th International Airborne Remote Sensing Conference and Exhibition / 21st Canadian Symposium on Remote Sensing*, 21–24 June 1999, Ottawa, Ontario. Vol. II, pp. 155–162.

Butson, C.R. and King, D.J. 2006. Lacunarity analysis to determine optimum extents for sample-based spatial information extraction from high-resolution forest imagery. *International Journal of Remote Sensing*, Vol. 27, No. 1, pp. 105–120. doi: 10.1080/01431160500238844.

Cohen, W.B. and Spies, T.A. 1992. Estimating structural attributes of Douglas-fir/Western hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment*, Vol. 41, No. 1, pp. 1 17. doi: 10.1016/0034-4257(92)90056-P.

Coops, N. and Catling, P. 1997. Predicting the complexity of habitat in forests from airborne videography for wildlife management. *International Journal of Remote Sensing*, Vol. 18, No. 12, pp. 2677–2682. doi: 10.1080/014311697217530.

Coops, N.C., Culvenor, D., Preston, R., and Catling, P.C. 1998. Procedures for predicting habitat and structural attributes in eucalypt forests using high spatial resolution remotely sensed imagery. *Australian Forestry*, Vol. 61, No. 4, pp. 244–252.

Cosmopoulos, P. and King, D.J. 2004. Temporal analysis of forest structural condition at an acid mine site using multispectral digital camera imagery. *International Journal of Remote Sensing*, Vol. 25, No. 12, pp. 2259–2275. doi:10.1080/0143116032000160507.

Danchuk, S. and Willson, C.S. 2010. Effects of shoreline sensitivity on oil spill trajectory modelling of the Lower Mississippi River. *Environmental Science and Pollution Research*, Vol. 17, No. 2, pp. 331–340. doi: 10.1007/s11356-009-0159-8.

Danson, F.M. and Curran, P.J. 1993. Factors affecting the remotely sensed response of coniferous forest plantations. *Remote Sensing of Environment*, Vol. 43, No. 1, pp. 55 65. doi: 10.1016/0034-4257(93)90064-5.

de la Cueva, A.V. 2008. Structural attributes of three forest types in central Spain and Landsat ETM plus information evaluated with redundancy analysis. *International Journal of Remote Sensing*, Vol. 29, No. 19, pp. 5657–5676. doi: 10.1080/01431160801891853.

Estes, L.D., Reillo, P.R., Mwangi, A.G., Okin, G.S., and Shugart, H.H. 2010. Remote sensing of structural complexity indices for habitat and species distribution modelling. *Remote Sensing of Environment*, Vol. 114, No. 4, pp. 792–804.

Fahrig, L. 2003. Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution and Systematics*, Vol. 34, No. 1, pp. 487–515. doi: 10.1146/annurev.ecolsys.34.011802.132419.

Franklin, J. and Strahler, A.H. 1988. Invertible canopy reflectance modelling of vegetation structure in semiarid woodland. *IEEE Transactions of Geoscience and Remote Sensing*, Vol. 26, No. 6, pp. 809–825. doi: 10.1109/36.7712.

Franklin, S.E., Hall, R.J., Moskal, L.M., Maudie, A.J., and Lavigne, M.B. 2000. Incorporating texture into classication of forest species composition from airborne multispectral images. *International Journal of Remote Sensing*, Vol. 21, No. 1, pp. 61–79. doi: 10.1080/014311600210993.

Gaston, K.J., Blackburn, T.M., and Goldewijk, K.K. 2003. Habitat conservation ad global avian biodiversity loss. *Proceedings of the Royal Society B: Biological Sciences*, Vol. 270, No. 1521, pp. 1293–1300. doi: 10.1098/rspb.2002.2303.

Guisan, A., Weiss, S.B., and Weiss. A.D. 1999. GLM versus CCA spatial modelling of plant species distribution. *Plant Ecology*, Vol. 143, No. 1, pp. 107–122. doi: 10.1023/A:1009841519580.

Hair, J.F., Anderson, R.E., Tatham, R.L., and Black, W.C. 1998. *Multivariate Data Analysis* (Fourth edition). Prentice Hall, Upper Saddle River, NJ.

Hansen, M.J., Franklin, S.E., Woudsma, C., and Peterson, M. 2001. Forest structure classification in the North Columbia mountains using the Landsat TM Tasseled Cap wetness component. *Canadian Journal of Remote Sensing*, Vol. 27, No. 1, pp. 20–32.

Haralick, R.M. 1979. Statistical and structural approaches to texture. *Proceedings to the IEEE*, Vol. 67, No. 5, pp. 786–804. doi: 10.1109/PROC.1979.11328.

Jakubauskas, M.E. 1996. Canonical correlation analysis of coniferous forest spectral and biotic relations. *International Journal of Remote Sensing*, Vol. 17, No. 12, pp. 2323–2332. doi: 10.1080/01431169608948775.

Kaiser, H.F. 1960. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, Vol. 20, No. 1, pp. 141–151. doi: 10.1177/001316446002000116.

