Mapping and Monitoring Indicators of Terrestrial Biodiversity with Remote Sensing

by

Shanley Dawn Thompson
BSc, Queen’s University, 2006
MSc, University of British Columbia, 2008

A Dissertation Submitted in Partial Fulfillment of the

Requirements of the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Geography

University of Victoria

Supervisory Committee

Mapping and Monitoring Indicators of Terrestrial Biodiversity with Remote Sensing

by

Shanley Dawn Thompson
BSc, Queen's University, 2006
MSc, University of British Columbia, 2008

Supervisory Committee

Dr. Trisalyn A. Nelson
(Department of Geography, University of Victoria)
Supervisor

Dr. Michael A. Wulder
(Department of Geography, University of Victoria, and
Natural Resources Canada (Canadian Forest Service - Pacific Forestry Centre))
Departmental Member

Dr. Trevor C. Lantz
(Department of Environmental Studies, University of Victoria)
Outside Member

Dr. Nicholas C. Coops
(Department of Forest Resources Management, University of British Columbia)
Additional Member

Abstract

Supervisory Committee

Dr. Trisalyn A. Nelson
(Department of Geography, University of Victoria)
Supervisor

Dr. Michael A. Wulder
(Department of Geography, University of Victoria, and
Natural Resources Canada (Canadian Forest Service - Pacific Forestry Centre))
Departmental Member

Dr. Trevor C. Lantz
(Department of Environmental Studies, University of Victoria)
Outside Member

Dr. Nicholas C. Coops
(Department of Forest Resources Management, University of British Columbia)
Additional Member

Biodiversity is a complex concept incorporating genes, species, ecosystems, composition, structure and function. The global scientific and political community has recognized the importance of biodiversity for human well-being, and has set goals and targets for its conservation, sustainable use, and benefit sharing. Monitoring biodiversity will help meet conservation goals and targets, yet observations collected *in-situ* are limited in space and time, which may bias interpretations and hinder conservation. Remote sensing can provide complementary datasets for monitoring biodiversity that are spatially comprehensive and repeatable. However, further research is needed to demonstrate, for various spatial scales and regions, how remotely sensed datasets represent different aspects of biodiversity. The overall goal of this dissertation is to advance the mapping and monitoring of biodiversity indicators, globally and within Canada, through the use of remote sensing. This research produced maps that were rich with spatially explicit, spatially continuous data, filling gaps in the availability and spatial resolution or scalability of information regarding ecosystem function (primary productivity) at global scales, tree species composition at regional scales (Saskatchewan, Canada), and ecosystem structure at local scales (coastal British Columbia, Canada). Further, the remotely sensed indicator datasets proposed and tested in each of the research chapters are repeatable, ecologically meaningful, translate to specific biodiversity targets globally and within Canada, and are calculable at multiple spatial scales. Challenges and opportunities for fully implementing these or similar remotely sensed biodiversity indicators and indicator datasets at a national level in Canada are discussed, contributing to the advancement of biodiversity monitoring science.

Co-Authorship Statement

Chapters 2 through 5 were co-authored, with research, analysis and writing lead by Shanley Thompson. Co-authors were involved in project scoping, and provided analytical guidance and editorial comments. Chapter 3 was proposed by Michael Wulder and Joanne White, who also provided the Landsat Best-Available-Pixel (BAP) composite data for the analyses. In Chapter 4, Ken Lertzman (Simon Fraser University) was also involved in project scoping. All data were provided by the Hakai Institute. LiDAR metrics used in Chapter 4 were calculated by Gordon Frazer, with additional remotely sensed data processed by the Hakai Institute. Nicholas Coops and Trevor Lantz provided helpful suggestions on an earlier manuscript draft. Michael Wulder and Nicholas Coops were particularly instrumental in the conception of Chapter 5, with input from all co-authors. A global MODIS fPAR mosaiced dataset used for calculating the fPAR metrics was provided to Nicholas Coops by Steve Running (University of Montana) and Maosheng Zhao (University of Maryland).

A version of Chapter 2 has been submitted for publication as:

Thompson, S.D., Nelson, T.A., Wulder, M.A., Coops, N.C. (In Review). Remote sensing opportunities for predicting species richness gradients at continental to global scales. *Remote Sensing in Ecology and Conservation.*

A version of Chapter 3 has been published in:

Thompson, S.D., Nelson, T.A., White, J.C., Wulder, M.A. 2015. Large area mapping of tree species using composited Landsat imagery. *Canadian Journal of Remote Sensing* 41(3): 203-218

A version of Chapter 4 has been submitted for publication as:

Thompson, S.D., Nelson, T.A., Giesbrecht, I, Frazer, G, Saunders, S. In Review. Data-driven regionalization of forested and non-forested ecosystems in coastal British Columbia with LiDAR and RapidEye imagery. *Applied Geography*.

A version of Chapter 5 has been submitted for publication as:

Thompson, S.D., Nelson, T.A., Coops, N.C., Wulder, M.A., Lantz, T.C. In Review. Spatial patterns of global primary productivity regimes from 2000-2012. *Environmental Monitoring and Assessment*.

Table of Contents

Contents

List of Tables

List of Figures

List of Acronyms

AET	Actual EvapoTranspiration
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AUC	Area Under the Curve
AVHRR	Advanced Very High Resolution Radiometer
BAP	Best Available Pixel
BEC	Biogeoclimatic Ecosystem Classification
CBD	Convention on Biological Diversity
DEM	Digital Elevation Model
fPAR	fraction of Photosynthetically Active Radiation
GBIF	Global Biodiversity Information Facility
GEO	BON Group on Earth Observations Biodiversity Observation Network
GEDI	Global Ecosystem Dynamics Investigation
GLAS	Geoscience Laser Altimeter System
GPP	Gross Primary Productivity
ICESAT	Ice, Cloud and land Elevation Satellite
IUCN	International Union for the Conservation of Nature
LiDAR	Light Detection And Ranging
MODIS	MODerate resolution Imaging Spectrometer
NDVI	Normalized Difference Vegetation Index
NFI	National Forestry Inventory
NPP	Net Primary Productivity
OOB	Out of Bag
PET	Potential EvapoTranspiration
RADAR	RAdio Detection And Ranging
ROC	Receiver Operating Characteristic
SPOT	Système Pour l'Observation de la Terre
SRTM	Shuttle Radar Topography Mission
TEM	Terrestrial Ecosystem Mapping
TPI	Topographic Position Index
TRASP	Topographic Radiation ASPect
TWI	Topographic Wetness Index
USGS	United States Geological Survey
WWF	World Wildlife Fund

Acknowledgements

I would like to thank my supervisor Trisalyn Nelson, and my supervisory committee members Michael Wulder, Trevor Lantz, and Nicholas Coops. I greatly appreciate the great ideas, suggestions and critiques at each stage of research, always provided in a timely manner. Their contributions undoubtedly helped me to learn and grow as a researcher, and guided this dissertation to successful completion.

I acknowledge the financial support provided by the Natural Sciences and Engineering Research Council (NSERC PGSD), and additional awards and assistantships provided by the University of Victoria, for which I am grateful.

I am hugely indebted to fellow members of the SPAR lab. Certainly the regular provision of technical advice was of benefit to me, but most importantly, I could not have completed this journey without your friendship and moral support! Finally, thank you to my family and friends who never seemed to doubt my abilities and for encouraging and supporting me along the way.

1. Introduction

1.1 Introduction to Biodiversity and Biodiversity Conservation

Biodiversity is defined by the United Nations as the variety of all life on Earth, including genetic diversity, species diversity and ecosystem diversity within terrestrial, freshwater and marine domains (United Nations 1992). At the same time, biodiversity also has compositional, structural and functional elements (Noss 1990). Composition refers to the identity and variety (e.g., species presence and abundance), structure refers to the physical or spatial pattern (e.g. forest height, canopy cover) and function refers to processes (e.g., primary productivity, nutrient cycling and disturbance) (Dale and Beyeler 2001; Franklin et al. 1981; Noss 1990). Ongoing research continues to improve knowledge of the importance of biodiversity for human well-being. As reviewed by Cardinale et al. (2012) and Hooper et al. (2005), the diversity of genes, species, and species functional groups influences ecosystem resilience and ecosystem functions such as biomass production and nutrient cycling. These ecosystem functions in turn influence the provision of ecosystem services such as the provision of clean air and clean water (Millennium Ecosystem Assessment 2005).

