

## Data-driven regionalization of forested and non-forested ecosystems in coastal British Columbia with LiDAR and RapidEye imagery



Shanley D. Thompson <sup>a,\*</sup>, Trisalyn A. Nelson <sup>a</sup>, Ian Giesbrecht <sup>b</sup>, Gordon Frazer <sup>b</sup>, Sari C. Saunders <sup>c</sup>

<sup>a</sup> Spatial Pattern Analysis and Research Lab, Department of Geography, University of Victoria, PO BOX 3060 STN CSC, Victoria, British Columbia, V8W 3R4, Canada

<sup>b</sup> Hakai Institute, Box 309, Heriot Bay, British Columbia, V0P 1H0, Canada

<sup>c</sup> BC Ministry of Forests, Lands and Natural Resources, 2100 Labieux Rd., Nanaimo British Columbia, V9T 6E9, Canada

### ARTICLE INFO

#### Article history:

Received 21 October 2015

Received in revised form

25 January 2016

Accepted 1 February 2016

Available online 1 March 2016

#### Keywords:

Ecosystem

Ecoregion

Classification

Cluster analysis

Ecosystem structure

Terrain

### ABSTRACT

Traditionally, forest inventory and ecosystem mapping at local to regional scales rely on manual interpretation of aerial photographs, based on standardized, expert-driven classification schemes. These current approaches provide the information needed for forest ecosystem management but constrain the thematic and spatial resolution of mapping and are infrequently repeated. The goal of this research was to demonstrate the utility of an unsupervised, quantitative technique based on Light Detection And Ranging (LiDAR) data and multi-spectral satellite imagery for mapping local-scale ecosystems over a heterogeneous landscape of forested and non-forested ecosystems. We derived a range of metrics characterizing local terrain and vegetation from LiDAR and RapidEye imagery for Calvert and Hecate Islands, British Columbia. These metrics were used in a cluster analysis to classify and quantitatively characterize ecological units across the island. A total of 18 clusters were derived. The clusters were attributed with quantitative summary statistics from the remotely sensed data inputs and contextualized through comparison to ecological units delineated in a traditional expert-driven mapping method using aerial photographs. The 18 clusters describe ecosystems ranging from open shrublands to dense, productive forest and include a riparian zone and many wetter and wetland ecosystems. The clusters provide detailed, spatially-explicit information for characterizing the landscape as a mosaic of units defined by topography and vegetation structure. This study demonstrates that using various types of remotely sensed data in a quantitative classification can provide scientists and managers with multivariate information unique from that which results from traditional, expert-based ecosystem mapping methods.

© 2016 Elsevier Ltd. All rights reserved.

### 1. Introduction

An ecosystem, as defined by the Convention on Biological Diversity (CBD) and the Millennium Ecosystem Assessment (MA), is a dynamic complex of biotic components, and the interaction between these components and their physical environment. Ecosystems can be conceptualized at a variety of spatial scales, in a hierarchical manner (Bailey, 1987; Franklin, 2013). As reviewed by Whittaker (1967) and Kent, Gill, Weaver, & Armitage (1997), boundaries between ecosystems are, in reality, most often gradual

and fuzzy; attributes such as species composition overlap along environmental gradients. The science of delineating regions (e.g., ecological or biogeographical) in geographical space is referred to as *regionalization* (Loveland & Merchant, 2004; Olstad, 2012) and is important for the understanding and management of the natural world (Mackey, Berry, & Brown, 2007; McMahon, Wiken, & Gauthier, 2004). Envisioned use of regions often determines the appropriate scale for mapping and thus which and how many regions can or will be delineated for a given project. National- or regional-scale conservation planning, resource management, and ecosystem services assessments benefit from maps that delineate spatial units on the basis of regional climate, large landforms, land cover class, and/or patterns of primary productivity (e.g., Handcock & Csillag, 2002; Leathwick, Overton, & McLeod, 2003; Sayre,

\* Corresponding author.

E-mail address: [sdthompson@uvic.ca](mailto:sdthompson@uvic.ca) (S.D. Thompson).

Comer, Harumi, & Cress, 2009). Differences in vegetation composition and structure, as influenced by micro-scale soil moisture and nutrient conditions, define ecological communities at plot to landscape scales, supporting more specific local and regional science and management activities (e.g., Banner, Meidinger, Lea, Maxwell, & Sacken, 1996).

Many jurisdictions rely on the manual interpretation of aerial photographs and field observations to delineate ecological regions and forest attributes at local to national scales. Interpreters use a set of methodological standards (e.g., Canadian Forest Service, 2001; Resource Information Management Branch, 2005; Resources Inventory Committee, 1998) that provide the information desired for forest or ecosystem management, but inherently limit the spatial and thematic resolutions of the resulting map. For instance, forest stand or ecosystem polygons are delineated and attributed with information regarding composition and structure that is considered representative of the entire polygon (Wulder et al., 2006). Alternative methods for mapping that incorporate other types of remotely sensed imagery in a quantitative classification may be used to delineate local regions with measurable attributes at improved spatial precisions. Further, relative to the manual delineation of ecological regions, a quantitative approach may be automated or semi-automated, offering increased consistency, repeatability, and cost-efficiency for monitoring over time and across large areas (MacMillan, Moon, & Coupé, 2007; Morgan, Gergel, & Coops, 2010).

Remote sensing can capture various structural, compositional, and functional aspects of ecosystems. Multispectral imagery at high to moderate spatial resolutions is frequently used to distinguish among vegetation life form and age class (e.g., Johansen, Coops, Gergel, & Stange, 2007; Valeria, Laamrani, & Beaudoin, 2014; Wulder, Skakun, Kurz, & White, 2004). Multispectral data are also used generate indices such as the Normalized Difference Vegetation Index (NDVI) (Rouse, Haas, Schell, & Deering, 1974), which correlates well with patterns of net primary productivity (Goward, Tucker, & Dye, 1985) and has been useful for wetland delineation (e.g., Barron, Emelyanova, VanNiel, Pollock, & Hodgson, 2014; Dechka, Franklin, Watmough, Bennett, & Ingstrup, 2002; White, Lewis, Green, & Gotch, 2015). Hyperspectral imagery is well suited to species composition mapping (e.g., Dalponte, Orka, Gobakken, Gianelle, & Naesset, 2013; Feret & Asner, 2013). Light Detection And Ranging (LiDAR) technology captures highly accurate, direct measurements of three-dimensional vegetation structure (Van Leeuwen & Nieuwenhuis, 2010; Lim, Treitz, Wulder, St-Onge, & Flood, 2003). In particular, LiDAR provides excellent measurements of forest structural properties such as tree height and canopy closure (Coops et al., 2007; Holmgren, 2004; Lefsky et al., 1999). The high spatial resolution of LiDAR-derived elevation data may (Hogg & Holland, 2008; Maxa & Bolstad, 2009) or may not (Knight, Tolcser, Corcoran, & Rampi, 2013) lead to improved wetland delineation compared to use of aerial photography, especially in areas of low topographic variation.

As no single sensor can capture all of the structural, functional and compositional characteristics of terrestrial ecosystems at the spatial resolution and extent typically desired by environmental managers, researchers often combine multiple types of remotely sensed data together to generate improved and more complex maps (e.g., Jones, Coops, & Sharma, 2010; Ke, Quackenbush, & Im, 2010; Wulder, Han, White, Sweda, & Tsuzuki, 2007). Combining various types of remotely sensed data in a quantitative (statistical) regionalization, researchers have effectively captured environmental domains over large regions (e.g., Coops, Wulder, & Iwanicka, 2009; Fitterer, Nelson, Coops, & Wulder, 2012; Handcock & Csillag, 2002; Powers et al. 2012). Drawing on the high spatial resolution and information content of remotely sensed data, regionalizations

are also possible that highlight ecosystem patterns, gradients, and ecotones at local scales (Hargrove & Hoffman, 2004; Kupfer, Gao, & Guo, 2012; Long, Nelson, & Wulder, 2010; Olstad, 2012). These smaller regions have higher internal homogeneity, a desirable property for local-scale management (Bryan, 2006) and for serving as strata for field-based sampling. As computing power increases and is more able to accommodate very large datasets, quantitative methods that combine numerous remotely sourced datasets at high to moderate spatial resolutions, including LiDAR, can produce information-rich (i.e., multivariate) categorical maps for environmental research and management investigated. Researchers continue to evaluate new techniques and data for regionalization in ecological contexts (e.g., Xu et al., 2014; Niesterowicz & Stepinski, 2013), however, research incorporating vegetation structural data from LiDAR with other remotely sensed data for local-scale regionalization has been limited.

The goal of this research was to demonstrate the utility of LiDAR and high spatial resolution multispectral imagery for mapping and characterizing the variety of local-scale ecosystems over a complex landscape of forested and non-forested ecosystems on the outer coast of British Columbia. Given that different ecosystem maps of the same area are possible depending on data and methodology, the objective was to use an unsupervised, quantitative regionalization method to determine what types of ecosystem units would be captured when using high resolution remotely sensed measures representing vegetation structure and topography. Following attribution of the resulting regions with the remotely sensed data, we contextualize our findings by comparing the results of the unsupervised classification to those of an expert-driven classification. We conclude with a discussion of implications for ecosystem management and recommendations for local-scale ecosystem mapping that will be useful in a variety of remote regions characterized by heterogeneous landscapes where field data collection is logistically challenging.

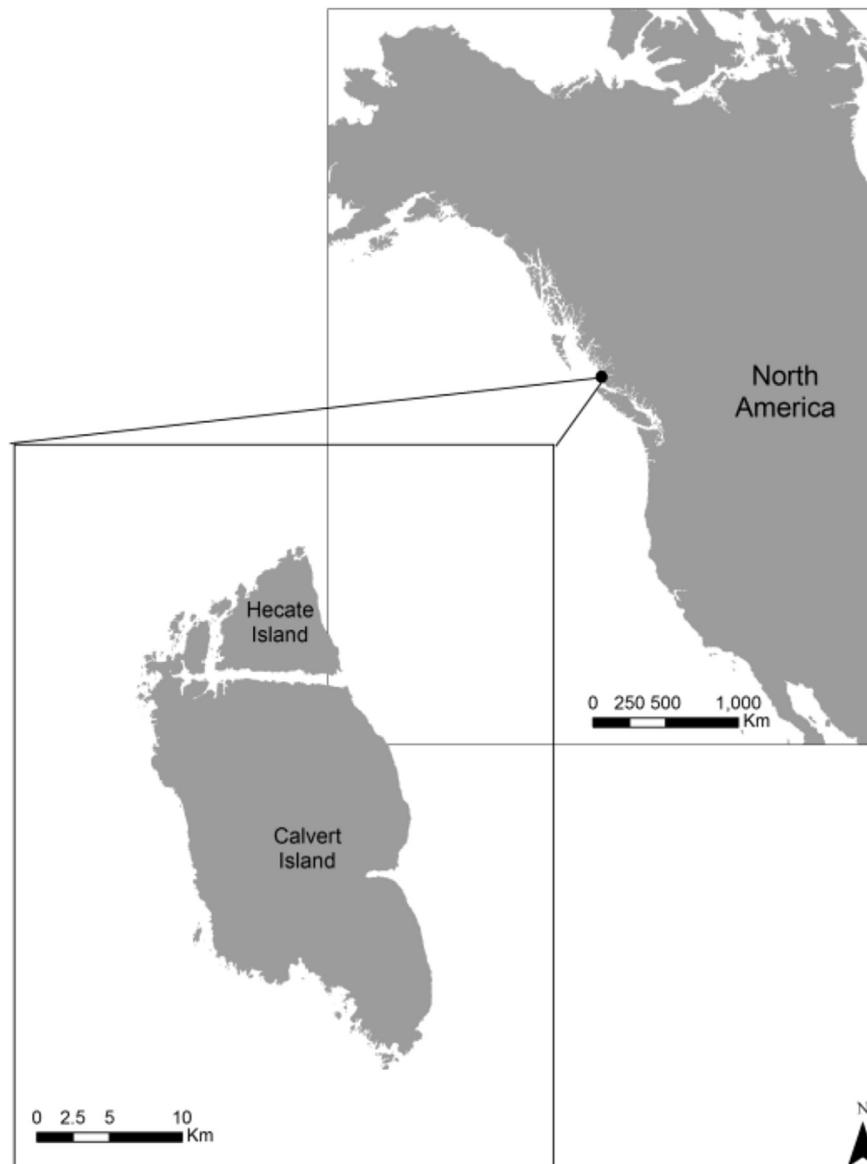
## 2. Methods

### 2.1. Study area

Calvert and Hecate Islands (total 37,433 ha) are remote islands on the central coast of British Columbia, Canada (Fig. 1). The islands are dominated by fairly subdued topography from low to moderate elevations in a biogeoclimatic unit classified as the Coastal Western Hemlock zone, Very Wet, Hypermaritime subzone, Central variant (CWHvh2). Elevations in the study area reach ~1000 m; these areas classified as the Mountain Hemlock zone, Wet Hypermaritime Subzone (MHvh1). Gridded climate data for the region (ClimateBC v5 – Wang, Hamann, Spittlehouse, & Murdock, 2012) indicate that average (1990–2012) annual temperature is 7.6°C for the CWHvh2 and 4.8 °C for the MHvh1. Average annual precipitation in these biogeoclimatic units is 3512.2 mm and 5140.7 mm, respectively. High precipitation, abundant fog, and low evapotranspiration result in an abundance of wet soils, wetland ecosystems, and relatively unproductive forests (Banner, LePage, Moran, & de Groot, 2005).