King, D.J., Olthof, I., Pellikka, P.K.E., Seed, E.D., and Butson, C. 2005. Modelling and mapping forest ice storm damage using remote sensing and environmental data. *Natural Hazards, Special Issue on Remote Sensing*, Vol. 35, No. 3, pp. 321–342.

Lassau, S.A., Hochuli, D.F., Cassis, G., and Reid, C.A.M. 2005. Effects of habitat complexity on forest beetle diversity: Do functional groups respond consistently. *Diversity and Distributions*, Vol. 11, No. 1, pp. 73–82. doi: 10.1111/j.1366-9516.2005.00124.x.

Lassueur, T., Joost, S., and Randin, C.F. 2006. Very high resolution digital elevation models. Do they improve models of plant species distribution? *Ecological Modelling*, Vol. 198, No. 1–2, pp. 139–153. doi: 10.1016/j.ecolmodel.2006.04.004.

Lawton, J.H., Bignell, D.E., Bolton, B., Bloemers, G.F., Eggleton, P., Hammond, P.M., Hodda, M., Holt, R.D., Larsen, T.B., Mawdsley, N.A., Stork, N.E., Srivastava, D.S., and Watt, A.D. 1998. Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest. *Nature*, Vol. 391, No. 6662, pp. 72–76. doi: 10.1038/34166.

Legendre, P. and Legendre, L. 1998. *Numerical ecology.* 2nd English edition. Elsevier Science BV, Amsterdam. pp. 870.

Lemay, V. and Newton, P., co-chairs. 2007. IUFRO, OMNR and CFS conference on Complex Stand Structures and Associated Dynamics. 29 July – 2 August 2007, Sault Ste. Marie, ON. OMNR Forest Research Information Paper No. 167, 117 p.

Lindenmayer, D.B., Margules, C.R., and Botkin, D.B. 2000. Indicators of biodiversity for ecologically sustainable forest management. *Conservation Biology*, Vol. 14, No. 4, pp. 941–950. doi: 10.1046/j.1523-1739.2000.98533.x.

Linke, J., Betts, M.G. et al. 2007. In: *Understanding Forest Disturbance and Spatial Pattern: Remote Sensing and GIS Approaches*. Wulder, M.A. and S.E. Franklin (eds.). Taylor and Francis, Boca Raton FL. pp. 1–29.

MacArthur, R.H. and MacArthur, J.W. 1961. On bird species diversity. *Ecology*, Vol. 42, No. 3,pp. 594–598. doi: 10.2307/1932254.

McDermid, G.J., Hall, R.J., Sanchez-Azofeifa, G.A., Franklin, S.E., Stenhouse, G.B., Kobliuk, T., and LeDrew, E.F. 2009. Remote sensing and forest inventory for wildlife habitat assessment. *Forest Ecology and Management*, Vol. 257, No. 11, pp. 2262–2269. doi: 10.1016/j.foreco.2009.03.005.

McElhinny, C., Gibbons, P., Brack, C., and Bauhus, J. 2005. Forest and woodland stand structural complexity: Its definition and measurement. *Forest Ecology and Management*, Vol. 218, No. 1–3, pp. 1–24. doi: 10.1016/j.foreco.2005.08.034.

McElhinny, C., Gibbons, P., and Brack, C. 2006. An objective and quantitative methodology for constructing an index of stand structural

complexity. *Forest Ecology and Management*, Vol. 235, No. 1–3, pp. 54–71. doi: 10.1016/j.foreco.2006.07.024.

Municipality of Chelsea, 2007. *Chelsea: Demographic profile.* Available from <http://www.chelsea.ca/renseignements_generaux/profil_e.php> [cited 28 January 2010].

NCC (National Capital Commission). 2002. 1:25,000 Scale Topographic Mapping – Contours.shp [computer file]. NCC, Ottawa, Ontario.

Neumann, M. and Starlinger, F. 2001. The significance of different indices for stand structure and diversity in forests. *Forest Ecology and Management*, Vol. 145, No. 1–2, pp. 91–106. doi: 10.1016/S0378-1127(00)00577-6.

Newsome, A.E. and Catling, P.C. 1979. Habitat preferences of mammals inhabiting heathlands of warm temperate coastal, montane and alpine regions of southeastern Australia. In *Ecosystems of the World. Vol. 9A. Heathlands and Related Shrublands of the World*. Edited by R.L. Specht. Elsevier, Amsterdam, Netherlands.

Noss, R.F. 1999. Assessing and monitoring forest biodiversity: A suggested framework and indicators. *Forest Ecology and Management*, Vol. 115, No. 2, pp. 135–146. doi: 10.1016/S0378-1127(98)00394-6.

Olthof, I., and King, D.J. 2000. Development of a forest health index using multispectral airborne digital camera imagery. *Canadian Journal of Remote Sensing*, Vol. 26, No. 3, pp. 166–176.

Pasher, J. 2009. Forest structural complexity in a temperate hardwood forest: a geomatics approach to modelling and mapping indicators of habitat and biodiversity. Ph.D. thesis, Carleton University, Ottawa, Ontario. 206p.