Biodiversity has declined dramatically over the past few hundred years (and particularly since 1950) as a result of human activities (Chapin et al. 2000; Millennium Ecosystem Assessment 2005). The greatest pressures on biodiversity are habitat loss (Sala 2000; Wilcove et al. 1998), especially related to agricultural expansion (Hoffmann et al. 2010), invasive alien species (Vitousek et al. 1997; Wilcove et al. 1998), and climate change (Bellard et al. 2012, Sala et al. 2000). International recognition of the need to protect biodiversity was formalized in the Convention on Biological Diversity (CBD), signed by 150 countries in 1992 at the Rio Earth

Summit who were committed to achieve a reduction in the loss of biodiversity. As of 2015, there were 168 signed party members committed. The CBD was intended to sustain biodiversity world-wide and promote sustainable development. Yet nearly 20 years after the Rio Earth Summit, the majority of targets initially set out had not yet been met (Butchart et al. 2010). For example, the global Red List Index (Butchart et al. 2004, 2005, 2007) indicates that population trends for many species at risk continue to deteriorate (e.g., Hoffmann et al. 2010). As well, the latest Global Forest Resource Assessment indicates that although global rates of deforestation have decreased, forests continue to be lost in many parts of the globe (Keenan et al. 2015). At a meeting in Aichi, Japan, 2010, five Strategic Goals and 20 “Aichi” Targets were established as part of the Convention on Biological Diversity’s Strategic Plan for Biodiversity for 2011-2020 (Table 1.1). While global in nature, they are designed to be implemented at national and local levels. Canada has signed the Convention on Biological Diversity, is committed to the Strategic Plan for Biodiversity 2011-2020, and has translated the global Aichi Biodiversity Targets framework to a national strategy and action plan (e.g., Table 1.1). Both global and Canadian targets follow the multifaceted and hierarchical nature of biodiversity by addressing each level and domain. For instance, Aichi Target 12 and Canadian Target 2 focus on species populations, aiming to improve population trends for species at risk, while Aichi Target 5 and Canadian Target 6 include a focus on ecosystem (e.g., forest) structure.

Table 1.1. Global and Canadian Biodiversity Goals and Targets for 2020

Aichi Global Goal	**Aichi Global Target**	**Canada Goal**^{2}	**Canada Target**^{2}
Address the underlying causes of BD loss by mainstreaming BD across government & society.	1. People are aware of BD values 2. BD values integrated into national planning & development 3. Incentives for conservation & sustainable use of BD applied; those harmful to BD eliminated 4. Sustainable production & consumption plans w/in gov't, businesses & stakeholders	Canada's lands & waters are planned & managed using an ecosystem approach to support BD conservation outcomes at local, regional & national scales.	1. Protected area coverage is ≥17% for terrestrial areas & inland waters & ≥ 10% coastal & marine areas 2. Mgmt. plans & recovery strategies improve trends for species at risk & maintain status of secure species 3. Wetland retention, restoration & mgmt. activities 4. Municipal planning accounts for BD 5. Climate change responses better understood & adaptation measures are underway
Reduce the direct pressures on BD & promote sustainable use.	5. Rates of habitat loss at least halved & fragmentation & degradation is sig. reduced 6. Marine & aquatic resource extraction is sustainable 7. Agriculture, forestry & aquaculture are sustainable 8. Pollution levels lowered 9. Invasive species mgmt. plans in place 10. Pressures on vulnerable ecosystems (e.g., coral reefs) lowered	Direct & indirect pressures as well as cumulative effects on BD are reduced, & production & consumption of Canada's biological resources are more sustainable.	6. Forests managed sustainably 7. Agricultural landscapes provide stable or improved habitat/BD capacity 8. Aquaculture is sustainable & science-based 9. Fisheries sustainable & ecosystem-based 10. Water is less polluted 11. Invasive species mgmt. plans in place 12. Aboriginal use of resources maintained & sustainable 13. Conservation & sustainable use mechanisms developed & applied
To improve the status of BD by safeguarding ecosystems, species & genetic diversity.	11. Protected area coverage is ≥17% for terrestrial areas & inland waters & ≥ 10% coastal & marine areas 12. Prevented extinction of threatened species (trends for species at risk are improved) 13. Genetic diversity maintained	Canadians have adequate & relevant information about BD & ecosystem services to support conservation planning & decision-making.	14. BD knowledge enhanced, integrated & accessible 15. Traditional knowledge promoted & used in mgmt. 16. Complete inventory of all the types of protected areas 17. National measures of natural capital in place

Enhance the benefits to all from BD & ecosystem services.	14. Ecosystem services safeguarded 15. Ecosystem resilience enhanced through conservation & restoration 16. Access & benefit sharing of genetic BD resources is operational	Canadians are informed about the value of nature & more actively engaged in its stewardship.	18. BD curricula developed for schools 19. Increased public participation in conservation activities
Enhance implementation through participatory planning, knowledge management & capacity building.	17. National BD Strategies & Action Plans are in place 18. Traditional knowledge & customary use promoted & used in mgmt. 19. BD knowledge improvement, transfer 20. Greater financial resource mobilization		

[1] See website for the Convention on Biological Diversity for more information about Aichi Biodiversity Goals and Targets: https://www.cbd.int/sp/targets/

[2] See "2020 Biodiversity Goals and Targets for Canada" (available at http://biodivcanada.ca/) for further information

1.2 Biodiversity data

In order to assess progress towards global conservation targets, and to facilitate environmentally sustainable decision making, biodiversity must be mapped and monitored at multiple scales (Braat and de Groot 2012; Ferrier et al. 2004; Pereira and Cooper 2006). Currently, biodiversity monitoring is a difficult task as knowledge of the distribution of biodiversity is globally inconsistent and incomplete (Guralnick and Hill 2009; Meyer et al. 2015; Scholes et al. 2012). Primary data regarding genetic composition is particularly lacking (Geijzendorffer et al. 2015; GEO BON 2011) and global biodiversity monitoring has traditionally been biased towards habitats and species (Laikre 2010). Projects such as Global Genome Initiative (Genome 10K Community of Scientists 2009) are rapidly increasing the availability of such data at scales ranging from within-species up to the ecosystem level (Yahara et al. 2010).

At the species level, knowledge of the spatial distribution of species world-wide is also limited (Ferrier et al., 2004). Discrete point data such as museum records and survey observations, known as primary occurrence data, are available from sources such as the Global Biodiversity Information Facility (GBIF) which, by September 2015, held over 576 million primary occurrence records for 1.6 million species (www.gbif.org). Nonetheless, a large number of species are absent, and records remain sparse for many regions (Meyer et al. 2015; Yesson et al. 2007). For example, a recent examination of the mammal data in the GBIF data showed that only 174 species had more than 100 records with georeferenced coordinates from the past 20 years (Boitani et al. 2011). In another study examining more than 2300 species of multiple taxa in Madagascar, more than half of the species had fewer than eight observations (Kremen et al. 2008). Existing occurrence data are biased towards certain, easily detectable organisms, specifically terrestrial vertebrates, plants and some insect groups (Newbold 2010) and the number of threatened species records are relatively low (Boakes et al. 2010). There is also considerable geographic bias in point occurrence datasets (e.g., Yang et al. 2013). For example, locations surveyed tend to be accessible areas close to roadways, rivers, urban areas, and biological stations (Graham et al. 2004; Rondinini et al. 2006).

Taxonomic, geographic, and temporal sampling bias may be particularly strong in large countries such as Canada (almost 10 million km^2 of which much is remote and inaccessible) (Boutin et al. 2009). Biodiversity data collection in Canada is currently conducted by different jurisdictions, at different spatial and time scales, measuring different parameters and using different methodologies, resulting in data gaps and uncertainties (Federal, Territorial and Provincial Governments of Canada 2010). A recent (2010) report concluded that biodiversity monitoring and research in the country was overall, “fair to poor… with some good data” (Federal, Territorial and Provincial Governments of Canada 2010). Biodiversity data availability in Canada lags behind other developed nations such as the United States and Western Europe (Meyer et al. 2015), particularly in the northern boreal and tundra biomes (Martin et al. 2012), but exceeds that of developing nations around the world (Amano and Sutherland 2013). Numerous national, international and regional institutes and activities exist to improve the organization, exchange, and availability of primary biological data for Canadian flora and fauna, including the Canadian Biodiversity Information Facility (www.cbif.gc.ca), and Canadensys (www.canadensys.net).

Primary occurrence data may be used to generate various secondary distribution products. For instance, a simple grid can be overlaid on the observation points to generate aerial estimate of species information. Grid sizes are generally chosen to be fairly large in order to overcome the spatial uncertainty, spatial bias and low numbers often associated with point occurrence data (Franklin 2009; Graham and Hijmans 2006). Thus gridded species distributions are generally only available for a small portion of Earth's species and are limited in their spatial extent (Jetz et al. 2012), which limits their usefulness for local and regional scale applications (Franklin 2009; Graham and Hijmans 2006). Alternatively, rather than generating gridded datasets, lists may be compiled of species for broad geopolitical or ecological region using observations from different areas and different times within general regions or biomes (Hortal 2008; Jetz et al. 2012).

Expert range maps are another type of species distribution data. Point occurrence data are used to determine and manually delineate the maximum geographic extent of a species' distribution (Graham and Hijmans 2006). Expert range maps of select species distributions are available at global scales from agencies such as the International Union for Conservation of Nature (IUCN) (www.iucnredlist.org), BirdLife International (www.birdlife.org), and NatureServe (www.natureserve.org). Expert range maps effectively indicate absence of a species outside of their boundaries, but greatly overestimate the presence of a species (Jetz et al. 2012). This high rate of commission error occurs because the maps are not designed to indicate habitat suitability at finer scales, that is, they assume uniform occurrence within the range (La Sorte and Hawkins 2007). Thus range maps are generally limited to use at broad scales and for generalist species (Hortal 2008; Hurlbert and Jetz 2007; Jetz et al. 2007).