### 2.2. Remotely sensed data

The quantitative regionalization used in this study relied on a combination of LiDAR data and multispectral satellite imagery with differing native spatial resolutions. All data were aggregated to a common spatial resolution of 20 m, a scale widely used in LiDAR-based forest inventories, soil mapping, and terrain analysis (Brosofske, Froese, Falkowski, & Banskota, 2014; Gillin, Bailey, McGuire, & Prisley, 2015). Airborne LiDAR data were acquired



**Fig. 1.** Study area in coastal British Columbia, Canada.

across Calvert and Hecate Islands in August 2012. Mounted on a fixed-wing aircraft flying at 1150 m above ground level with a maximum scan angle of  $\pm 26^\circ$ , the discrete-return (4 returns/pulse), small-footprint (0.3 mrad) V-Gen LiDAR system acquired data with an average point density of 2.32 pt/m<sup>2</sup> and a standard deviation of 1.07 pt/m<sup>2</sup>. Relative and absolute vertical accuracies at one standard deviation were estimated by the data provider to be  $\pm 15$  cm and  $\pm 30$  cm, respectively. A small number of data voids in the LiDAR coverage were present and were excluded from the analysis.

Several data processing steps were required to extract candidate topographic indices and vegetation metrics from the LiDAR dataset. First, LiDAR returns were separated into ground and non-ground feature classes using TerraScan's (Terrasolid Ltd.) automated ground-filtering routine. Manual refinement of the point classification was undertaken in areas of steep, complex topography to correct any obvious classification errors created by the automated ground filter (Merrick Advanced Remote Sensing—MARS—software suite). Second, we used the BLAST extension pack (blast2dem) of LAStools (rapidlasso GmbH) to construct a gridded one-meter,

'bare-earth' Digital Terrain Model (DTM) from a Triangulated Irregular Network (TIN) of ground-classified points. The gridded one-meter DTM was subsequently resampled to a spatial resolution of 20 m using the cell-area weighted mean elevation prior to computing several topographic indices (System for Automated Geoscientific Analyses—SAGA). Third, we converted (i.e., height normalized) each elevation in the LiDAR point cloud to an estimated height in meters above the ground surface using the difference in elevation between a point and its corresponding elevation on the TIN of ground-classified points (LAStools). We extracted the following terrain indices from the 20 m DTM: slope (%), the Topographic Radiation ASPect (TRASP) (Roberts & Cooper, 1989), the Topographic Position Index (TPI) normalized to the local standard deviation in elevation as per De Reu et al. (2013), and the Topographic Wetness Index (TWI) calculated using the D-infinity algorithm (Tarboton, 1997) (Table 1). Elevation, slope, TRASP, TPI, and TWI are considered useful inputs in automated ecosystem classification and predictive vegetation mapping because these variables directly and indirectly influence plant growth and

**Table 1**

Description of remotely sensed variables used in analyses.

Variable	Derivation	Interpretation
Elevation	Elevation above sea level in metres, derived from DEM.	Elevations in the study site range from 0 (sea level) to 1012 m.
Gap fraction	The ratio of total number of laser points <2 m in height to the total number of laser points within a grid cell.	A density-based statistic, higher values represent lower canopy cover.
Height – coefficient of variation	2 <sup>nd</sup> central moment about the mean	A high coefficient of variation means the vegetation has a strong vertical heterogeneity.
Height – maximum	LiDAR Canopy Height Model = value of maximum laser point return per grid cell.	Height of the tallest tree or shrub in the grid cell.
Height – mean	Arithmetic mean height (m) of laser points ≥ the minimum height threshold (2 m in this study).	Average height of tree canopy.
Normalized Difference Vegetation Index (NDVI)	NDVI = (NIR reflectance – red reflectance)/(NIR reflectance + red reflectance) (Rouse et al. 1974).	A measure of vegetation greenness, widely used as a proxy for productivity. Values near zero are considered non-vegetated or sparsely vegetated.
Slope	Slope in percent rise, derived from the DEM	Slopes in the study area range from level (0%) to very steep (~120%).
Topographic Position Index (TPI)	TPI = $z_o - z_{avg}$ where $z_o$ = elevation of focus pixel and $z_{avg}$ = average elevation neighbouring pixels (Weiss, 2001). Calculated here for a 100 m neighbourhood and standardized to z-scores as per De Reu et al. (2013).	Positive values (>0.5 in this study) represent ridges and upper slopes, whereas negative values (<-0.5 in this study) represent lower slopes and valley bottoms; values near zero are mid-slope or flat (De Reu et al., 2013; Tagil & Jenness, 2008; Weiss, 2001). Values near 0 (~< 0.5) are N, E, NE or NW facing (cooler slopes) while those near 1 (or ~ >0.5) are S, W, SW, or SE facing (warm slopes).
Topographic Radiation ASPECT (TRASP)	TRASP = $(1 - \cos((3.1416/180)(\text{aspect}-30))) / 2$ where aspect is in degrees (Roberts & Cooper, 1989).	
Topographic Wetness Index (TWI)	TWI = $\ln(\text{specific catchment area}/\tan(\text{slope}))$ . Calculated using the D-infinity algorithm (Tarboton, 1997).	Higher values are considered wetter, lower values drier.

community composition via effects on temperature, precipitation, soil moisture, soil nutrients, and wind exposure (Franklin, 1995). These types of topographic derivatives are commonly used in predictive ecosystem mapping (e.g., Chastain & Struckhoff, 2008; Dobrowski, Safford, Cheng, & Ustin, 2008; Fraser, McLennan, Ponomarenko, & Olthof, 2012; MacMillan et al. 2007), and standard ecosystem classification in British Columbia at the local scale relies heavily on topographic concepts (Banner et al., 1996).

We used the height normalized LiDAR point cloud to compute several area-based (gridded) canopy height and density metrics on the same 20 m grid shared by the DTM and derivative topographic indices. Area-based height metrics included estimates of the mean, maximum, standard deviation, coefficient of variation (CV), and percentiles of canopy height computed using all laser points with heights greater than or equal to 2 m within a single grid cell (Magnussen & Boudewyn, 1998; Næsset, 2002). A measure of canopy density or gap fraction was estimated for each grid cell as the ratio of the number laser points found below 2 m to the total number of laser points located within the cell (Hopkinson & Chasmer, 2009). All grid cells devoid of laser returns greater than 2 m in height were considered to be non-forested in this study.

Multispectral reflectance data (5 m spatial resolution) were acquired from the RapidEye satellite sensor (BlackBridge) in 2011. Using PCI Geomatica, the RapidEye imagery was first converted to Top of Atmosphere units to adjust for sun angle and earth-sun distance and further corrected for atmospheric noise using the Dark Object Subtraction method (Chavez, 1996). We calculated the NDVI as an overall proxy for vegetation greenness or productivity. The NDVI was resampled to 20 m spatial resolution to match the LiDAR data. Finally, as our focus was on terrestrial systems, we used a provincial water body database,<sup>1</sup> as well as the NDVI layer (values < 0.2) to remove water and un-vegetated pixels.

### 2.3. Existing ecosystem data

The study area was recently mapped using British Columbia's provincial standard of Terrestrial Ecosystem Mapping (TEM) (Green, 2014; Resources Inventory Committee, 1998). TEM uses a well-defined hierarchical classification system referred to as Biogeoclimatic Ecosystem Classification (BEC) (Banner et al., 1996; Pojar, Klinka, & Meidinger, 1987). At the regional level, vegetation, soils, and topography are used to infer climatic zones. Within these regional climate zones, *subzones* are defined by the climax plant association typical of a zonal site (zonal sites are those on which vegetation is primarily influenced by climate, not edaphic features). *Variants* may be used to distinguish finer-scaled variation in climate within the Subzone. Within each subzone or variant, a set of ecosystems is defined, which are referred to as *site series*. The site series (each given a unique two letter and two- or three-digit code) are classified with reference to established rules that associate each with a particular combination of terrain, soil and vegetation characteristics.

TEM relies on analysts to interpret aerial photographs, topographic maps, and other available geospatial information to identify site series. Interpreted TEM polygons may contain up to three site series each, where individual patches of these site series are smaller than the minimum mapping unit (typically 0.5–2.0 ha). In these complex polygons, the relative proportion of each site series is noted, although the exact location is not. The TEM for the study area identified 38 site series, including six non-vegetated site series, seven non-forested site series, and 25 forested site series (Green, 2014). Ground sampling is used to calibrate and assess the TEM. Field surveys were conducted in 2013 and 2014 by provincial ecologists and Hakai Institute and affiliated researchers, primarily in the northern part of Calvert Island and the southern part of Hecate Island following provincial BEC standards (BC Ministry of Forests and Range and BC Ministry of Environment, 2010). Field plots identified 10 ecosystem types, including several sub-types not explicitly captured in the TEM (e.g., shallow minerotrophic blanket bogs, and deep ombrotrophic blanket bogs).

<sup>1</sup> The Freshwater Atlas (FWA): <http://geobc.gov.bc.ca/base-mapping/atlas/fwa/index.html>.

## 2.4. Unsupervised regionalization

Multivariate clustering was used to determine the types of local-scale ecosystems that are distinguishable on Calvert and Hecate Islands with key metrics derived from LiDAR and RapidEye data. Clustering is an unsupervised, multivariate technique, which can be used to group observations or sample units that are similar with respect to the variables used to define them. These techniques generally seek to minimize within-group variability and maximize between-group variability.

A Spearman's rank correlation analysis was conducted to assess collinearity among all the aforementioned terrain (slope, TRASP, TWI, and TPI) and vegetation variables (NDVI, mean height, maximum height, and gap fraction). Correlations were fairly low (<0.5) between the terrain variables and the vegetation variables, as well as among the terrain variables (Table 2). Correlations were moderate to high (>0.7) among the vegetation variables. From these four vegetation variables, we chose to retain one of the LiDAR vegetation variables (mean height) and the multispectral vegetation variable (NDVI) as inputs for the clustering.

With each of the chosen variables having been clipped and resampled to a consistent spatial extent and spatial resolution (20 m), and being projected to the same datum (NAD 83), a "raster stack" was built (using the "raster" package in R) to combine the variables into a single data frame for clustering. Clustering was conducted separately for the forested and non-forested regions of the island, since the average height and coefficient of variation of height from LiDAR were only available for the forested region (defined here as vegetation > 2 m in height). All data were standardized to z-scores prior to clustering in order to ensure approximately equal weight for all variables. A clustering methodology called TwoStep, employing a probability-based distance measure, was implemented in SPSS (v22). The TwoStep approach was chosen because it is well suited to very large datasets. It begins with an initial partition of the data, followed by a hierarchical clustering of these partitions. The number of resulting clusters may be explicitly specified by the analyst. To select an appropriate number, we first considered several statistical criteria: the Bayesian Information Criterion (BIC), the Calinski–Harabasz (CH) criterion (Calinski & Harabasz, 1974) and the Average Silhouette Width (ASW) (Table 3). Whereas the BIC is calculated internally during the TwoStep clustering procedure, we calculated the CH and ASW statistics based on k-means partitioning using the pamk() function from the fpc library in R (v3.1.2). In choosing the number of clusters, we also took into consideration the number of site series identified in the existing ecosystem map and in field surveys conducted in 2013 and 2014.