Pasher, J. and King, D.J. 2010. Multivariate forest structure modelling and mapping using high resolution airborne imagery and topographic information. *Remote Sensing of Environment*, Vol. 114, No. 8, pp. 1718–1732. doi: 10.1016/j.rse.2010.03.005.

Pasher, J. and King, D.J. 2011. Development of a forest structural complexity index based on multispectral airborne remote sensing and topographic data. *Canadian Journal of Forest Research*, Vol. 41, No. 1, pp. 44–58. doi: 10.1139/X10-175.

Pielou, E.C. 1984. The interpretation of ecological data: a primer on classification and ordination. John Wiley and Sons, New York.

Pisaric, M.F.J., King, D.J., MacIntosh, D., and Bemrose, R. 2008. Impact of the 1998 ice storm on the health of sugar maple (*Acer saccharum* Marsh) dominated forest stands in Gatineau Park, Quebec. *Journal of the Torrey Botanical Society*, Vol. 135, No. 4, pp. 530–539. doi: 10.3159/08-RA-053R.1.

Plan Environnemental de la Municipalité de Chelsea. 1991. *Environmental Plan.* Available from <http://www.chelsea.ca/environnement/plan_e.php> [cited 19 January 2010].

Rapport, D.J. and Whitford, W.G. 1999. How ecosystems respond to stress: common properties of arid and aquatic systems. *BioScience*, Vol. 49, No. 3, pp. 193–203. doi: 10.2307/1313509.

Seed, E.D. and King, D.J. 2003. Shadow brightness and shadow fraction relations with effective LAI: Importance of canopy closure and view angle in mixedwood boreal forest. *Canadian Journal of Remote Sensing,* Vol. 29, No. 3, pp. 324–335. doi: 10.5589/m03-003

Smith, G.F., Gittings, T., Wilson, M., French, L., Oxbrough, A., O'Donoghue, S., O'Halloran, J., Kelly, D.L., Mitchell, F.J.G., Kelly, T., Iremonger, S., McKee, A.M., and Giller, P. 2008. Identifying practical indicators of biodiversity for stand-level management of plantation forests. *Biodiversity and Conservation*, Vol. 17, No. 5, pp. 991–1015. doi: 10.1007/s10531-007-9274-3.

Staudhammer, C., and Lemay, V.M. 2001. Introduction and evaluation of possible indices of stand structural diversity. *Canadian Journal of Forest Research*, Vol. 31, No. 7, pp. 1105–1115. doi: 10.1139/x01-033.

Steffan-Dewenter, I., Kessler, M., Barkmann, J., Bos, M.M., Buchori, D., Erasmi, S., Faust, H., Gerold, G., Glenk, K., Gradstein, S.R., Guhardja, E., Harteveld, M., Hertel, D., Höhn, P., Kappas, M., Köhler, S., Leuschner, C., Maertens, M., Marggraf, R., Migge-Kleian, S., Mogea, J., Pitopang, R., Schaefer, M., Schwarze, S., Sporn, S.G., Steingrebe, A., Tjitrosoedirdjo, S.S., Tjitrosoemito, S., Twele, A., Weber, R., Woltmann, L., Zeller, M., and Tscharntke, T. 2007. Tradeoffs between income, biodiversity, and ecosystem functioning during tropical rainforest conversion and agroforestry intensification. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 104, No. 12, pp. 4973–4978. doi: 10.1073/pnas.0608409104.

ter Braak, C.J.F. and Šmilauer, P. 2002. CANOCO reference manual and CanoDraw for Windows. User's guide: Software for Canonical Community Ordination (v.4.5). Microcomputer Power, Ithaca, NY. 500 pp.

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M.C., Schwager, M., and Jeltsch, F. 2004. Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. *Journal of Biogeography*, Vol. 31, No. 1, pp. 79–92. doi: 10.1046/j.0305-0270.2003.00994.x.

U.S. Forest Service. 2006. *Forest management guides.* Available from <http://www.ncrs.fs.fed.us/fmg/nfmg/fm101/silv/p4_ecology.html> [cited 2 January 2010].

Williams, S.E., Marsh, H., and Winter, J. 2002. Spatial scale, species diversity, and habitat structure: small mammals in Australian tropical rain forest. *Ecology*, Vol. 83, No. 5, pp. 1317–1329. doi: 10.1890/0012-9658(2002)083[1317:SSSDAH]2.0.CO;2.

Wunderle, A.L., Franklin, S.E., and Guo, X.G. 2007. Regenerating boreal forest structure estimation using SPOT-5 pansharpened imagery. *International Journal of Remote Sensing*, Vol. 28, No. 19, pp. 4351–4364. doi: 10.1080/01431160701244849.

Zenner, E.K. 2004. Does old-growth condition imply high live-tree structural complexity? *Forest Ecology and Management*, Vol. 195, No. 1–2, pp. 243–258. doi: 10.1016/j.foreco.2004.03.026.