Species distribution models are another form of secondary distribution data. These models aim to either a) spatially interpolate between known occurrences within a species' current range, or b) to extrapolate beyond known occurrences in space or time (Elith and Leathwick 2009). Most common are correlative models that relate species occurrence data to various environmental variables to predict the spatial distribution of a species. The environmental variables are those that are known or hypothesized to affect where a species can be found, either through direct resource and physiological needs such as temperature and water, or through indirect surrogates of those such as altitude, slope and aspect (Austin 2002, 2007; Guisan and Zimmermann 2000). The predictive ability of these models is affected by factors such as the quality and quantity of species occurrence data (Moudrý and Šímová 2012; Wisz et al. 2008), the selection and type of environmental predictor variables (Austin 2007; Stockwell and Peterson 2002), model type (Elith et al. 2006; Tsoar et al. 2007) and species traits (Guisan et al. 2007; Kadmon et al. 2003).

Knowledge of the global distribution of ecosystems is more developed than that of species distributions. Various classifications of global, broad-scale biogeographic units were created as early as the 18th century (Cox 2001). Early regionalizations had roots in evolutionary biology and biogeography (Kreft and Jetz 2010). Beginning in the mid-20th century, finer-scaled global regionalizations were sought for purposes of biological conservation and land management (Jepson and Whittaker 2002) and were made possible due to greater information about global climate and vegetation patterns (Holdridge 1967). A more recent and widely referenced system, especially in conservation literature, was developed by the World Wildlife Fund (WWF) and recognizes more than 800 terrestrial ecoregions globally (Olson et al. 2001). Ecosystem mapping is also routinely carried out at local and regional scales for natural resource management and conservation planning purposes (Banner et al. 1996; Crins et al. 2009).

1.2 Biodiversity indicators

Given the complexity and scope of the term “biodiversity” (Noss 1990), and due to the paucity of biodiversity data, biodiversity monitoring must focus on select surrogates, or indicators. Furthermore, biodiversity is complex and multifaceted, and monitoring cannot focus on a single measure, scale, or level of organization (Dale and Beyeler 2001; Scholes et al. 2012) Dozens of indicators of biodiversity of been proposed in recent years to monitor biodiversity, some examples of which are given in Table 1.2. Many authors have recently commented on the desirable characteristics of biodiversity indicators. Ideally, indicators should be cost-effective and repeatable, informative or calculable at multiple spatial extents and resolutions, responsive to perturbations, easy to compute and easy to understand, related to a particular target or management question, and proven to relate to one or more elements of biodiversity (Feld et al. 2010; Jones et al. 2011; Mace and Baillie 2007; Noss 1990). Despite the emerging consensus on the importance of indicators and careful indicator selection, the poor spatial and temporal coverage of biodiversity data limits the development or applicability of many indicators (Walpole et al. 2009). The indicators, like the primary biodiversity data they draw on, are biased towards temperate regions and developed countries, and many are from only one or two points in time (Jones et al. 2011; Walpole et al. 2009).

Table 1.2. Example biodiversity indicators and indicator data

Biodiversity component	**GENERAL**		**REMOTE SENSING**	
	Examples	**Limitations**	**Examples**	**Limitations**
Genetic composition	-Allelic diversity (allelic richness of selected species) -Breed & variety diversity (e.g., # of breeds of domesticated livestock species)	-incomplete taxonomic coverage -incomplete geographic sampling	n/a	n/a
Species composition	-Red List Index (Butchart et al. 2004, 2005, 2007) -Living planet index (Collen et al. 2009; Loh et al. 2005) -COSEWIC listings	-temporal resolution -incomplete taxonomic coverage -incomplete spatial coverage	-Biodiversity Intactness Index (Scholes and Biggs 2005) -Biodiversity Habitat Index (GEO BON 2015) -Species Habitat Indices (GEO BON 2015) -species distribution models	-spatial resolution & coverage (global indicator needs to be adapted to finer scales, and fine scale models need to be applied to larger areas) -temporal coverage / consistency limited by remotely sensed data availability -species observation data limits what species can be modelled and model accuracy
Ecosystem structure	-Local *in-situ* ecological studies, including long-term ecological research (LTER) plots -FAO Global Forest Resources Assessment of forest cover (e.g., Keenan et al. 2015)	-in-situ studies limited by spatial coverage, temporal resolution and cost -Global data not spatially explicit and not detailed enough	-Canada's National Forest Inventory (Gillis et al. 2005) -A New Map of Global Ecological Land Units (Sayre et al. 2014) -global land cover maps	-spatial resolution of global datasets -spatial, thematic and methodological consistency, especially of global land cover maps -spatial coverage of forest inventories

It is increasingly recognized that remotely sensed data are needed to help address gaps in biodiversity data. Remotely sensed data are ideal for biodiversity monitoring because they can repeatedly provide systematically collected data over the entire earth. In comparison, cost and logistics limit the collection of *in-situ* observations in space and time, especially in large and remote areas or politically insecure areas (Buchanan et al. 2009; Duro et al. 2007). With the exception of genetic biodiversity, remote sensing can characterize multiple attributes of biodiversity over multiple geographic extents, spatial resolutions, and temporal resolutions (Duro et al. 2007; Gillespie et al. 2008; Nagendra 2001; Turner et al. 2003). The type of information that can be derived from a remote sensing image is a function of the sensor's temporal, spectral, and spatial resolution (Turner et al. 2003). Specifically, these attributes determine how often a given area is re-visited by the sensor, and the type and size of features on the ground that can be resolved (Table 1.3). For instance, imagery is generally referred to high, or H-resolution when image pixels are smaller than the objects under investigation, and as low, or L-resolution, when image pixels are larger than the features of interest (Strahler et al. 1986). Spatial and temporal resolution are also also related to the areal extent of a given image (the footprint). Specifically, sensors with a high temporal resolution are in orbit very high above the earth and have large footprints but low to moderate spatial resolution. For local and landscape scale applications, highly detailed information can be derived from small-footprint, high spatial resolution (< 10 m) imagery, hyperspectral imagery and active sensors such as lidar (light detection and ranging). Nation-wide biodiversity monitoring in Canada incorporates remotely sensed data from multispectral, moderate-to-low resolution sensors such as Landsat and AVHRR to assess trends in such things as land cover, forest cover, forest fragmentation, disturbance, and primary productivity (Ahern et al. 2011; Duro et al. 2007). Globally, a number of indicators capitalize on remotely sensed data, including several new “Essential Biodiversity Variables” (GEO BON 2015; Pereira et al. 2013) proposed by the Group on Earth Observations – Biodiversity Observation Network (GEO BON) (Table 1.2).

Table 1.3. Characteristics of common multispectral remotely sensed imagery

Sensor	Class (spatial resolution)	Spatial resolution	Swath width	Launch / Operating Dates	Approximate revisit time	Potential uses
Airborne multispectral scanners	Very High	< 1 m			Deployed on demand	Identification of large individual trees and detailed structural information; species identification sometimes possible
WorldView-2	Very High	1.8 m	16.4 km	2009	1.1-3.7 days	
GeoEye-1	Very High	1.84 m	15.2 km	2008	2.1-8.3 days	
QuickBird-2	Very High	2.44 m	16.5 km	2001	1-3.5 days	
IKONOS	High	4 m	11.3 km	1999	3 days	Landscape scale vegetation mapping, including forest stand structure (e.g., successional stage); species identification sometimes possible
RapidEye	High	5 m	77 km	2008	1-5.5 days	
Landsat MSS	Medium	60 m	185 km	1972–1992	16-18 days	Separation of coniferous vs deciduous forest stands; landscape disturbance / successional stage mapping; species identification may be possible
Landsat TM /ETM+	Medium	30 m	185 km	1984–2013 (TM) 1999-2003 (ETM+)	8-16 days	
Landsat 8	Medium	30 m	185 km	2013	8-16 days	
Sentinel-2	Medium	10-60 m	290 km	2014	5-10 days	
Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER)	Medium	15-90 m	60 km	1999	16 days	
MODerate resolution Imaging Spectroradiometer (MODIS) Terra and Aqua	Low	250-1000 m	2330 km	1999 (Terra) 2002 (Aqua)	1-2 days	National to global land cover types, biome and biomass mapping; species identification not possible
ENVISAT MERIS	Low	300 m	1150 km	2002-2012	3 days	
Advanced Very High Resolution Radiometer (AVHRR)	Very low	≥ 1 km	2400 km	1981	1 day	

1.3 Research objectives

Despite the frequent and increasing use of remote sensing for biodiversity monitoring globally and within Canada, there are opportunities for expansion and improvement at both levels. Firstly, remote sensing scientists need to communicate more directly with ecologists and biologists (Pettorelli et al. 2014; Skidmore et al. 2015) and need to more effectively demonstrate the link between remote sensing–based indicators (e.g., land cover) and particular aspects of biodiversity (e.g., species distributions or ecosystem function) (Buchanan et al. 2009; Han et al. 2014; Walpole et al. 2009). Secondly, work is needed to address gaps and uncertainties with respect to spatial resolution and spatial scaling. The spatial resolutions of current indicator data do not always match what is needed or desired by users (e.g., Han et al. 2014), with many indicators developed at the global scale unable to be scaled to sub-global scales, or vice-versa (Secades et al. 2014). The cost of higher spatial resolution datasets are often prohibitive to monitoring at national and especially international scales. Thus while research that develops high spatial resolution indicators needs to continue, case studies demonstrating the capabilities and limitations of free and open-source remotely sensed data for diverse regions and scales will increase confidence and interest in the remote sensing indicators (Skidmore et al. 2015).