Once the desired number of clusters was determined, the

clusters were mapped in geographic space and a majority filter of 4 × 4 pixels was applied to remove speckle in the final product. Otherwise, no minimum mapping unit was imposed. Clusters were attributed with summary statistics for each data layer input into the cluster analysis (e.g., NDVI, elevation, and slope). Additionally, the LiDAR-derived variables of gap fraction and maximum height (excluded from the cluster analysis because of high collinearity with each other as well as with NDVI and average mean height) were also used to describe the clusters.

## 2.5. Map comparisons

We used several approaches to contextualize our unsupervised ecosystem map relative to the expert-driven TEM. First, each cluster was attributed with field survey information by overlaying the field survey points on the clusters in ArcGIS, extracting the cluster value at each point, and summarizing by cluster type. Second, TEM information was attributed to each cluster as follows: the majority (modal) cluster value occurring within each TEM polygon was extracted and joined to the TEM database. Next, the total area of each TEM site series within each cluster type was derived using information on the proportional area of each site series within each TEM polygon. We used pie charts to summarize the detailed composition of our clusters in terms of TEM units and the composition of TEM units in terms of our clusters.

Third, we calculated the diversity of different TEM site series comprising each cluster, as well as the inverse (diversity of clusters within each TEM site series) to determine correspondence between the two mapping methodologies. More diverse clusters or TEM site series represent poor agreement between the two maps. We assessed this relationship both graphically (via pie-charts) and numerically, via Simpson's Diversity Index, which is commonly used in ecological studies to indicate species diversity in a way that captures both richness (a count of species) and evenness (relative abundance of each species). The Simpson's Diversity Index is calculated using the formula

$$D = 1 - \left[ \sum(n/N)^2 \right]$$

where n is the number of individuals of a particular species, and N is the total number of all individuals for a given location. In our case, rather than species diversity, we calculated the index to measure the diversity of site series in each cluster and again to measure the diversity of clusters within each site series. The values ranged from 0 to 1, with higher values indicating a poor association between the two mapping methodologies. We note, however, that the index of diversity used to compare the maps at this level does not take into account ecological similarity among site series and that a cluster

**Table 2**  
Spearman's rank correlation coefficients.

	Elevation (m)	Gap fraction	Height – coefficient of variation	Height – maximum (m)	Height – mean (m)	NDVI	Slope (%)	TPI	TRASP	TWI
Elevation (m)	1.00									
Gap fraction	-0.02	1.00								
Height – coefficient of variation	-0.06	-0.32	1.00							
Height – maximum (m)	-0.03	-0.90	0.38	1.00						
Height – mean (m)	-0.06	-0.92	0.41	0.93	1.00					
NDVI	0.09	-0.76	0.30	0.73	0.75	1.00				
Slope (%)	0.48	-0.44	0.17	0.40	0.39	0.41	1.00			
TPI	0.11	0.07	-0.18	-0.12	-0.15	-0.09	0.03	1.00		
TRASP	0.04	-0.09	0.06	0.10	0.12	0.11	0.08	0.03	1.00	
TWI	-0.14	0.20	-0.02	-0.15	-0.13	-0.15	-0.46	-0.59	-0.02	1.00

NDVI = Normalized Difference Vegetation Index (NDVI), TPI = Topographic Position Index, TRASP = Topographic Radiation ASPect, TWI = Topographic Wetness Index.

containing many similar site series may not be as severely mismatched to the TEM as a cluster containing the same number of ecologically dissimilar site series. To avoid bias due to the variation in the size and number of each TEM polygon and each TEM site series, TEM site series that are rare on the landscape were excluded (those below the 5th percentile in number of polygons as well as area). Individual TEM polygons below the 5th percentile in area, regardless of site series, were also removed.

Finally, to allow direct comparison of the two maps with a single classification scheme, we assigned each TEM site series and each cluster to one of six broad ecosystem classes. We used broad ecosystem classes suggested by the local TEM authors as well as the regional handbook for ecosystem classification (Green & Klinka, 1994). Clusters were assigned to broad classes based on interpretation of the calculated terrain metrics, particularly topographic wetness, slope and topographic position. The generalized classes maintain distinctions between forested and non-forested ecosystems and emphasize differences along a moisture gradient. To compare the unsupervised regionalization with the expert-driven TEM, we calculated the spatial extent of each of these generalized ecosystem types for both systems.

### 3. Results

#### 3.1. Unsupervised regionalization

We identified 12 forested clusters (Clusters 1–12) and six non-forested clusters (Clusters 13–18) (Figs. 2 and 3), drawing on several lines of information to judge an appropriate and useful number of clusters in each group. The statistically optimal number of clusters varied from two to three for non-forested regions and two to 25 for forested regions (Table 3), while the TEM and field plot data suggested upwards of 32 vegetated site series may be present on the island. Upon visual inspection of the resulting map, 18 clusters were chosen to provide the desired spatially-detailed characterization of the landscape without sub-dividing beyond the limits of the data or our ability to interpret, describe, and—potentially—sample the landscape mosaic in the field.

Cluster 1 described a forest of intermediate wetness (mean TWI = 6.07) and moderate productivity (mean NDVI = 0.75, mean height = 7.7 m), relative to the study area (Table 4, Figs. 4 and 5). It occurred in the middle of the range of elevations found in the study area (average of 292 m) on steep slopes (average of 44%), and cool aspects (mean TRASP = 0.16). Cluster 2 was more productive, and had relatively tall trees (mean NDVI = 0.78, mean maximum height = 20.3, mean height = 13 m). It was found at low elevations (average of 86 m). Clusters 3 and 4 were the most abundant classes across the study area (3903 ha and 3696 ha, or 11.2% and 10.6% of the study area, respectively). They were characterized as forests of intermediate to high wetness (mean TWI = 6.52 and 6.93, respectively) with moderate productivity (mean NDVI = 0.74). Both occurred at low elevations (average < 82 m). Cluster 3 occurred on warm aspects (mean TRASP = 0.83), while Cluster 4 occurred on cool aspects (mean TRASP = 0.16).

Cluster 5 occurred primarily on ridges (mean TPI = 1.52) at low elevations (average 119 m) and was predicted to be the driest cluster (mean TWI = 4.4). Productivity was moderate (mean NDVI = 0.75) and tree heights averaged 9 m. Clusters 6 and 8 had a similar topographic wetness to Clusters 3 and 4, but were considerably less productive (mean NDVI values were 0.64 and 0.65 for Clusters 6 and 8, respectively). Both had open canopies (gap fractions of 0.69) and short trees (average mean and max heights < 6 m). Cluster 6 represented cool aspects, while Cluster 8 was associated with warm aspects. Cluster 7 was found at low elevations (mean = 85 m), at fairly level sites (mean slope = 8%), and

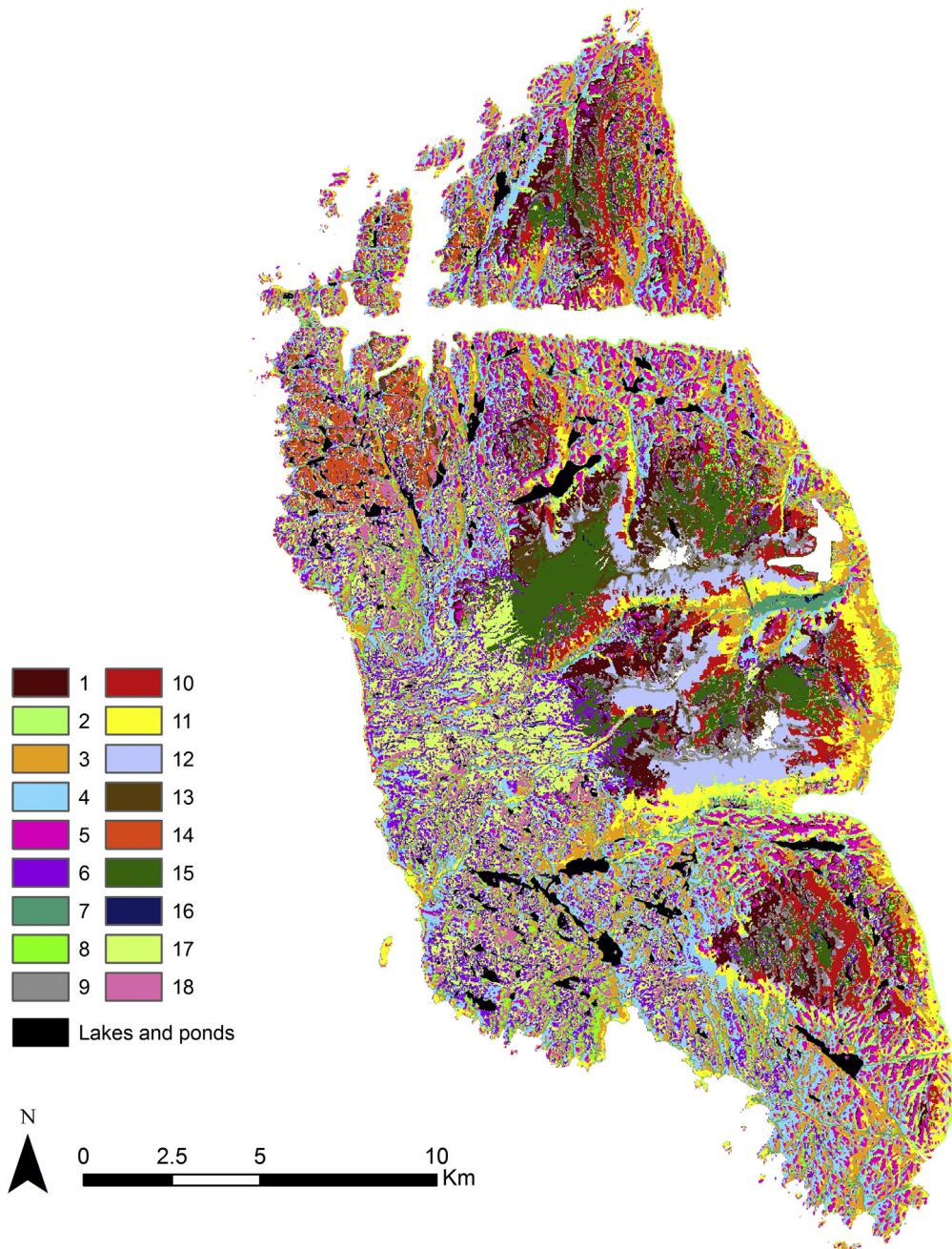
at lower slope positions (mean TPI = −0.88). It was predicted to have very high wetness (mean TWI = 12.89) with moderate productivity (mean NDVI = 0.74). Cluster 9 was predicted to be relatively dry, similar to Cluster 5 (mean TWI = 4.59), with moderate productivity (mean NDVI = 0.71) and short vegetation (mean height 5.73 m). Cluster 9 was found at the highest slope positions among our clusters (mean TPI = 1.55), at high elevations (506 m on average) on very steep slopes (51% on average).

Cluster 10 was found in moderately productive forest (mean NDVI = 0.73) with average tree heights of 6.5 m. It occurred at middle elevations (average 311 m), on moderate slopes (average 33%) of distinctly warm aspect (mean TRASP = 0.86), and was predicted to be found on sites with moderate wetness (mean TWI = 5.96). Clusters 11 and 12 were the tallest, densest, and most productive forests on the island: maximum tree heights averaged >21 m, mean gap fractions were ≤0.16, and NDVI averaged 0.79 and 0.82, respectively. Cluster 11 occurred at low elevations (average of 120 m) and moderate slopes (average 33%), while Cluster 12 was found at higher elevations (average of 461 m) on very steep slopes (average 65%).

The remaining six clusters were characterized as non-forested (mean and maximum vegetation height less than 2 m). All non-forested clusters had low mean NDVI values (≤0.6) and very high mean gap fractions (≥0.93), relative to the forested clusters. Despite this low NDVI and high gap fraction, NDVI boxplots suggest that some non-forested pixels reach forest-like levels of productivity (Fig. 5). Cluster 13 was relatively scarce (849 ha of the study area) and was found across a wide range of elevations (mean 361 m) and at relatively steep slopes (38%). It was predicted to be relatively dry, similar to Clusters 5 and 9 (mean TWI = 4.86). Cluster 14 was also relatively scarce (696 ha in total) and generally restricted to low elevations (79 m average). It had the lowest productivity and lowest canopy cover of any cluster (average NDVI of 0.42 and average gap fraction of 0.98). Cluster 15 was a wetter cluster (mean TWI = 7.03) that occurred at high elevations (average 413 m) across a wide range of aspects. Cluster 16 was predicted to have very high wetness (mean TWI = 11.95), and it was uncommon across the study site (720 ha). Generally it occurred at low elevations (99 m average) and on level sites (mean slope 4%). Cluster 17 and 18 were predicted to have moderately high wetness (mean TWI ranged from 6.82 to 7.20) and were found on very gentle slopes (8% average). Cluster 17 occurred on cool aspects (mean TRASP = 0.16), and was extensive across the island (covering 2449 ha), while Cluster 18 was less abundant (1674 ha) and occurred on warm aspects (mean TRASP = 0.80).

#### 3.2. Map comparisons

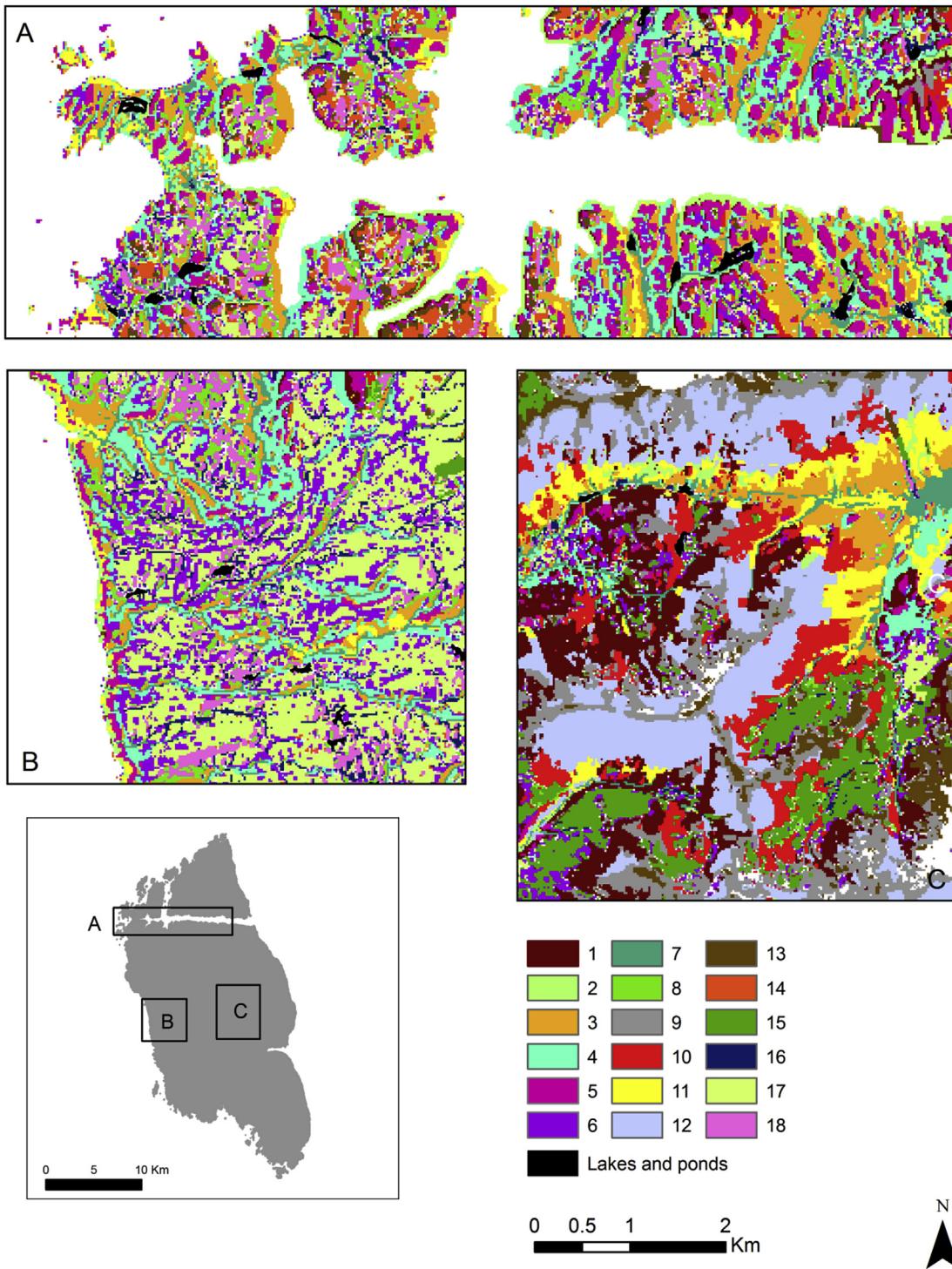
The clusters contained information that was unique from that contained in the TEM. Most clusters in our established set of 18 were comprised of a variety of TEM site series (Table 5, Supplementary Figure A.1) and each TEM site series contained multiple clusters (Supplementary Figure A.2). Measured in terms of relative area, Cluster 14 was the most homogenous, with two site series (bedrock and blanket bogs) comprising 84% its total area, with few site series comprising the remaining area. When measured in terms of the Simpson's diversity index, which was based on counts (i.e., number of TEM site series within each cluster, and number of cluster types within each TEM site series), Cluster 16 was the least diverse (most homogenous), with a Diversity Index value of 0.53 (Fig. 6). Less extensive classes (defined by either method) were often less diverse, presumably in part because the chance of including additional classes is smaller when area is very limited. Cluster 9 was the most diverse, followed by Cluster 7, with values of 0.89 and 0.87, respectively (Fig. 6).



**Fig. 2.** Eighteen clusters representing a range of forested and non-forested terrestrial ecosystems on Calvert and Hecate Islands. Clusters 1 through 12 are forested and Clusters 13 through 18 are non-forested. White areas within the land mass are data voids.

Areas mapped as rivers (RI) in the TEM were entirely captured by Cluster 7 (Fig. A.2, Fig. 7), which we have identified as a riparian ecosystem; Cluster 7 thus incorporated the actual river as well as the riparian vegetation. Aside from rivers, the most homogenous TEM site series were high elevation ecosystem types occurring in the Mountain Hemlock (MHwh1) biogeoclimatic subzone (e.g., MHwh1/01(MB), MHwh1/06(MD), MHwh1/02(MM), MHwh1/03(MR), MHwh1/00( MS), and MHwh1/09(YC). Across their full spatial extent, these site series consistently overlapped the same one to three cluster types (particularly Cluster 9). When measured in terms of counts, Simpson's diversity indices for these montane classes were also low, ranging from 0 to 0.44 (Fig. 6). Conversely,

the lower elevation forested site series such as CWHvh2/00(TS), CWHvh2/12(LS), and CWHvh2/02(LR), overlapped with many different cluster types, likely due in part to their larger extents. Despite the heterogeneous composition of the 18 clusters in terms of TEM classes and vice versa, a generalized (six-class) comparison showed greater similarity (Table 6, Fig. 7). Specifically, the islands were dominated by wet and wetland forests with low productivity (~35% of the total area). Drier forests, and especially drier non-forested ecosystems, comprised considerably less area (~15% of the total area combined). Both mapping approaches indicated that both fresh to very moist forests and shrub/herb wetland ecosystems occupy ~22%–27% of the islands.



**Fig. 3.** Zoomed in view of 18 clusters on Calvert and Hecate Islands. Clusters 1 through 12 are forested and Clusters 13 through 18 are non-forested.

#### 4. Discussion

Our analysis generated a total of 18 clusters (ecological regions) including open, shrub-herb dominated ecosystems (e.g., Clusters 13 to 18), a riparian zone (Cluster 7), wetter and wetland ecosystems (e.g., Clusters 3, 4, 6 to 8, 14 through 18), and less wet, more productive forests (e.g., Clusters 2, 11, 12). We assessed what unique ecological information is contained in these clusters derived from a variety of high spatial resolution remotely sensed data and what

information complements an expert-driven classification. For a generalized (six-class) classification, both the quantitative classification and the expert-driven classification tell a similar story. Namely, there is a high abundance of wetter and wetland ecosystems on the island, and productive forests are limited in extent.

At a more detailed level, the 18 clusters complement or augment the expert-driven map. First, whereas the TEM for our study area does not explicitly differentiate based on aspect, several of the 18 clusters were differentiated primarily by aspect. For instance,

**Table 3**

The statistical optimum number of clusters according to three different criteria.

Criterion	Method details	Optimum cluster number (forested model)	Optimum cluster number (non-forested model)
Average Silhouette Width (ASW) <sup>a</sup>	$ASW =  \sum S_i/n $ The similarity (e.g., distance) of each data point $i$ to other points in the same cluster $a$ , relative to distance to points in other clusters, $b$ averaged for all points. In an optimal solution, values are close to 1.0, meaning clusters are very homogenous and well separated.	25	3
Calinski–Harabasz (CH) <sup>a</sup>	$BSS(K-1)/WSS(n-K)$ , where $BSS$ is between group sum of squares, $WSS$ is within group sum of squares, $K$ is number of clusters and $n$ is number of data points. Optimal number of clusters maximizes this index.	2	2
Bayesian Information Criterion (BIC) <sup>b</sup>	BIC identifies the model (cluster solution) that would most likely produce the observed data, penalizing for complexity (number of clusters $\times$ number of observations). Optimal number of clusters maximizes this index (i.e., BIC will be higher for smaller numbers of clusters).	2	2

<sup>a</sup> ASW and CH were calculated in R based on  $k$ -means clustering.

<sup>b</sup> BIC was calculated in SPSS based on Two-Step clustering.

**Table 4**

Mean values of remotely sensed inputs and descriptors of 18 clusters. Clusters 1 through 12 are forested and Clusters 13 through 18 are non-forested.