The overall goal of this dissertation is to advance the mapping and monitoring of biodiversity indicators, globally and within Canada, through the use of remote sensing. In particular, to advance knowledge of the information content of various types of remotely sensed data for biodiversity monitoring needs at the species and ecosystem levels. The biodiversity monitoring needs addressed by each research chapter are described in the following paragraphs, but common to all are i) the need for cost efficiency, ii) the need to increase spatial and temporal resolution, spatial coverage, and scalability of indicator data, and iii) better integration of remote sensing with ecology. The majority of data assessed in this dissertation are free and open-source remotely sensed datasets, and research included spans local to global scales, while also including compositional, structural and functional attributes of biodiversity at both the species and ecosystem levels (Figure 1.2).

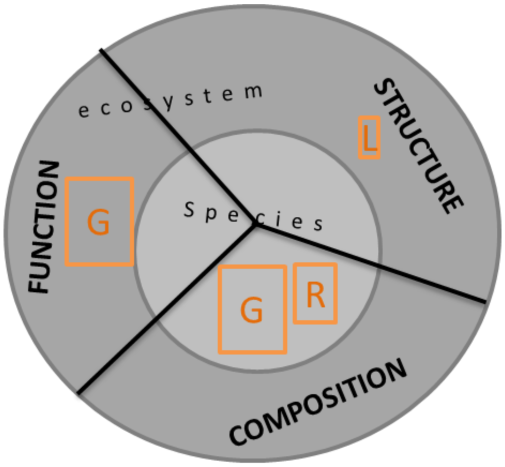


Figure 1.1. Biodiversity incorporates structure, composition, and function for genes (not shown), species (inner circle) and ecosystems (outer circle). This dissertation will study species composition at global (G) and regional (R) scales (Chapters 2 and 3, respectively), ecosystem structure at local (L) scales (Chapter 4), and ecosystem function at global scales (Chapter 5).

Starting at the species level, perhaps one of the easiest to comprehend measures of biodiversity is species richness, which is simply the number of different species found in a given area. Species richness for a given area may be modelled in several ways, including through a spatial overlap and summing of individual expert range maps (Graham and Hijmans 2006), or alternatively through modeling richness directly as a function of environmental variables, which may or may not be sourced from remote sensing (e.g., Ferrier et al. 2004). To monitor species richness over time, indicators such as the Biodiversity Intactness Index (Scholes and Biggs 2005) and the Biodiversity Habitat Index (GEO BON 2015) have been proposed, both utilizing remotely sensed land use and/or forest cover change datasets to infer effects on richness. These indicators are directly relevant to Aichi Target 5. For individual countries or researchers to adopt these types of indicators for national-level reporting, knowledge of which datasets are available and appropriate at sub-global scales will be important (e.g., Geijzendorffer et al. 2015). The goal of Chapter 2 is to determine which remotely sensed data products are available and appropriate for modelling species richness for each of six global biogeographic realms. The research also provides an explicit list of remotely sensed data products along with a discussion of some general considerations for monitoring changes to richness that will help address the gap between ecologists and remote sensing scientists, and where appropriate, inform the selection of data in subsequent research chapters.

Continuing at the species level of biodiversity**,** trends in species populations are currently monitored globally with indicators such as the Red List Index (Butchart et al. 2004, 2005, 2007) and the Living Planet Index (Collen et al. 2009; Loh et al. 2005), with relevance to Aichi Target 12 (Table 1.1, Table 1.2). Progress within Canada towards Target 2 is currently assessed based on species at risk population trends as supplied by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC). Recently “The Species Habitat Indices” (GEO BON 2015) were proposed as a way to improve the geographic and temporal coverage of the aforementioned indices at the global scale (Table 1.2). These indices model the suitable habitat of species using remotely sensed data (e.g., land cover), are validated by species occurrence data, and can be assessed on an annual basis. A similar national-level indicator that utilizes remotely sensed data has not been developed for Canada. The goal of Chapter 3 is to generate spatially detailed distribution maps for six tree species over an unprecedently large forested area of Canada ((>39 million ha). Presented as a study that tests the effect of Best-Available-Pixel (BAP) Landsat composite imagery on species distribution models, the approach in Chapter 3 may be adaptable to standardized biodiversity monitoring in Canada as discussed in Chapter 6. Chapter 3 addresses the call for further case studies of remotely sensed biodiversity monitoring and questions about scale.

At the ecosystem structure level, land cover and forest cover change, and forest fragmentation are among the indicators that may be used to assess progress towards Aichi Targets 5 (Table 1.1). As with species-level mapping, monitoring of ecosystem structure is challenged by spatial resolution and data availability. A number of different global land cover products have been produced over the past two decades with different sensors, spatial resolutions and classification schemes (Herold et al. 2008). Until recently, global products were limited to sensors with broad footprints and low spatial resolutions (> 300 m), however, recent global datasets have been produced with Landsat data at 30 m spatial resolution (Chen et al. 2015; Hansen et al. 2013). Within Canada, a land cover time series is available from the Canada Centre for Remote Sensing (Ahern et al. 2011), produced with data from the Advanced Very High Resolution Radiometer (AVHRR) at 1 km spatial resolution for every 5 years between 1985 and 2005. More detailed vegetation structural information at finer spatial resolutions are available from forest inventories and regional ecosystem mapping efforts. However, these systems rely on the manual interpretation of aerial photographs, limiting their cost effectiveness, spatial coverage, and repeat frequency. In Chapter 4, a thematically rich, relatively detailed (20 m spatial resolution) map is created to characterize the variety of local-scale ecosystem structural classes in coastal British Columbia using remotely sensed data. The results are compared to an existing mapping framework and methodology to provide explicit examples to ecologists of the opportunities present in new remotely sensed data. The potential for adapting the mapping methodology demonstrated in Chapter 4 for Canada-wide biodiversity monitoring is discussed in Chapter 6.

Finally, also at the at the ecosystem level, but turning to ecosystem function, Net Primary Productivity (NPP) is a fundamental ecosystem function and one of the best established of the remote-sensing based biodiversity indicators (e.g., Kerr and Ostrovsky 2003; Turner et al. 2003). Global remotely-sensed estimates of NPP or proxies such as the Normalized Difference Vegetation Index (NDVI) are free and readily available, with data coverage extending back over 30 years in some cases. Monitoring trends in NPP can be used to assess progress towards Aichi Targets 5, 14, and 15 (e.g., GEO BON 2015) and Canadian biodiversity Targets 6 and 17 (e.g., Coops et al. 2014; Ahern et al. 2011). At the global scale, trends in NPP are typically reported for pre-defined ecoregions or land cover types, which may limit the ability for the information to be downscaled to finer spatial resolutions and extents, and may also limit the ecological meaning of the results if global change is sufficiently large to question the resilience of those units over time. In Chapter 5, an alternative method to assessing change in land productivity is proposed. The method focuses on spatial pattern within data-driven regions, rather than using *a priori* defined land cover types, and generates new spatial indicators from remotely sensed data to assess variability in productivity and phenology between 2000 and 2012 across the globe. Applicability of this approach to biodiversity monitoring in Canada is discussed in Chapter 6.

1.4 References

Ahern, F., Frisk, J., Latifovic, R., & Pouliot, D. (2011). *Monitoring ecosystems remotely: a selection of trends measured from satellite observations of Canada. Canadian Biodiversity: Ecosystem Status and Trends 2010, Technical Thematic Report No. 17.* Ottawa, Ontario.

Amano, T., & Sutherland, W. (2013). Four barriers to the global understanding of biodiversity conservation: wealth, language, geographical location and security. *Proceedings of the Royal Society B: Biological Sciences*, *280*.

Austin, M. P. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, *157*(2-3), 101–118. doi:10.1016/S0304-3800(02)00205-3

Austin, M. P. (2007). Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*, *200*(1-2), 1–19. doi:10.1016/j.ecolmodel.2006.07.005

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. doi:10.1007/BF00396139

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. *Ecology Letters*, *15*(4), 365–377. doi:10.1111/j.1461-0248.2011.01736.x

Boakes, E. H., McGowan, P. J. K., Fuller, R. a, Chang-qing, D., Clark, N. E., O'Connor, K., & Mace, G. M. (2010). Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS biology*, *8*(6), e1000385. doi:10.1371/journal.pbio.1000385

Boitani, L., Maiorano, L., Baisero, D., Falcucci, A., Visconti, P., & Rondinini, C. (2011). What spatial data do we need to develop global mammal conservation strategies? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *366*(1578), 2623–32. doi:10.1098/rstb.2011.0117

Boutin, S., Haughland, D. L., Schieck, J., Herbers, J., & Bayne, E. (2009). A new approach to forest biodiversity monitoring in Canada. *Forest Ecology and Management*, *258S*, S168–S175. doi:10.1016/j.foreco.2009.08.024

Braat, L. C., & de Groot, R. (2012). The ecosystem services agenda: bridging the worlds of natural science and economics, conservation and development, and public and private policy. *Ecosystem Services*, *1*(1), 4–15. doi:10.1016/j.ecoser.2012.07.011

Buchanan, G. M., Nelson, A., Mayaux, P., Hartley, A., & Donald, P. F. (2009). Delivering a global, terrestrial, biodiversity observation system through remote sensing. *Conservation Biology*, *23*(2), 499–502. doi:10.1111/j.1523-1739.2008.01083.x