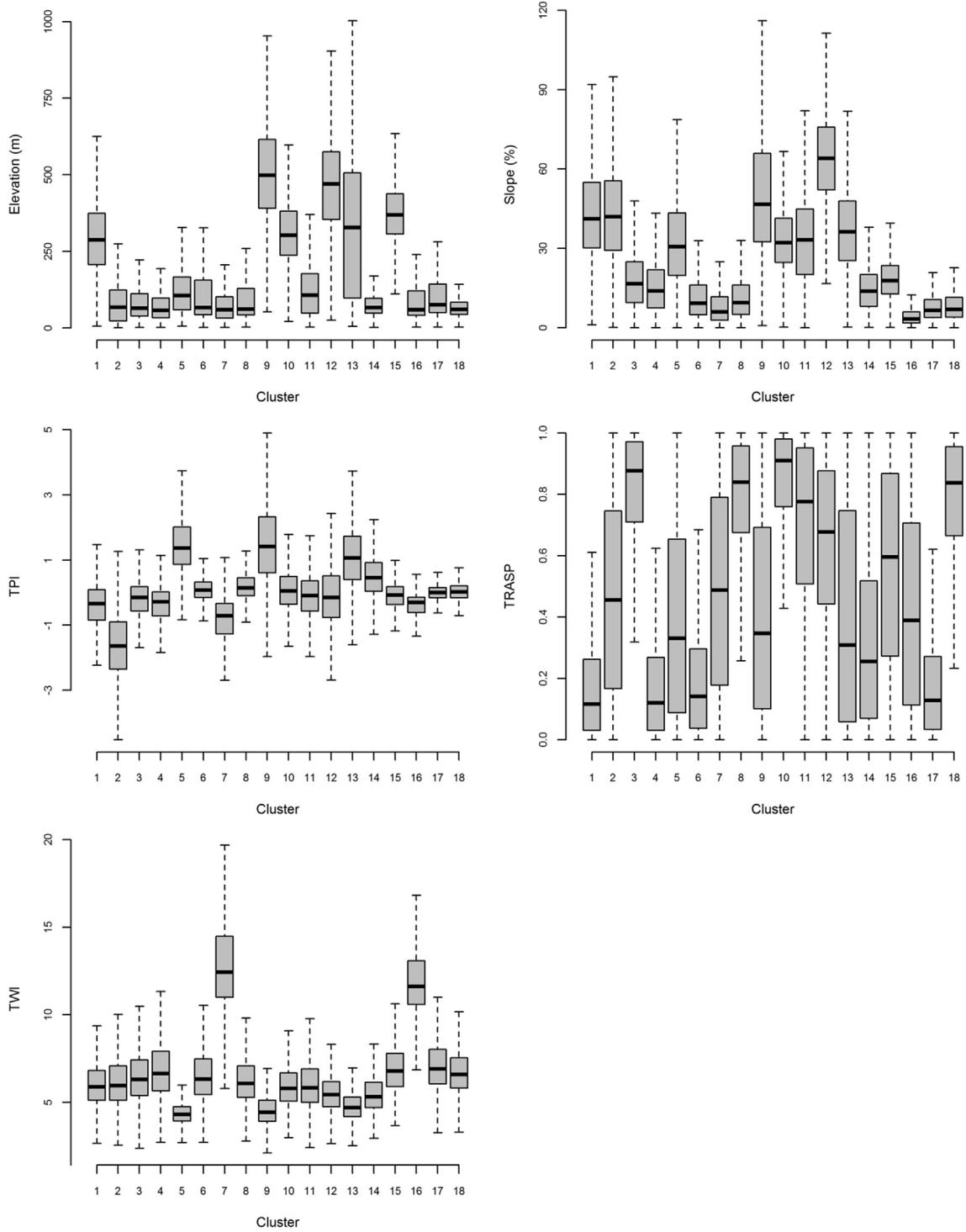
Cluster	Extent (ha)	Elevation (m)	Gap fraction	Height – coefficient of variation	Height – maximum (m)	Height – mean (m)	NDVI	Slope (%)	TPI	TRASP	TWI
1	1936	292	0.39	0.46	12.08	7.72	0.75	44	-0.41	0.16	6.07
2	1160	86	0.24	0.52	20.30	13.00	0.78	43	-1.65	0.47	6.29
3	3903	81	0.42	0.46	12.80	8.48	0.74	18	-0.20	0.83	6.52
4	3696	73	0.41	0.44	12.65	8.41	0.74	15	-0.38	0.16	6.93
5	2900	119	0.32	0.42	13.74	9.04	0.75	33	1.52	0.39	4.39
6	2908	113	0.69	0.37	5.99	4.73	0.64	12	0.08	0.18	6.67
7	1473	85	0.40	0.45	14.29	9.77	0.74	8	-0.88	0.49	12.89
8	1921	100	0.69	0.38	5.88	4.76	0.65	11	0.20	0.80	6.32
9	1020	506	0.48	0.38	7.97	5.73	0.71	51	1.55	0.41	4.59
10	2039	311	0.50	0.45	9.67	6.51	0.73	34	0.07	0.86	5.96
11	2206	120	0.16	0.36	23.38	16.61	0.79	33	-0.10	0.69	6.13
12	1326	463	0.12	0.33	21.11	15.25	0.82	65	-0.13	0.63	5.54
13	849	361	0.93	n/a	1.89	n/a	0.59	38	1.10	0.38	4.86
14	696	79	0.98	n/a	0.83	n/a	0.42	13	0.45	0.32	5.73
15	1902	413	0.95	n/a	1.60	n/a	0.57	18	-0.11	0.52	7.03
16	720	99	0.94	n/a	1.57	n/a	0.57	4	-0.44	0.45	11.95
17	2449	105	0.93	n/a	1.93	n/a	0.60	8	-0.02	0.16	7.20
18	1674	76	0.93	n/a	1.91	n/a	0.60	8	0.03	0.80	6.82

NDVI = Normalized Difference Vegetation Index (NDVI), TPI = Topographic Position Index, TRASP = Topographic Radiation ASPect, TWI = Topographic Wetness Index.

Clusters 3 and 4 were found to have similar values of elevations, height, NDVI, slope, and TWI, but are different in terms of TRASP (Table 4, Figs. 4 and 5). Likewise, Cluster 6 and 8 had similar values of elevation, slope, gap fractions, similar height, and NDVI profiles, and were predicted to be similar in topographic wetness. However, Cluster 6 was associated with cooler aspects while Cluster 8 was associated with warmer aspects. Another example is Clusters 17 and 18, which were similar with respect to the majority of remotely sensed variables but were quite different with respect to TRASP. We included TRASP as an input variable because it is known to influence ecological patterns and processes. In particular, in temperate regions, localized aspect-dependent differences in sun exposure, temperature, and moisture can lead to differences in plant growth, plant composition and structure, growth and response to disturbance (Åström, Dynesius, Hylander, & Nilsson, 2007; Diggins & Catterlin, 2014; Fekedulegn, Hicks, & Colbert, 2003; Holland & Steyn, 1975), as well as to soil microbial communities (Carletti et al., 2008).

Another unique and promising aspect of the quantitative regionalization demonstrated in this research relates to the delineation of wetland (defined here to include wet forest) ecosystems. Wetlands provide a range of important ecological processes and ecosystem services, including critical habitat, water quality and quantity regulation, and nutrient cycling and climate regulation,

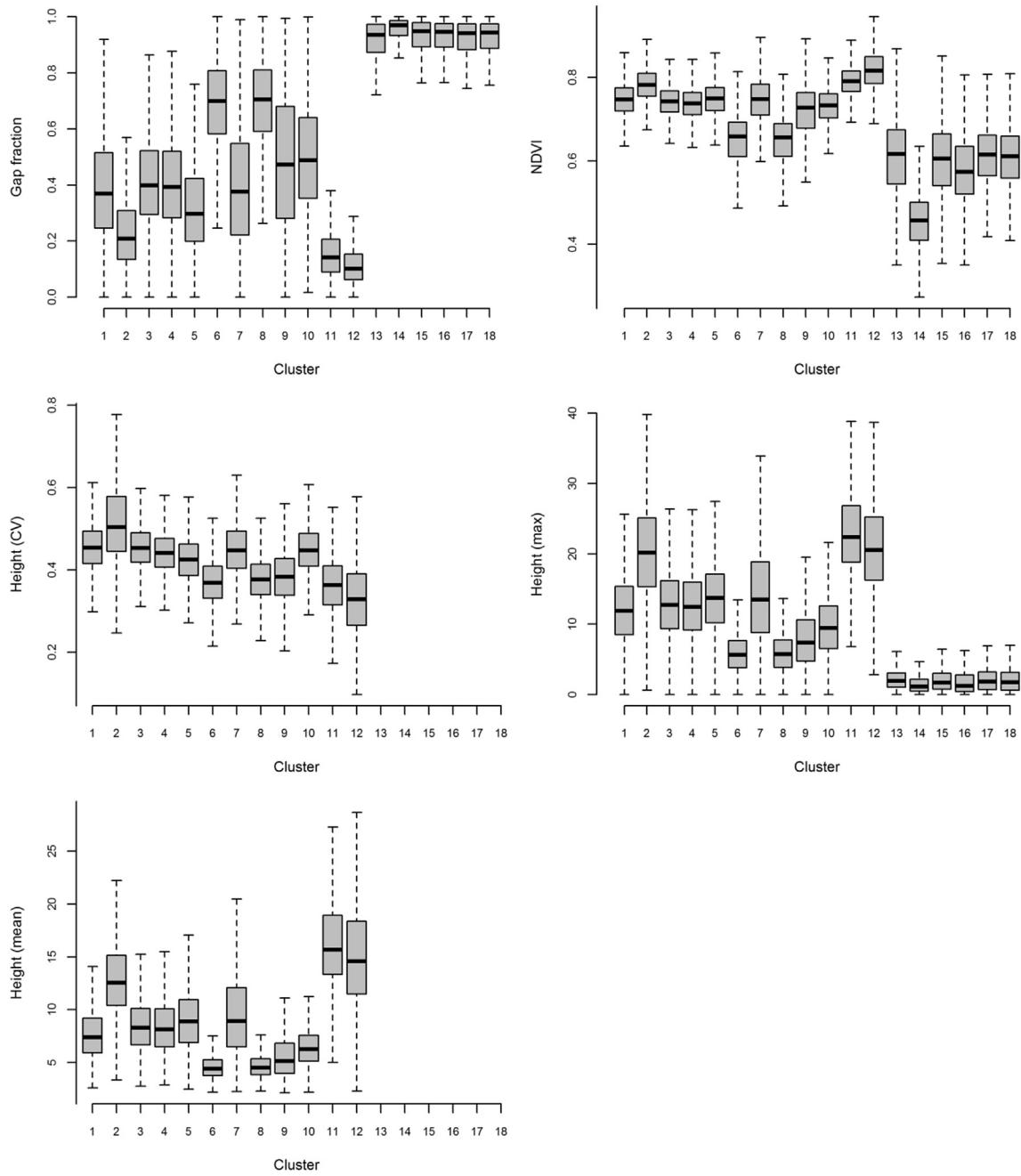
which may vary according to wetland extent, type, and location within a watershed (Brinson, 1993; Emili, 2003; Fennessy, 2014; Zedler & Kercher, 2005). Consequently, a mapping methodology that readily identifies wetland distributions at high spatial resolutions is desirable. Yet these ecosystems are often poorly mapped remotely because of their small size relative to the spatial resolution of commonly used satellite imagery such as Landsat (30 m) (Congalton, Birch, Jones, & Schriever, 2002; Ozesmi & Bauer, 2002) or relative to a standard minimum mapping unit, such as the 0.5–2 ha unit typical of TEM. Further, classification of wetlands with multispectral imagery is driven by vegetation or land cover spectral response (Dechka et al., 2002). Additional features used to classify wetlands to a particular class or site series in the field such as acidity/alkalinity and magnitude of lateral and vertical water flow (MacKenzie & Moran, 2004) are not apparent in multispectral imagery, particularly under heavy vegetation cover (Dechka et al., 2002; Rosenqvist, Finlayson, Lowry, & Taylor, 2007). Terrain indices such as the Topographic Wetness Index may be useful for wetland delineation by estimating surface water flow and accumulation but cannot capture all the hydrological properties typically used to classify wetland type in the field. In this study, we predicted the location of wetlands using terrain and vegetation structural indices generated from the LiDAR data, combined with multi-spectral imagery from RapidEye. Our study demonstrates



**Fig. 4.** Distribution of clusters across each LiDAR-derived terrain index. Clusters 1 through 12 are forested, and Clusters 13 through 18 are non-forested.

that in addition to multispectral imagery and terrain indices provided by a 20 m LiDAR-derived DEM, 3D vegetation structural information provided by the LiDAR data may also be of great value. For instance, in our study the wetland vegetation was shown to be quite distinct from less wet vegetation with regards to productivity and structure (e.g., canopy height and openness). The clusters we assigned to the broad wetland classes (Clusters 3, 4, 6, 7, 8, and 15 through 18) overlapped considerably with various types of BEC-TEM (expert-delineated) wetlands or water bodies (Fig. 7, and

A.1). Thus, although a precise ecological interpretation of these clusters is challenging in the absence on-the-ground verification, we can be reasonably confident that they do represent wetlands (e.g., Figs. 2, 3, 7 and A.2). Future research could examine the potential of using even higher spatial resolution (e.g., <5 m) LiDAR DEMs, the effect of neighbourhood sizes for metrics such as TPI, and alternative terrain metrics for generating variables that are a stronger proxy for hydrological processes, and thus for enabling more detailed wetland mapping and classification.



**Fig. 5.** Distribution of clusters across LiDAR and RapidEye vegetation data. Clusters 1 through 12 are forested, and Clusters 13 through 18 are non-forested.