Butchart, S. H. M., Resit Akçakaya, H., Chanson, J., Baillie, J. E. M., Collen, B., Quader, S., et al. (2007). Improvements to the Red List Index. *PloS ONE*, *2*(1), e140. doi:10.1371/journal.pone.0000140

Butchart, S. H. M., Stattersfield, a J., Baillie, J., Bennun, L. a, Stuart, S. N., Akçakaya, H. R., et al. (2005). Using Red List Indices to measure progress towards the 2010 target and beyond. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *360*(1454), 255–68. doi:10.1098/rstb.2004.1583

Butchart, S. H. M., Stattersfield, A. J., Bennun, L. a, Shutes, S. M., Akçakaya, H. R., Baillie, J. E. M., et al. (2004). Measuring global trends in the status of biodiversity: red list indices for birds. *PLoS biology*, *2*(12), e383. doi:10.1371/journal.pbio.0020383

Butchart, S. H. M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., Almond, R. E. a, et al. (2010). Global biodiversity: indicators of recent declines. *Science*, *328*(5982), 1164–8. doi:10.1126/science.1187512

Canadian Council of Forest Ministers. (2006). *Criteria and indicators of sustainable forest management in Canada: national status 2005*. Ottawa, Ontario.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., et al. (2012). Corrigendum: Biodiversity loss and its impact on humanity. *Nature*, *489*(7415), 326–326. doi:10.1038/nature11373

Chapin, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., et al. (2000). Consequences of changing biodiversity. *Nature*, *405*(6783), 234–42. doi:10.1038/35012241

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., et al. (2015). Global land cover mapping at 30m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, *103*, 7–27. doi:10.1016/j.isprsjprs.2014.09.002

Collen, B., Loh, J., Whitmee, S., Mcrae, L., Amin, R., & Baillie, J. E. M. (2009). Monitoring change in vertebrate abundance: the Living Planet Index. *Conservation Biology*, *23*(2), 317–327. doi:10.1111/j.1523-1739.2008.01117.x

Coops, N. C., Fontana, F. M. A., Harvey, G. K. A., Nelson, T. A., & Wulder, M. A. (2014). Monitoring of a national-scale indirect indicator of biodiversity Using a long time-series of remotely sensed imagery. *Canadian Journal of Remote Sensing*, *40*(3), 179–191. doi:10.1080/07038992.2014.945826

Cox, B. (2001). The biogeographic regions reconsidered. *Journal of Biogeography*, *28*, 511–523.

Crins, W., Gray, P., Uhlig, P., & Wester, M. (2009). *The ecosystems of Ontario, Part 1: Ecozones and ecoregions. Technical Report SIB TER IMA TR-01. Technical Report SIB TER IMA TR-01* (Vol. Science an). Ontario MNR. Science & Information Branch Inventory, Monitoring and Assessment Section.

Dale, V. H., & Beyeler, S. C. (2001). Challenges in the development and use of ecological indicators. *Ecological Indicators*, *1*(1), 3–10. doi:10.1016/S1470-160X(01)00003-6

Duro, D. C., Coops, N. C., Wulder, M. A., & Han, T. (2007). Development of a large area biodiversity monitoring system driven by remote sensing. *Progress in Physical Geography*, *31*(3), 235–260. doi:10.1177/0309133307079054

Elith, J., Graham, C., Anderson, R., Dudik, M., Ferrier, S., Guisan, A., et al. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), 129–151.

Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677–697. doi:10.1146/annurev.ecolsys.110308.120159

Federal Territorial & Provincial Governments of Canada. Canadian Biodiversity: Status and Trends 2010 (2010). Ottawa, Ontario.

Feld, C. K., Sousa, J. P., Silva, P. M., & Dawson, T. P. (2010). Indicators for biodiversity and ecosystem services: towards an improved framework for ecosystems assessment. *Biodiversity and Conservation*, *19*(10), 2895–2919. doi:10.1007/s10531-010-9875-0

Ferrier, S., Powell, G. V. N., Richardson, K. S., Manion, G., Overton, J. M., Allnutt, T. F., et al. (2004). Mapping more of terrestrial biodiversity for global conservation assessment. *BioScience*, *54*(12), 1101–1109.

Franklin, J. (2009). *Mapping species distributions: spatial inference and prediction*. Cambridge University Press, Cambridge UK.

Franklin, J., Cromack, K. J., Denison, W., McKee, A., Maser, C., Sedell, J., et al. (1981). *Ecological characteristics of old-growth Douglas-fir forests*. *General Technical Report PNW-118*. doi:10.1007/s13398-014-0173-7.2

Geijzendorffer, I. R., Regan, E. C., Pereira, H. M., Brotons, L., Brummitt, N., Gavish, Y., et al. (2015). Bridging the gap between biodiversity data and policy reporting needs: An Essential Biodiversity Variables perspective. *Journal of Applied Ecology*, n/a–n/a. doi:10.1111/1365-2664.12417

Genome 10K Community of Scientists. (2009). Genome 10K: a proposal to obtain whole-genome sequence for 10,000 vertebrate species. *The Journal of heredity*, *100*(6), 659–74. doi:10.1093/jhered/esp086

GEO BON. (2011). Adequacy of biodiversity observation systems to support the CBD 2020 targets. A report for prepared by the Group on Earth Observations Biodiversity Observation Network for the Convention on Biological Diversity. Pretoria, South Africa.

GEO BON. (2015). *Global Biodiversity Change Indicators. Version 1.1.* Leipzig, Germany.

Gillespie, T., Foody, G. M., Rocchini, D., Giorgi, a. P., & Saatchi, S. S. (2008). Measuring and modelling biodiversity from space. *Progress in Physical Geography*, *32*(2), 203–221. doi:10.1177/0309133308093606

Gillis, M. D., Omule, A. Y., & Brierley, T. (2005). Monitoring Canada's forests: The National Forest Inventory. *The Forestry Chronicle*, *81*(2), 214–221. doi:10.5558/tfc81214-2

Graham, C. H., Ferrier, S., Huettman, F., Moritz, C., & Peterson, a T. (2004). New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology & Evolution*, *19*(9), 497–503. doi:10.1016/j.tree.2004.07.006

Graham, C. H., & Hijmans, R. J. (2006). A comparison of methods for mapping species ranges and species richness. *Global Ecology and Biogeography*, *15*, 578–587. doi:10.1111/j.1466-822x.2006.00257.x

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2-3), 147–186. doi:10.1016/S0304-3800(00)00354-9

Guisan, A., Zimmermann, N. E., Elith, J., Graham, C. H., Phillips, S. J., & Peterson, A. T. (2007). What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? *Ecological Monographs*, *77*(4), 615–630.

Guralnick, R., & Hill, A. (2009). Biodiversity informatics: automated approaches for documenting global biodiversity patterns and processes. *Bioinformatics*, *25*(4), 421–8. doi:10.1093/bioinformatics/btn659

Han, X., Smyth, R. L., Young, B. E., Brooks, T. M., Sánchez de Lozada, A., Bubb, P., et al. (2014). A Biodiversity Indicators Dashboard: Addressing Challenges to Monitoring Progress towards the Aichi Biodiversity Targets Using Disaggregated Global Data. *PLoS ONE*, *9*(11), e112046. doi:10.1371/journal.pone.0112046

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., et al. (2013). High-resolution global maps of 21st-Century forest cover change. *Science*, *342*(6160), 850–853. doi:10.1126/science.1244693

Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A., & Schmullius, C. (2008). Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment*, *112*(5), 2538–2556. doi:10.1016/j.rse.2007.11.013

Hoffmann, M., Hilton-taylor, C., Angulo, A., Böhm, M., Brooks, T. M., Butchart, S. H. M., et al. (2010). The Impact of Conservation on the Status of the World 's Vertebrates. *Science, 330*(6010), 1503–1509. doi:10.1126/science.1194442

Holdridge, L. R. (1967). *Life Zone Ecology*. San Jose, Costa Rica: Tropical Science Centre.

Hooper, D., Chapin, F. S., Ewel, J., Hector, A., Inchausti, P., Lavorel, S., et al. (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological Monographs*, *75*(1), 3–35.

Hortal, J. (2008). Uncertainty and the measurement of terrestrial biodiversity gradients. *Journal of Biogeography*, *35*(8), 1335–1336. doi:10.1111/j.1365-2699.2008.01955.x

Hurlbert, A. H., & Jetz, W. (2007). Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(33), 13384–13389. doi:10.1073/pnas.0704469104

Jepson, P., & Whittaker, R. J. (2002). Ecoregions in context: a critique with special reference to Indonesia. *Conservation Biology*, *16*(1), 42–57. doi:10.1046/j.1523-1739.2002.01143.x

Jetz, W., McPherson, J. M., & Guralnick, R. P. (2012). Integrating biodiversity distribution knowledge: toward a global map of life. *Trends in Ecology & Evolution*, *27*(3), 151–159. doi:10.1016/j.tree.2011.09.007

Jetz, W., Sekercioglu, C. H., & Watson, J. E. M. (2007). Ecological correlates and conservation implications of overestimating species geographic ranges. *Conservation Biology*, *22*(1), 110–119. doi:10.1111/j.1523-1739.2007.00847.x

Jones, J. P. G., Collen, B., Atkinson, G., Baxter, P. W. J., Bubb, P., Illian, J. B., et al. (2011). The why, what, and how of global biodiversity indicators beyond the 2010 target. *Conservation Biology*, *25*(3), 450–7. doi:10.1111/j.1523-1739.2010.01605.x

Kadmon, R., Farber, O., & Danin, A. (2003). A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*, *13*(3), 853–867.

Keenan, R. J., Reams, G. A., Achard, F., Freitas, J. V. De, Grainger, A., & Lindquist, E. (2015). Dynamics of global forest area: Results from the FAO Global Forest Resources Assessment 2015. *Forest Ecology and Management*, *352*, 9–20.

Kerr, J. T., & Ostrovsky, M. (2003). From space to species: ecological applications for remote sensing. *Trends in Ecology & Evolution*, *18*(6), 299–305. doi:10.1016/S0169-5347(03)00071-5

Krausmann, F., Erb, K.-H., Gingrich, S., Haberl, H., Bondeau, a., Gaube, V., et al. (2013). Global human appropriation of net primary production doubled in the 20th century. *Proceedings of the National Academy of Sciences*, *110*(25), 10324–10329. doi:10.1073/pnas.1211349110

Kreft, H., & Jetz, W. (2010). A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography*, *37*(11), 2029–2053. doi:10.1111/j.1365-2699.2010.02375.x

Kremen, C., Cameron, A., Moilanen, A., Phillips, S. J., Thomas, C. D., Beentje, H., et al. (2008). Aligning conservation priorities across taxa in Madagascar with high-resolution planning tools. *Science*, *320*(5873), 222–6. doi:10.1126/science.1155193

La Sorte, F. A., & Hawkins, B. A. (2007). Range maps and species richness patterns: errors of commission and estimates of uncertainty. *Ecography*, *30*(5), 649–662. doi:10.1111/j.2007.0906-7590.05084.x

Laikre, L. (2010). Genetic diversity is overlooked in international conservation policy implementation. *Conservation Genetics*, *11*(2), 349–354. doi:10.1007/s10592-009-0037-4

Loh, J., Green, R. E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V., & Randers, J. (2005). The Living Planet Index: using species population time series to track trends in biodiversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1454), 289–295. doi:10.1098/rstb.2004.1584

Ma, T., Zhou, C., & Pei, T. (2012). Simulating and estimating tempo-spatial patterns in global human appropriation of net primary production (HANPP): A consumption-based approach. *Ecological Indicators*, *23*, 660–667. doi:10.1016/j.ecolind.2012.05.026

Mace, G. M., & Baillie, J. E. M. (2007). The 2010 biodiversity indicators: challenges for science and policy. *Conservation Biology*, *21*(6), 1406–13. doi:10.1111/j.1523-1739.2007.00830.x

Martin, L. J., Blossey, B., & Ellis, E. (2012). Mapping where ecologists work: Biases in the global distribution of terrestrial ecological observations. *Frontiers in Ecology and the Environment*, *10*(4), 195–201. doi:10.1890/110154

Meyer, C., Kreft, H., Guralnick, R., & Jetz, W. (2015). Global priorities for an effective information basis of biodiversity distributions. *Nature Communications*, *6*, 8221. doi:10.1038/ncomms9221

Millennium Ecosystem Assessment. (2005). Ecosystems and human well-being: biodiversity synthesis. Washington, DC: World Resources Institute.

Moudrý, V., & Šímová, P. (2012). Influence of positional accuracy, sample size and scale on modelling species distributions: a review. *International Journal of Geographical Science, 26*(11), 2083–2095.

Nagendra, H. (2001). Using remote sensing to assess biodiversity. *International Journal of Remote Sensing, 22*(12), 2377–2400.

Newbold, T. (2010). Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. *Progress in Physical Geography*, *34*(1), 3–22. doi:10.1177/0309133309355630

Noss, R. F. (1990). Indicators for monitoring biodiversity: a hierarchical approach. *Conservation Biology*, *4*(4), 355–364. doi:10.1111/j.1523-1739.1990.tb00309.x

Olson, D. M., Dinerstein, E., Wikramanaya, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., et al. (2001). Terrestrial ecoregions of the world: a new map of life on Earth. *BioScience*, *51*(11), 933–938. doi:http://dx.doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Pereira, H. M., & Cooper, H. D. (2006). Towards the global monitoring of biodiversity change. *Trends in Ecology & Evolution*, *21*(3), 123–9. doi:10.1016/j.tree.2005.10.015

Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jognman, R. H. G., Scholes, R. J., et al. (2013). Essential Biodiversity Variables. *Science*, *339*(6117), 277–278.

Pettorelli, N., Safi, K., Turner, W., & B, P. T. R. S. (2014). Satellite remote sensing, biodiversity research and conservation of the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1643), 20130190. doi:10.1098/rstb.2013.0190

Rondinini, C., Wilson, K. a, Boitani, L., Grantham, H., & Possingham, H. P. (2006). Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology Letters*, *9*(10), 1136–45. doi:10.1111/j.1461-0248.2006.00970.x

Sala, O. E. (2000). Global Biodiversity Scenarios for the Year 2100. *Science, 287*(5459), 1770–1774. doi:10.1126/science.287.5459.1770

Sayre, R., Dangermond, J., Frye, R., Vaughan, A., Aniello, P., Breyer, S., et al. (2014). *A new map of global ecological land units - an ecophysiographic stratification approach*. Washington, DC.

Scholes, R. J., & Biggs, R. (2005). A biodiversity intactness index. *Nature*, *434*(7029), 45–9. doi:10.1038/nature03289

Scholes, R. J., Walters, M., Turak, E., Saarenmaa, H., Heip, C. H., Tuama, É. Ó., et al. (2012). Building a global observing system for biodiversity. *Current Opinion in Environmental Sustainability*, *4*(1), 139–146. doi:10.1016/j.cosust.2011.12.005

Secades, C., O'Connor, B., Brown, C., & Walpole, M. (2014). *Earth Observation for Biodiversity Monitoring: A review of current approaches and future opportunities for tracking progress towards the Aichi Biodiversity Targets.* CBD Technical Series No 72. Montreal, Canada: Secretariat of the Convention on Biological Diversity.

Skidmore, A. K., Pettorelli, N., Coops, N. C., Geller, G., Hansen, M. C., Lucas, R., et al. (2015). Environmental science: Agree on biodiversity metrics to track from space. *Nature*, *523*(7561), 403–405. doi:10.1038/523403a

Stockwell, D. R., & Peterson, A. T. (2002). Effects of sample size on accuracy of species distribution models. *Ecological Modelling, 148*(1), 1–13. doi:10.1016/S0304-3800(01)00388-X

Strahler A. H., Woodcock C. E., Smith J. A. (1986). On the nature of models in remote sensing. *Remote Sensing of Environment*, 20(2), 121–39. doi:10.1016/0034-4257(86)90018-0

Tsoar, A., Allouche, O., Steinitz, O., Rotem, D., & Kadmon, R. (2007). A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions, 13*(4), 397–405. doi:10.1111/j.1472-4642.2007.00346.x

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., & Steininger, M. (2003). Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution, 18*(6), 306–314. doi:10.1016/S0169-5347(03)00070-3

United Nations. Convention on Biological Diversity (1992). Rio de Janeiro, 5 June 1992.

Vitousek, P. M., D'Antonio, C. M., Loope, L. L., Rejmánek, M., & Westbrooks, R. (1997). Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology*, *21*(1), 1–16.

Walpole, M., Almond, R. E. a, Besançon, C., Butchart, S. H. M., Campbell-Lendrum, D., Carr, G. M., et al. (2009). Tracking progress toward the 2010 biodiversity target and beyond. *Science*, *325*(5947), 1503–4. doi:10.1126/science.1175466

Wilcove, D. S., Rothstein, D., Dubow, J., Phillips, A., & Losos, E. (1998). Quantifying Threats to Imperiled Species in the United States. *BioScience*, *48*(8), 607–615. doi:10.2307/1313420

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., & Guisan, A. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions, 14*(5), 763–773. doi:10.1111/j.1472-4642.2008.00482.x

Yahara, T., Donoghue, M., Zardoya, R., Faith, D., & Cracraft, J. (2010). Genetic diversity assessments in the century of genome science. *Current Opinion in Environmental Sustainability*, *2*(1-2), 43–49. doi:10.1016/j.cosust.2010.03.008

Yang, W., Ma, K., & Kreft, H. (2013). Geographical sampling bias in a large distributional database and its effects on species richness-environment models. *Journal of Biogeography*, *40*(8), 1415–1426. doi:10.1111/jbi.12108

Yesson, C., Brewer, P. W., Sutton, T., Caithness, N., Pahwa, J. S., Burgess, M., et al. (2007). How global is the global biodiversity information facility? *PLoS ONE*, *2*(11), e1124. doi:10.1371/journal.pone.0001124

2. Remote sensing opportunities for predicting species richness gradients at continental to global scales

Abstract

The limited quantity and quality of species observation data globally results in species richness frequently being modeled as a function of environmental covariates. These covariates can come from a variety of sources, including remote sensing. Remote sensing provides spatially explicit, systematic, repeatable data over large areas, yet its widespread use for biodiversity monitoring has been constrained by factors such as cost, data availability, lack of capacity, and ineffective demonstration of the link to components of biodiversity. The goal of this research is to identify gaps in the current usage of remote sensing data in broad scale species richness models and to highlight opportunities to further leverage the information content of remotely sensed data for these models. We reviewed 38 studies that predicted species richness gradients at broad scales using one or more remotely sensed datasets. In the models we reviewed, remotely sensed estimates of land cover and greenness, especially the Normalized Difference Vegetation Index (NDVI) were commonly used predictors. Remotely sensed topographic data were rarely used, and remotely sensed measures of climate were never included in models. We highlight free, open-source, and analysis-ready remotely sensed data that may be used to predict species richness at broad scales, emphasizing which are available and appropriate for each of six global biogeographic realms. In highlighting the capabilities of geographic technologies for species richness modelling, this study will help bridge the gap between ecologists, managers and remote sensing scientists, and facilitate the advancement of biodiversity monitoring.