A final key and uniquely informative aspect of the quantitative regionalization approach is the ability to quantify and map spatial variation in forest height and productivity at 20 m spatial resolution. Forest inventories generally do measure stand height, basal area, canopy cover, and so on, however these attributes are aggregated over larger polygons. A system in British Columbia known as Site Index by BEC Site Series (SIBEC) estimates heights at age 50 for particular tree species within each site series. The SIBEC system adds to the utility of TEM by allowing a comparison of productivity across sites, yet when mapped, is constrained by the same spatial ambiguity as is TEM, and is based on a model rather than direct measurement. Our study has shown that tree heights for the study area are quite low (most clusters have a mean height <10 m). Low tree heights were also measured and predicted for the

CWHvh2 in a nearby SIBEC study within the CWHvh2, where mean tree heights in old-growth stands ranged from 5.2 m for Western redcedar (*Thuja plicata*) in the site series 04, to 2.0 m for Western hemlock (*Tsuga heterophylla*) in site series 12 (Banner et al., 2005). Our research also indicated that canopies are generally very open (forested clusters have, on average, a canopy closure of <60%), and are variable in terms of vertical heterogeneity. Combined with NDVI, a measure of vegetation greenness, our clusters contained rich and detailed structural and productivity information.

Understanding the unique information contained within an expert-driven compared to a quantitative ecological classification using newer geographic technologies has been the subject of several contemporary studies (e.g., Thomas et al. 2002). There are several reasons for differences between an expert-driven and a

**Table 5**

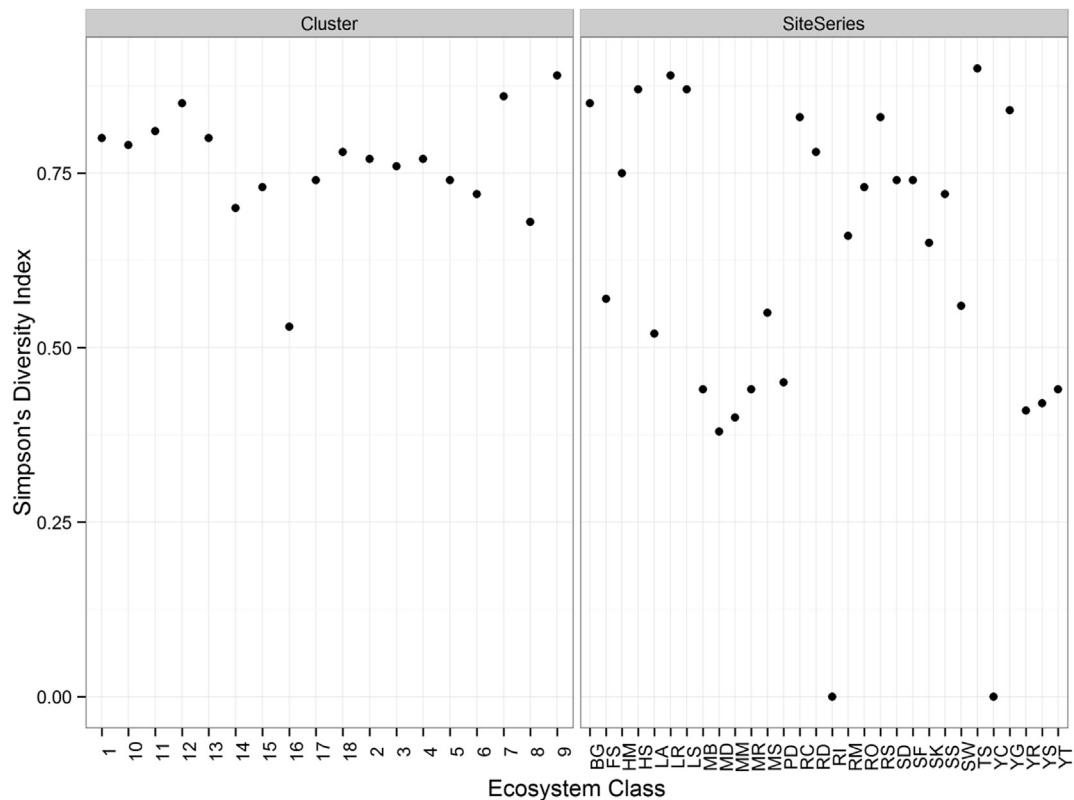
Comparison of 18 clusters to existing Terrestrial Ecosystem Mapping data.

Cluster	Dominant TEM unit <sup>a,b</sup> (% of cluster)
1	Zonal forest (36)
2	Zonal forest (49)
3	Bog forest (42)
4	Bog forest (41)
5	Cedar – salal forest (38)
6	Bog woodland (42)
7	Zonal forest (28)
8	Bog woodland (52)
9	Cedar – salal (31)
10	Bog forest (38)
11	Zonal forest (47)
12	Zonal forest (35)
13	Bedrock (52)
14	Bedrock (42), Blanket bog (42)
15	Blanket bog (65)
16	Fen (28)
17	Blanket bog (65)
18	Blanket bog (56)

<sup>a</sup> Zonal forest = CWHvh2/01(HS); Blanket bog = CWHvh2/00(TS); Bog forest = CWHvh2/11(YG); Cedar – salal forest = CWHvh2/03(RS); Bog woodland = CWHvh2/12(LS); Fen = CWHvh2/00(FS); Bedrock = RO.

<sup>b</sup> See pie charts in Fig A.1 for further details.

automatically extracted from remotely sensed imagery to approximate but not replicate some of these same visual cues for ecosystem classification, including spectral reflectance, image texture, shape information derived from object-based classifiers, as well as landscape pattern indices. With high spatial resolution LiDAR data, an additional dimension of data is available for interpretation that is not present in standard multispectral satellite imagery. The three-dimensional forest structure provided by LiDAR and the high spatial resolution terrain models providing a range of ecological mapping applications (Lefsky, Cohen, Parker, & David, 2002; Vierling, Vierling, Gould, Martinuzzi, & Clawges, 2008). Through this study we aim to demonstrate the utility of a quantitative, unsupervised remote-sensing based regionalization, and to assess the unique and complementary information content of mapping, relative to conventional expert-driven ecosystem mapping. Future work should focus on the use of high-spatial resolution remotely sensed data (especially LiDAR) in a supervised classification approach (e.g., Predictive Ecosystem Mapping (MacMillan et al. 2007)) to assess the ability to model and map ecosystem distributions, including BEC site series. Such an approach should also focus on incorporating data and metrics chosen to capture compositional attributes of ecosystems. In particular, it may be



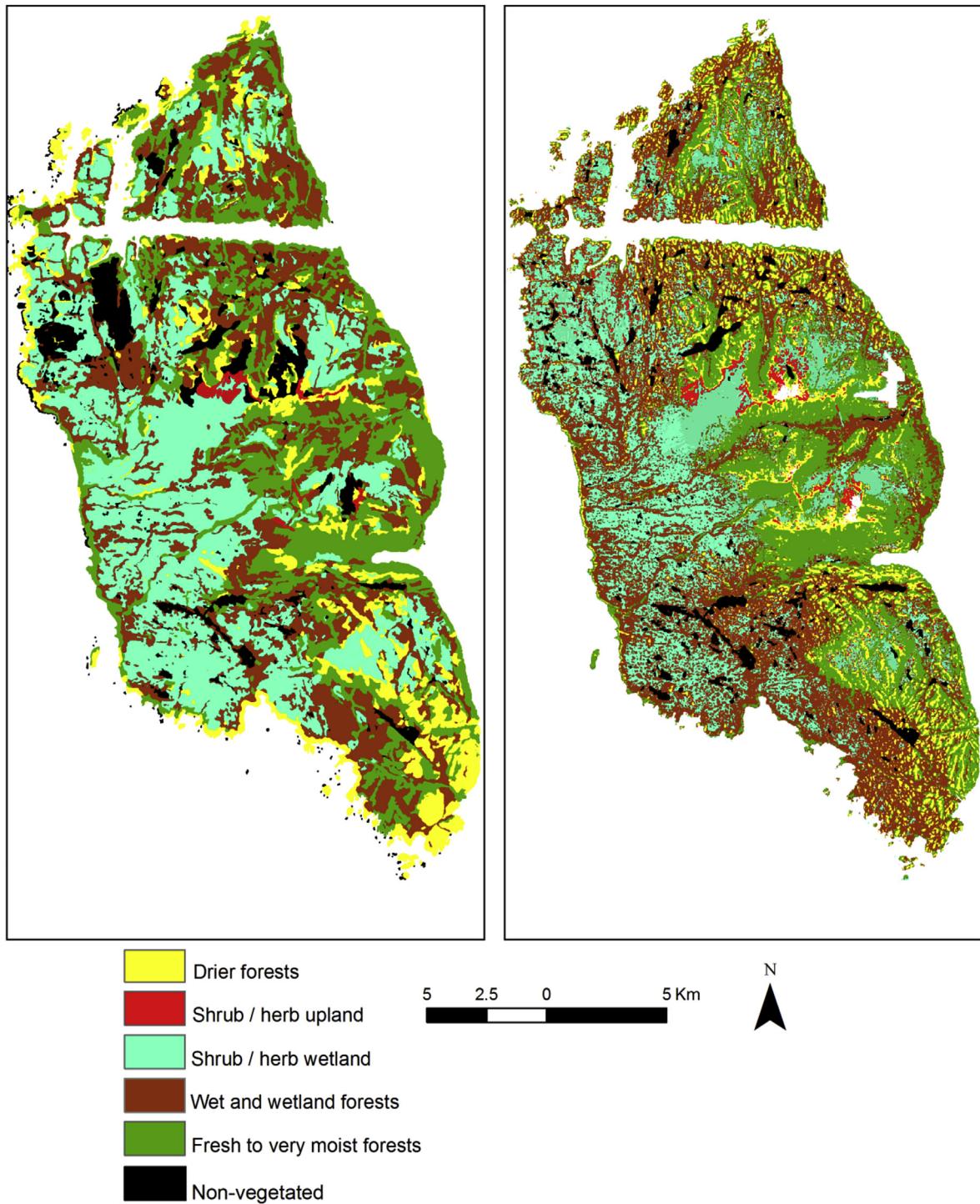
**Fig. 6.** Diversity of TEM site series (with respect to cluster composition) and diversity of clusters (with respect to TEM site series). Lower values of the Simpson's Diversity Index represent greater similarity between TEM and cluster types.

quantitative remotely sensed classification, including the fact that expert systems may be better at separating ecologically meaningful information from within large amounts of (non-ecologically meaningful) variation (Schmidlein, Tichý, Feilhauer, & Faude, 2010). Experts manually interpreting aerial photos for ecosystem mapping inherently incorporate multiple characteristics such as tone or colour, shape, size, pattern, texture, shadow, and landscape context (Morgan et al., 2010). Various types of information can be

useful to incorporate hyperspectral remotely sensed data, which is well suited to mapping vegetation composition (e.g., Jones et al. 2010).

## 5. Conclusion

This study has demonstrated that LiDAR and high spatial resolution multi-spectral imagery can be combined using a quantitative



**Fig. 7.** Generalized ecosystem classes as depicted by the expert-based classification (Terrestrial Ecosystem Map) in the left panel, and the unsupervised regionalization on the right.

regionalization to map a variety of ecosystem types across a heterogeneous landscape. Given the quantitative nature of our approach, it is transparent and easily repeatable, and likely to succeed in other regions where topography is subdued and forest structure complex. The ecosystem classes highlight vegetation structure and productivity, refined by topography, and may have a variety of applications beyond—yet complementary to—those provided by expert-based ecosystem mapping from aerial photography. For instance, we have identified clusters that are

attributed with quantitative vegetation height information (mean height and variation in mean height) and are spatially explicit. Height information can be used for carbon accounting, because carbon storage in vegetation can be estimated through allometric equations that use variables such as tree height to estimate biomass. As well, estimates regarding relative primary productivity are readily available in our dataset because of the NDVI incorporated into the clustering. Productivity is an essential ecosystem function that supports a range of ecosystem services; while

**Table 6**

Extent of generalized ecosystem classes on Calvert and Hecate Islands, British Columbia, using an unsupervised classification of remotely sensed data (cluster analysis) and an expert-driven classification (Terrestrial Ecosystem Mapping).