2.1 Introduction

The variability of species richness among global regions, and especially along latitudinal gradients, has been studied by ecologists for many years (Pianka 1966; Rohde 1992). Research to fully describe and explain broad-scale patterns with respect to both historical effects and the contemporary environment is ongoing (Beck et al. 2012). As detailed spatial distribution information is limited in quality and quantity for the majority of species on earth

(Jetz et al. 2012), researchers often use surrogates (such as ecosystems), or build a model relating covariates to generalized range maps or point observation data (Ferrier et al. 2004). Environmental covariates typically include some combination of climate, primary productivity, and habitat heterogeneity metrics (e.g., Field et al., 2008; Hawkins et al., 2003). These variables, in combination with data regarding historical factors such as glaciation, tectonics and geographic isolation, as well as biotic interactions, influence processes such as speciation, extinction, and dispersal, which in turn, affect which species will co-exist at any given location (Belmaker and Jetz 2012; Currie 1991; Qian et al. 2007; White and Hurlbert 2010). Global maps of species richness may be used to help set global biodiversity conservation goals (e.g., Myers et al 2000). Models of richness that incorporate environmental covariates may be also to assess the potential impacts of climate and land use change on predicted richness gradients over time, allowing goals to be adaptively managed and prioritized.

Environmental covariate or proxy datasets for species richness modelling are often acquired from multiple sources. Remotely sensed imagery is one source of data. Remote sensing provides spatially explicit, systematic, repeatable data over large areas, yet its widespread use for biodiversity monitoring in general has been constrained by factors such as cost and data availability (Secades et al. 2014; Skidmore et al. 2015), lack of capacity (Han et al. 2014; Pettorelli et al. 2014), and ineffective demonstration of the link to components of biodiversity (Buchanan et al. 2009; Walpole et al. 2009). The goal of this research is to identify gaps in the current usage of remote sensing data in broad scale species richness models and to highlight opportunities to further leverage the information content of remotely sensed data for these models. To meet this goal, we conduct a literature review to determine the breadth and the frequency with which remote sensing data have been utilized in global species diversity monitoring in recent years. Based on this review, we then identify and recommend remote sensing data and products that can capture the key environmental predictors of richness models, emphasizing those data which are currently underutilized. Our review builds upon recent discussions of the application of remote sensing for individual species distribution modeling (e.g., Bradley and Fleishman, 2008; Cord et al. 2013; He et al., 2015), and biodiversity monitoring in general (e.g., Duro et al. 2007; Nagendra, 2001; Nagendra et al., 2013; Turner et al., 2003) by providing specific and current product recommendations. In highlighting the capabilities of geographic technologies for species richness modelling, this study will help bridge the gap between ecologists, managers and remote sensing scientists, and facilitate the advancement of biodiversity monitoring.

2.2. Background

The environmental covariates used to predict species richness gradients at broad scales may be grouped into one of four categories: ambient energy, water or water energy, productive energy, and habitat heterogeneity. Ambient energy in a macroecological context refers to temperature-related variables, and can convey information related to climatic stability, harshness, and seasonality (Willig et al., 2003). Ambient energy may influence species richness gradients because of the physiological tolerance or intolerance to very hot or very cold temperatures, or by influencing speciation rates (Clarke and Gaston 2006; Currie et al. 2004; Evans et al. 2005; Wiens and Donoghue 2004). For migratory species, the effect of climatic seasonality on resource availability may be more influential than direct physiological effects on richness (Lennon et al. 2000). Ambient energy variables used in species richness models commonly include metrics of temperature and potential evapotranspiration (PET).

Ambient energy is thought to be particularly limiting to species richness at high latitudes, whereas water availability is more important at southern latitudes (Hawkins et al. 2003; Whittaker et al. 2006). Even so, water is essential for animals everywhere, with water having both a direct physiological role in animal survival and reproduction, as well as an indirect role through Net Primary Productivity, that is, through the provision of food resources. Broad scale species richness models commonly include metrics such as precipitation Actual Evapotranspiration (AET) to represent water and water-energy, respectively. Primary productivity is typically included as a separate variable in richness models even though it is correlated to AET. NPP is the rate at which energy is converted through photosynthesis or chemosynthesis to organic substances per unit time, minus the energy lost for respiration. Beginning with studies by Pianka (1966) and Wright (1983), the relationship between NPP and species richness continues to find wide-spread support (Mittelbach et al. 2001; Waide et al. 1999; Whittaker 2010). As reviewed in Currie et al. (2004) and Evans et al. (2005), primary productivity has been hypothesized to directly affect the number of individuals that an area can support, and thus in turn, species richness at any given location.

Variables representing habitat heterogeneity at broad scales such as the diversity of ecosystem or land cover types, or topographic complexity, are often included as predictors of species richness, despite often being considered less important at global scales than climate variables (Field et al. 2008; Whittaker et al. 2001). Habitat heterogeneity is thought to relate to species richness because species are associated with particular ecological niches, and thus more variable habitats are likely to host more species (Currie 1991). Topographic complexity has also been hypothesized to affect species richness from a historical perspective because, at least in the tropics where climate is generally more stable than in temperate areas, mountain ranges can act as barriers to dispersal, and the geographic isolation can lead to speciation (Janzen 1967). The topographic influence on species richness may be particularly important at tropical or subtropical latitudes where available energy is high (Kerr and Packer 1997; Qian 2010; Rahbek and Graves 2001). Topography is also a particularly important predictor of species richness in areas with high levels of endemism (Szabo et al. 2009), and especially where endemism is largely congruent with richness, as in sub-Saharan Africa (Jetz et al. 2004). In areas of low topographic relief, the relative influence of land cover diversity may be increased (Kerr et al. 2001).

2.3. Current uses of remote sensing in species richness modelling

We reviewed studies that model broad scale (continental to global) geographic gradients in species richness as a function of environmental covariates. To enable remote sensing related recommendations, all studies reviewed derived at least one environmental covariate from remote sensing. We constrained our review to studies using spatial units of analysis of at least 0.25° latitude x 0.25° longitude (~725 km^2 at the equator), a spatial resolution that not only allows, but indeed is often recommended (e.g., Hurlbert and White 2005; Rahbek 2004) for modelling across very large areas. Our focus was on animal species richness within the terrestrial realm, thus studies of plant diversity and aquatic or marine environments were excluded. A total of 38 studies were included in the review.

Based on a review of this literature, we compiled a list of remote sensing products that can capture the variables important for modelling species richness. There is emerging consensus that species richness-environment relationships vary regionally (Buckley and Jetz 2007; Davieset al. 2007; Hawkins 2010; Hortal et al. 2008; Kissling et al. 2012; Qian 2010), therefore we also provide recommendations at sub-global scales, using a framework of global biogeographic realms. The sensors we highlight have free and open-source data policies, and many are associated with pre-processed, ready to use products. Sensors and products suitable for global scale mapping are characterized by high temporal resolutions (e.g., daily to annually), medium to low spatial resolutions (>10 m), and medium to large image extents (>100 km^2). Higher spatial resolution imagery are generally only available from commercial vendors, and will often have limited spatial extents (many less than 20 km wide), and thus more individual images (with different dates and viewing angles) are required to cover large areas. Project complexity and cost are thus much higher (Wulder et al. 2008a). Therefore, although commercial sources of high spatial resolution remotely sensed data have been shown to be valuable when relating information in support of biodiversity studies (e.g., St-Louis et al., 2006; Viedma et al., 2012), the applicability of these data to global studies is often limited to provision of detailed calibration or validation information (Falkowski et al. 2009; Morisette et al. 2003).

2.4. Opportunities for remote sensing in species richness modelling

We found a narrow scope of remotely sensed data usage among the 38 studies reviewed (Table 2.1). Most studies utilized only two types of remotely sensed information: a vegetation index, most often the Normalized Difference Vegetation Index (NDVI) (e.g., Hawkins and DeVries, 2009; Qian, 2010; Ruggiero and Kitzberger, 2004), and land cover (e.g., Davies, et al., 2007; Kerr et al., 2001; Rahbek et al., 2007). All of the studies we reviewed used climate datasets (e.g., temperature, precipitation, AET and PET) derived from meteorological stations (e.g., WorldClim, Hijmans et al., 2005), rather than from spaceborne sensors. However, a few studies used annual variability of NDVI or NDVI within a particular season (e.g., Carnicer and Díaz-Delgado, 2008; Hurlbert and Haskell, 2003; Keil et al. 2008) to represent climatic seasonality or variability. Studies overwhelmingly used non-remotely sensed topographic data sources such as digitized contour maps, or other unspecified sources. Two studies used topography data that were exclusively remotely sensed based, for instance, from the Shuttle Radar Topography Mission (SRTM, see Table 2.2) (Carnicer and Díaz-Delgado 2008; Luo et al. 2012). In Sections 4.1 through 4.4, we highlight underutilized remotely sensed sources of climate, productivity, and habitat heterogeneity data that could be used in broad-scale models of species richness (Figure 2.1, Table 2.2).