General ecosystem class	Corresponding clusters	Corresponding BEC units present in study area TEM <sup>a</sup>	Percentage of study area (clusters)	Percentage of study area (TEM)
Shrub/herb upland	13	MHwh1/00(MS) CWHvh2/00(SA)	1.9	0.4
Shrub/herb wetland	14, 15, 16, 17, 18	CWHvh2/00(TS) CWHvh2/00(BG) CWHvh2/00(FS) MHwh1/00(TS) CWvh2/00(WI)	27.2	21.5
Wetter and wetland forests	3, 4, 6, 7, 8	CWHvh2/00(RH) CWHvh2/11(YG) CWHvh2/12(LS) CWHvh2/13(RC) MHwh1/06(MD) MHwh1/08(YS) MHwh1/09(YC)	34.7	34.4
Drier forests	5, 9	CWHvh2/00(RM) CWHvh2/02(LR) CWHvh2/03(RS) CWHvh2/14(SS) CWHvh2/15(SK) CWHvh2/16(SR) MHwh1/00(YR) MHwh1/02(MM)	12.3	14.9
Fresh to very moist forests	1, 2, 10, 11, 12	CWHvh2/00(RD) CWHvh2/01(HS) CWHvh2/04(HM) CWHvh2/06(SF) CWHvh2/07(SD) CWHvh2/17(SW) CWHvh2/18(SE) MHwh1/01(MB) MHwh1/03(MR) MHwh1/05(YT)	23.9	22.4
Non-vegetated	Excluded from analysis to extent possible	Bedrock (RO), rivers (RI), ponds (PD), lakes (LA), shallow open water (OW), exposed soil (ES)	n/a	6.2

<sup>a</sup> MHwh1 = Mountain Hemlock zone (Wet Hypermaritime Subzone); CWHvh2 = Coastal Western Hemlock zone (Very Wet Hypermaritime Subzone, Central variant). The two-digit numbers (and associated two-letter acronyms) following the forward slash refer to Site Series; those with numbers 00 are generally non-forested or not formalized in the current version of the provincial classification for the area (Banner et al., 1993; Green & Klinka, 1994; MacKenzie & Moran, 2004).

approximations of site productivity (or potential) within traditional mapped polygons can be made based on species composition, our methodology may facilitate monitoring because of the increased spatial precision with which it is estimated. As LiDAR and high-spatial resolution multispectral imagery become increasingly available in the future, they will be important complementary datasets in local-scale ecosystem mapping.

## Acknowledgements

We are grateful to the Tula Foundation for supporting this research. Additional funding was provided by the Natural Sciences and Engineering Research Council of Canada (PGS-D and Discovery grants). The LiDAR data were acquired by Terra Remote Sensing Inc. for the Hakai Institute. Thank you to Keith Holmes and Luba Reshitnyk of the Hakai Institute, members of the Coastal Ecosystem Dune Dynamics (CEDD) lab at the University of Victoria, and Richard Fournier (Université de Sherbrooke) for additional data processing, preparation, and management. Will MacKenzie (Provincial BEC Ecologist) summarized the ClimateBC data for the BEC subzones. Finally, thank you to Ken Lertzman (Simon Fraser University), Trevor Lantz (University of Victoria), Nicholas Coops (University of British Columbia), and Mike Wulder (Natural Resources Canada) for valuable comments and discussions.

## Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.apgeog.2016.02.002>.

## References

- Åström, M., Dynesius, M., Hylander, K., & Nilsson, C. (2007). Slope aspect modifies community responses to clear-cutting in boreal forests. *Ecology*, 88(3), 749–758. <http://doi.org/10.1890/06-0613>.
- Bailey, R. G. (1987). Suggested hierarchy of criteria for multi-scale ecosystem mapping. *Landscape and Urban Planning*, 14, 313–319.
- Banner, A., LePage, P., Moran, J., & de Groot, A. (2005). *The Hyp3 project: Pattern, process and productivity in hypermaritime forests of coastal British Columbia*.
- Banner, A., MacKenzie, W., Haeussler, S., Thomson, S., Pojar, J., & Trowbridge, R. (1993). *A field guide to site identification and interpretation for the Prince Rupert Forest Region*. Land Manage. Handb. 26. BC Min. For., Victoria, BC.
- Banner, A., Meidinger, D. V., Lea, E. C., Maxwell, R. E., & Sacken, B. C. (1996). Ecosystem mapping methods for British Columbia. *Environmental Monitoring and Assessment*, 39(1–3), 97–117. <http://doi.org/10.1007/BF00396139>.
- Barron, O. V., Emelyanova, I., Van Niel, T. G., Pollock, D., & Hodgson, G. (2014). Mapping groundwater-dependent ecosystems using remote sensing measures of vegetation and moisture dynamics. *Hydrological Processes*, 28(2), 372–385. <http://doi.org/10.1002/hyp.9609>.
- BC Ministry of Forests and Range and BC Ministry of Environment. (2010). Field manual for describing Terrestrial ecosystems. In *Land management handbook* (2nd ed., Vol. 25) (Victoria, BC).
- Brinson, M. (1993). Changes in the functioning of wetlands along environmental gradients. *Wetlands*, 13(2), 65–74. <http://doi.org/10.1007/BF03160866>.
- Brosofske, K. D., Froese, R. E., Falkowski, M. J., & Banskota, A. (2014). A review of methods for mapping and prediction of inventory attributes for operational forest management. *Forest Science*, 60(4), 733–756. <http://dx.doi.org/10.5849/forsci.12-134>.

- Bryan, B. A. (2006). Synergistic techniques for better understanding and classifying the environmental structure of landscapes. *Environmental Management*, 37(1), 126–140. <http://doi.org/10.1007/s00267-004-0058-1>.
- Calinski, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, 3(1), 1–27.
- Canadian Forest Service. (2001). *Canada's National Forest Inventory - National standards for photo plots: Compilation procedures (Vol. Version 1)*. Victoria, BC: Natural Resources Canada.
- Carletti, P., Vendramin, E., Pizzighello, D., Concheri, G., Zanella, A., Nardi, S., et al. (2008). Soil humic compounds and microbial communities in six spruce forests as function of parent material, slope aspect and stand age. *Plant and Soil*, 315(1–2), 47–65. <http://doi.org/10.1007/s11104-008-9732-z>.
- Chastain, R. J., & Struckhoff, M. (2008). Mapping vegetation communities using statistical data fusion in the Ozark National Scenic Riverways, Missouri, USA. *Photogrammetric Engineering and Remote Sensing*, 74(2), 247–264. <http://doi.org/> <http://dx.doi.org/10.14358/PERS.74.2.247>.
- Chavez, P. (1996). Image-based atmospheric corrections-revisited and improved. *Photogrammetric Engineering and Remote Sensing*, 62(9), 1025–1036. <http://doi.org/> [http://10.1016/S0168-1699\(02\)00108-4](http://10.1016/S0168-1699(02)00108-4).
- Congalton, R., Birch, K., Jones, R., & Schriever, J. (2002). Evaluating remotely sensed techniques for mapping riparian vegetation. *Computers and Electronics in Agriculture*, 37, 113–126. [http://doi.org/10.1016/S0168-1699\(02\)00108-4](http://doi.org/10.1016/S0168-1699(02)00108-4).
- Coops, N. C., Hilker, T., Wulder, M. A., St-Onge, B., Newnham, G., Siggins, A., et al. (2007). Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees*, 21(3), 295–310. <http://doi.org/10.1007/s00468-006-0119-6>.
- 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(1), 8–22. <http://doi.org/10.1016/j.ecoinf.2008.09.005>.
- Dalponte, M., Orka, H. O., Gobakken, T., Gianelle, D., & Naesset, E. (2013). Tree species classification in boreal forests with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 51(5), 2632–2645. <http://dx.doi.org/10.1109/TGRS.2012.2216272>.
- Dechka, J., Franklin, S., Watmough, M., Bennett, R., & Ingstrup, D. (2002). Classification of wetland habitat and vegetation communities using multi-temporal Ikonos imagery in southern Saskatchewan. *Canadian Journal of Remote Sensing*, 28(5), 679–685. <http://doi.org/10.5589/m02-064>.
- Diggins, T. P., & Catterlin, R. G. (2014). Topographic patterns in forest composition and diversity on slopes of Zoar Valley Canyon, western New York. *Northeastern Naturalist*, 21(3), 337–350. <http://doi.org/> <http://dx.doi.org/10.1656/045.021.0301>.
- Dobrowski, S. Z., Safford, H. D., Cheng, Y. B., & Ustin, S. L. (2008). Mapping mountain vegetation using species distribution modeling, image-based texture analysis, and object-based classification. *Applied Vegetation Science*, 11(4), 499–508. <http://doi.org/10.3170/2008-7-18560>.
- Emili, L. (2003). *Hydrochemical characteristics of hypermaritime forest-peatland complexes. North Coast British Columbia*: University of Waterloo (PhD thesis).
- Fekedulegn, D., Hicks, R. R., & Colbert, J. J. (2003). Influence of topographic aspect, precipitation and drought on radial growth of four major tree species in an Appalachian watershed. *Forest Ecology and Management*, 177(1–3), 409–425. [http://doi.org/10.1016/S0378-1127\(02\)00446-2](http://doi.org/10.1016/S0378-1127(02)00446-2).
- Fennessy, M. S. (2014). Wetland ecosystems and global change. In B. Freedman (Ed.), *Global environmental change* (pp. 255–261). Netherlands: Springer. <http://doi.org/10.1007/978-94-007-5784-4>.
- Feret, J., & Asner, G. P. (2013). Tree species discrimination in tropical forests using airborne imaging spectroscopy. *IEEE Transactions on Geoscience and Remote Sensing*, 51(1), 73–84. <http://dx.doi.org/10.1109/TGRS.2012.2199323>.
- Fitterer, J. L., Nelson, T. A., Coops, N. C., & Wulder, M. A. (2012). Modelling the ecosystem indicators of British Columbia using earth observation data and terrain indices. *Ecological Indicators*, 20, 151–162. <http://doi.org/10.1016/j.ecolind.2012.02.024>.
- Franklin, J. (1995). Predictive vegetation mapping: geographic modelling of biogeographical patterns in relation to environmental gradients. *Progress in Physical Geography*, 19(4), 474–499. <http://doi.org/10.1177/03091339501900403>.
- Franklin, J. (2013). Mapping vegetation from landscape to regional scales. In E. van der Maarel, & J. Franklin (Eds.), *Vegetation ecology* (2nd ed.). John Wiley & Sons, Ltd.
- Fraser, R., McLennan, D., Ponomarenko, S., & Olthof, I. (2012). Image-based predictive ecosystem mapping in Canadian arctic parks. *International Journal of Applied Earth Observation and Geoinformation*, 14(1), 129–138. <http://doi.org/10.1016/j.jag.2011.08.013>.
- Gillin, C. P., Bailey, S. W., McGuire, K. J., & Prisley, S. P. (2015). Evaluation of lidar-derived DEMs through terrain analysis and field comparison. *Photogrammetric Engineering & Remote Sensing*, 81(5), 387–396. <http://dx.doi.org/10.14358/PERS.81.5.387>.
- Goward, S. N., Tucker, C. J., & Dye, D. G. (1985). North American vegetation patterns observed with the NOAA-7 advanced very high resolution radiometer. *Vegetatio*, 64(1), 3–14. <http://doi.org/10.1007/BF00033449>.
- Green, R. (2014). *Reconnaissance level terrestrial ecosystem mapping of priority landscape units of the coast EBM planning area: Phase 3* (Prepared for BC Min. Forests, Lands & Natural Resource Ops., by B.A. Blackwell & Associates, Vancouver, BC).
- Green, R., & Klinka, K. (1994). *A field guide to site identification and interpretation for the Vancouver Forest Region*. Victoria, BC: BC Ministry of Forests Research Program.
- Hancock, R. N., & Csillag, F. (2002). Ecoregionalization assessment: spatio-temporal analysis of net primary production across Ontario. *Ecoscience*, 9(2), 219–230.
- Hargrove, W. W., & Hoffman, F. M. (2004). Potential of multivariate quantitative methods for delineation and visualization of ecoregions. *Environmental Management*, 34(Suppl 1), S39–S60. <http://doi.org/10.1007/s00267-003-1084-0>.
- Hogg, A. R., & Holland, J. (2008). An evaluation of DEMs derived from LiDAR and photogrammetry for wetland mapping. *Forestry Chronicle*, 84(6), 840–849. <http://doi.org/10.5558/tfc84840-6>.
- Holland, P. G., & Steyn, D. G. (1975). Vegetational responses to latitudinal variations in slope angle and aspect. *Journal of Biogeography*, 2(3), 179–183.
- Holmgren, J. (2004). Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. *Scandinavian Journal of Forest Research*, 19(6), 543–553. <http://doi.org/10.1080/02827580410019472>.
- Hopkinson, C., & Chasmer, L. (2009). Testing LiDAR models of fractional cover across multiple forest ecozones. *Remote Sensing of Environment*, 113(1), 275–288. <http://doi.org/10.1016/j.rse.2008.09.012>.
- Johansen, K., Coops, N. C., Gergel, S. E., & Stange, Y. (2007). Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. *Remote Sensing of Environment*, 110(1), 29–44. <http://doi.org/10.1016/j.rse.2007.02.014>.
- Jones, T. G., Coops, N. C., & Sharma, T. (2010). Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. *Remote Sensing of Environment*, 114(12), 2841–2852. <http://doi.org/10.1016/j.rse.2010.07.002>.
- Kent, M., Gill, W., Weaver, R., & Armitage, R. (1997). Landscape and plant community boundaries in biogeography. *Progress in Physical Geography*, 21(3), 315–335. <http://doi.org/10.1177/03091339702100301>.
- Ke, Y., Quackenbush, L. J., & Im, J. (2010). Synergistic use of QuickBird multispectral imagery and LiDAR data for object-based forest species classification. *Remote Sensing of Environment*, 114(6), 1141–1154. <http://dx.doi.org/10.1016/j.rse.2010.01.002>.
- Knight, J. F., Tolcser, B. P., Corcoran, J. M., & Rampi, L. P. (2013). The effects of data selection and thematic detail on the accuracy of high spatial resolution wetland classifications. *Photogrammetric Engineering & Remote Sensing*, 79(7), 613–623. <http://doi.org/> <http://dx.doi.org/10.14358/PERS.79.7.613>.
- Kupfer, J. A., Gao, P., & Guo, D. (2012). Regionalization of forest pattern metrics for the continental United States using contiguity constrained clustering and partitioning. *Ecological Informatics*, 9, 11–18. <http://doi.org/10.1016/j.ecoinf.2012.02.001>.
- Leathwick, J. R., Overton, J. M., & McLeod, M. (2003). An environmental domain classification of New Zealand and its use as a tool. *Conservation Biology*, 17(6), 1612–1623. <http://doi.org/10.1111/j.1523-1739.2003.00469.x>.
- Van Leeuwen, M., & Nieuwenhuis, M. (2010). Retrieval of forest structural parameters using LiDAR remote sensing. *European Journal of Forest Research*, 129(4), 749–770. <http://doi.org/10.1007/s10342-010-0381-4>.
- Lefsky, M. A., Cohen, W. B., Acker, S. A., Parker, G. G., Spies, T., & Harding, D. (1999). Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment*, 70(3), 339–361. [http://doi.org/10.1016/S0034-4257\(99\)00052-8](http://doi.org/10.1016/S0034-4257(99)00052-8).
- Lefsky, M. A., Cohen, W. B., Parker, G. G., & David, J. (2002). Lidar remote sensing for ecosystem studies. *BioScience*, 52(1), 19–30. [http://doi.org/10.1641/0006-3568\(2002\)052\[0019:LRSSES\]2.0.CO;2](http://doi.org/10.1641/0006-3568(2002)052[0019:LRSSES]2.0.CO;2).
- Lim, K., Treitz, P., Wulder, M. A., St-Onge, B., & Flood, M. (2003). LiDAR remote sensing of forest structure. *Progress in Physical Geography*, 27(1), 88–106. <http://doi.org/10.1191/030913303pp360ra>.
- Long, J., Nelson, T. A., & Wulder, M. A. (2010). Regionalization of landscape pattern indices using multivariate cluster analysis. *Environmental Management*, 46(1), 134–142. <http://doi.org/10.1007/s00267-010-9510-6>.
- Loveland, T. R., & Merchant, J. M. (2004). Ecoregions and ecoregionalization: geographical and ecological perspectives. *Environmental Management*, 34(S1), S1–S13. <http://doi.org/10.1007/s00267-003-5181-x>.
- MacKenzie, W., & Moran, J. (2004). *Wetlands of British Columbia: A guide to identification*. In *Land management handbook* (Vol. 52). Victoria, BC: BC Ministry of Forests Research Branch.
- Mackey, B. G., Berry, S. L., & Brown, T. (2007). Reconciling approaches to biogeographical regionalization: a systematic and generic framework examined with a case study of the Australian continent. *Journal of Biogeography*, 35(2), 213–229. <http://doi.org/10.1111/j.1365-2699.2007.01822.x>.
- MacMillan, R. A., Moon, D. E., & Coupé, R. A. (2007). Automated predictive ecological mapping in a forest region of B.C., Canada, 2001–2005. *Geoderma*, 140(4), 353–373. <http://doi.org/10.1016/j.geoderma.2007.04.027>.
- Magnussen, S., & Boudevyn, P. (1998). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research*, 28(7), 1016–1031. <http://doi.org/10.1139/x98-078>.
- Maxa, M., & Bolstad, P. (2009). Mapping northern wetlands with high resolution satellite images and LiDAR. *Wetlands*, 29(1), 248–260. <http://doi.org/10.1672/08-911>.
- McMahon, G., Wiken, E. B., & Gauthier, D. A. (2004). Toward a scientifically rigorous basis for developing mapped ecological regions. *Environmental Management*, 34(S1), S111–S124. <http://doi.org/10.1007/s00267-004-0170-2>.
- Morgan, J. L., Gergel, S. E., & Coops, N. C. (2010). Aerial photography: a rapidly evolving tool for ecological management. *BioScience*, 60(1), 47–59. <http://doi.org/10.1525/bio.2010.60.1.9>.
- Niesterowicz, J., & Stepinski, T. F. (2013). Regionalization of multi-categorical

- landscapes using machine vision methods. *Applied Geography*, 45, 250–258. <http://dx.doi.org/10.1016/j.apgeog.2013.09.023>.
- Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80(1), 88–99. [http://doi.org/10.1016/S0034-4257\(01\)00290-5](http://doi.org/10.1016/S0034-4257(01)00290-5).
- Olstad, T. A. (2012). Understanding the science and art of ecoregionalization. *The Professional Geographer*, 64(2), 303–308. <http://doi.org/10.1080/00330124.2011.603656>.
- Ozesmi, S. L., & Bauer, M. E. (2002). Satellite remote sensing of wetlands. *Wetlands Ecology and Management*, 10, 381–402. <http://doi.org/10.1023/A:1020908432489>.
- Pojar, J., Klinka, K., & Meidinger, D. (1987). Biogeoclimatic ecosystem classification in British Columbia. *Forest Ecology and Management*, 22, 119–154.
- Powers, R. P., Coops, N. C., Morgan, J. L., Wulder, M. A., Nelson, T. A., Drever, C. R., et al. (2012). A remote sensing approach to biodiversity assessment and regionalization of the Canadian boreal forest. *Progress in Physical Geography*, 37(1), 36–62. <http://doi.org/10.1177/0309133312457405>.
- Resource Information Management Branch. (2005). Alberta vegetation inventory interpretation standards. Version 2.1.1. In *Chapter 3 – Vegetation inventory standards and data model documents*. Edmonton, AB: Alberta Sustainable Resource Development.
- Resources Inventory Committee. (1998). *Standard for terrestrial ecosystem mapping in British Columbia*. Province of British Columbia: Ecosystems Working Group, Terrestrial Ecosystems Task Force.
- De Reu, J., Bourgeois, J., Bats, M., Zwertyaegher, A., Gelorini, V., De Smedt, P., ... Cromb  , P. (2013). Application of the topographic position index to heterogeneous landscapes. *Geomorphology*, 186, 39–49. <http://doi.org/10.1016/j.geomorph.2012.12.015>.
- Roberts, D. W., & Cooper, S. V. (1989). Concepts and techniques of vegetation mapping. In *Land classifications based on vegetation: Applications for resource management* (pp. 90–96). Ogden, UT: USDA Forest Service GTR INT-257.
- Rosenqvist, A., Finlayson, C., Lowry, J., & Taylor, D. (2007). The potential of long-wavelength satellite-borne radar to support implementation of the Ramsar Wetlands Convention. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 17, 229–244. <http://doi.org/10.1002/aqc>.
- Rouse, J. W. J., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the great plains with ERTS. In *Third earth resources technology satellite symposium*. NASA SP-351 (pp. 309–317).
- Sayre, R., Comer, P., Harumi, W., & Cress, J. (2009). A new map of standardized terrestrial ecosystems of the conterminous United States. *US Geological Survey Professional Paper*, 1768.
- Schmidlein, S., Tichy, L., Feilhauer, H., & Faude, U. (2010). A brute-force approach to vegetation classification. *Journal of Vegetation Science*, 21(6), 1162–1171. <http://doi.org/10.1111/j.1654-1103.2010.01221.x>.
- Tagil, S., & Jenness, J. S. (2008). GIS-based automated landform classification and topographic, landcover and geologic attributes of landforms around the Yazoren Polje, Turkey. *Journal of Applied Sciences*, 8(6), 910–921.
- Tarboton, G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33(2), 309–319. <http://doi.org/10.1029/96WR03137>.
- Thomas, V., Treitz, P., Jelinski, D., Miller, J., Lafleur, P., & McCaughey, J. H. (2002). Image classification of a northern peatland complex using spectral and plant community data. *Remote Sensing of Environment*, 84, 83–99. [http://doi.org/http://dx.doi.org/10.1016/S0034-4257\(02\)00099-8](http://doi.org/http://dx.doi.org/10.1016/S0034-4257(02)00099-8).
- Valeria, O., Laamrani, A., & Beaudoin, A. (2014). Monitoring the state of a large boreal forest region in eastern Canada through the use of multitemporal classified satellite imagery. *Canadian Journal of Remote Sensing*, 38(1), 91–108. <http://doi.org/10.5589/m12-014>.
- Vierling, K. T., Vierling, L. A., Gould, W. A., Martinuzzi, S., & Clawges, R. M. (2008). Lidar: shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment*, 6(2), 90–98. <http://doi.org/10.1890/070001>.
- Wang, T., Hamann, A., Spittlehouse, D. L., & Murdoch, T. (2012). ClimateWNA – high-resolution spatial climate data for western North America. *Journal of Applied Meteorology and Climatology*, 51, 16–29. <http://doi.org/http://dx.doi.org/10.1175/JAMC-D-11-043.1>.
- Weiss, A. (2001). Topographic position and landforms analysis. *Poster Presentation, ESRI User Conference, San Diego, CA*.
- White, D. C., Lewis, M. M., Green, G., & Gotch, T. B. (2015). A generalizable NDVI-based wetland delineation indicator for remote monitoring of groundwater flows in the Australian Great Artesian Basin. *Ecological Indicators*, 1–12. <http://doi.org/10.1016/j.ecolind.2015.01.032>.
- Whittaker, R. (1967). Gradient analysis of vegetation. *Biological Reviews*, 49, 207–2684.
- Wulder, M. A., Han, T., White, J. C., Sweda, T., & Tsuzuki, H. (2007). Integrating profiling LIDAR with Landsat data for regional boreal forest canopy attribute estimation and change characterization. *Remote Sensing of Environment*, 110(1), 123–137. <http://dx.doi.org/10.1016/j.rse.2007.02.002>.
- Wulder, M. A., Skakun, R. S., Kurz, W. A., & White, J. C. (2004). Estimating time since forest harvest using segmented Landsat ETM+ imagery. *Remote Sensing of Environment*, 93(1–2), 179–187. <http://doi.org/10.1016/j.rse.2004.07.009>.
- Wulder, M. A., White, J. C., Luther, J. E., Strickland, G., Remmell, T. K., & Mitchell, S. W. (2006). Use of vector polygons for the accuracy assessment of pixel-based land cover maps. *Canadian Journal of Remote Sensing*, 32(3), 268–279. <http://doi.org/10.5589/m06-023>.
- Xu, C., Sheng, S., Chi, T., Yang, X., An, S., & Liu, M. (2014). Developing a quantitative landscape regionalization framework integrating driving factors and response attributes of landscapes. *Landscape and Ecological Engineering*, 10(2), 295–307. <http://dx.doi.org/10.1007/s11355-013-0225-8>.
- Zedler, J. B., & Kercher, S. (2005). Wetland resources: status, trends, ecosystem services, and restorability. *Annual Review of Environment and Resources*, 30(1), 39–74. <http://doi.org/10.1146/annurev.energy.30.050504.144248>.