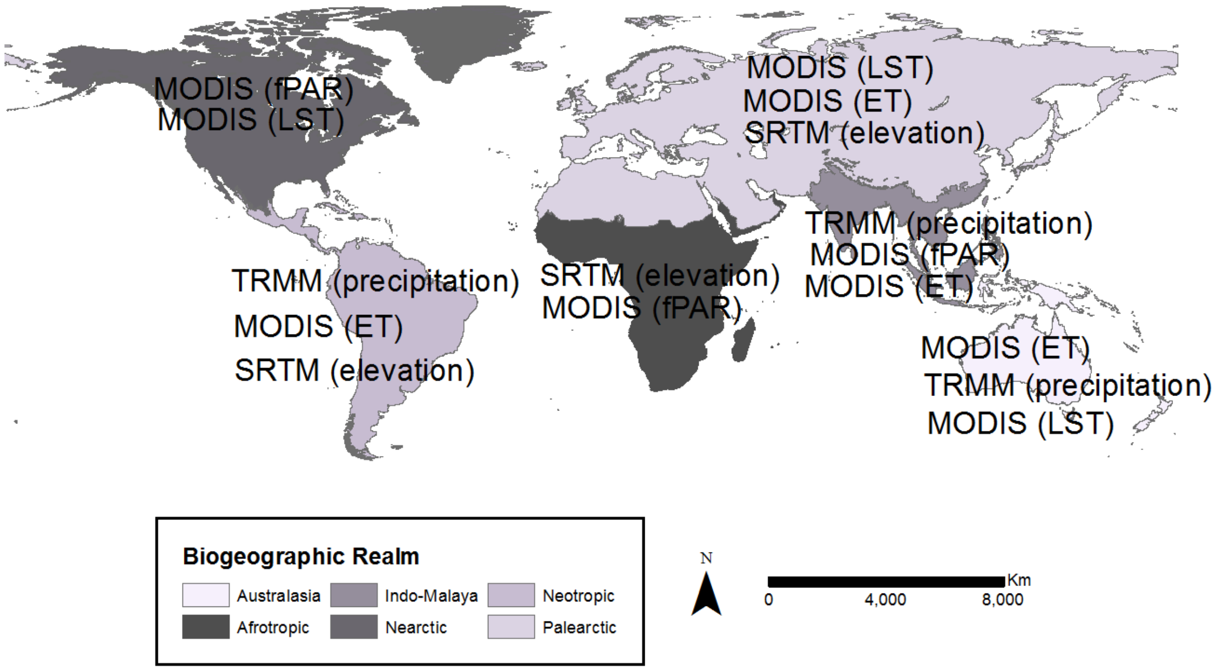


Figure 2.1. The relative importance of environmental predictors of species richness varies by global geographic realm. Our recommendations of remotely sensed data to capture some of the most important variables in each realm are indicated.

Table 2.1. Review of studies modelling the geographic distribution of species richness at continental to global scales, using one or more sources of remotely sensed data.

Study Reference	**Remotely Sensed Predictors**^{a}	**Other Predictors**^{a}
(Ballesteros-Mejia et al. 2013)	% herb cover and % tree cover	PET, AET, ELEVR
(Buckley and Jetz 2007)	Number of vegetation land cover classes	Annual AET, PET, mean annual temperature, mean annual diurnal temperature, frost-day frequency, annual precipitation, mean precipitation of the driest three months, wet-day frequency, relative humidity, mean annual NPP, mean three-month minimum NPP, number of elevation bands, realm
(Carnicer and Díaz-Delgado 2008)	Land cover type, mean June and mean December NDVI, seasonal range in NDVI, spatial variation in NDVI, interannual NDVI, ELEVR, ELEVMN	AET, PET, NPP, max and min June and Dec temps, daily range of June temps, annual rainfall, spatial variation of June and Dec max temps, lat and long, area, # of sampling routes
(Davies, Orme, Storch, et al. 2007)	Land cover diversity, agricultural land area, mean annual NDVI	Mean annual temp, ELEVR, human population density, glacial history, MDE
(Davies, Orme, Webster, et al. 2007)	Land cover diversity	NPP, AET, PET, ELEVR, mean annual temp, mean annual precip
(Diniz-Filho et al. 2013)	Land cover diversity	PET, AET, mean annual temp, mean annual precip, NPP, ELEVR, CV annual temp, PET of the driest month, # of dry months, min temp of the coldest month
(Diniz-Filho et al. 2004)	Land cover diversity	AET, PET, mean daily temp in the coldest month, annual mean temp, annual precip, ELEVR, ELEVR x min temp, genera age
(Diniz-Filho and Bini 2005)	Land cover diversity	PET, AET, mean annual temp, mean daily temp in coldest month, annual rainfall, ELEVR, ELEVR*min temp

(Fløjgaard et al. 2011)	Land cover evenness, NDVI	AET, PET, Mean annual temp, min temp, annual precip, water balance, GDD, long-term climatic stability, dispersal from ice age refugia, distance from Asia, ELEVR, human influence index
(H-Acevedo and Currie 2003)	Mean NDVI (Dec-Feb, May-July and annual)	Mean temp and total precip (Dec-Feb, and May-July), mean annual temp, total annual precip, mean temp x mean precip, ELEVR, soil fertility, land cover diversity
(Hawkins and Porter 2003a)	Mean annual and mean summer NDVI, land cover diversity	AET, PET, mean daily temp in the coldest month, mean annual temp, annual precip, ELEVR, plant species richness, age since glaciation
(Hawkins and Porter 2003b)	Land cover diversity	Annual AET, annual PET, mean monthly temp, mean January temp, annual precip, ELEVR, age since glaciation
(Hawkins et al. 2005)	Mean annual NDVI, land cover diversity	Annual AET, annual PET, ELEVR
(Hawkins et al. 2007)	Mean annual NDVI	Annual AET, mean annual temperature, elevation range
(Hawkins and DeVries, 2009)	Mean summer and annual NDVI	Annual AET, mean Jan and July temps and spatial range in mean annual temp, time since glaciation, biome type, distance from USA-Mexico border and southern Florida
(Hawkins 2010)	Mean annual NDVI	Maximum annual AET, mean annual temperature, local variability in mean annual temperatures, region
(Hortal et al. 2008)	Land cover diversity	AET, PET, mean, maximum and minimum monthly temperatures, mean annual precipitation, winter, spring, summer and autumn precipitation, water balance, elevation range, mean slope, aspect variability, habitat (biome) type, global realm, area
(Hurlbert and Haskell 2003)	Mean monthly NDVI, seasonal variability in NDVI (max-min and	ELEVR, number of biomes

	ratio of max-min/max)	
(Huang et al. 2011)	NDVI	AET, PET, mean annual max temp, mean annual min temp, mean sum of temp > 0°, mean sum of temp > 10°, mean annual solar radiation, mean annual temp, mean annual sunshine, mean annual diurnal temp range, annual frost day frequency, mean annual wind speed, mean annual wet day frequency, mean annual relative humidity, ELEVR, # of veg classes
(Jetz and Rahbek 2002)	Mean monthly NDVI, diversity of NDVI classes	Annual AET, annual PET, mean daily temp, annual monthly temp range, annual precip, annual precip range, NPP, mean solar radiation, ELEVM, ELEVR
(Keil et al. 2008)	Mean annual and mean summer NDVI, land cover diversity	AET, PET, mean annual temp, mean annual precip, ELEVR, plant species richness, age since glaciation, water surface area
(Kerr et al. 2001)	Land cover diversity	Mean, min, max and range each of AET, PET, water deficit, ELEVR, ELEVM, ELEVX, ELEVMN, sampling intensity
(Kerr et al. 2006)	NPP, land cover heterogeneity, forest extent, mean, minimum and maximum elevation	Mean, minimum and maximum temperature and precipitation, MDE
(Kissling et al. 2012)	Land cover diversity	AET, coefficient of variation of monthly NPP, mean annual temperature, annual precipitation, elevation range, biogeographic realm
(Kreft and Jetz 2007)	Land cover diversity	Annual AET, annual PET, mean annual NPP, mean annual temperature, precipitation, relative humidity, wind speed, # of wet and dry days, # of frost days, sunshine duration, water balance, growing months, number of 300 m elevation belts per unit, soil types, floristic kingdom, biome, latitude, area
(Luo et al. 2012)	Mean annual NDVI, ELEVR, slope, aspect, land cover diversity	Mean annual temp, annual precip, temp and precip seasonality, annual temp range, vegetation type diversity

(McPherson and Jetz 2007)	Mean annual temp, mean annual precip, annual CV precip, annual CV temp, land cover diversity	NPP, topographic heterogeneity
(Qian, Wang, et al. 2007)	Mean monthly NDVI and EVI	Mean annual temp, mean temp of coldest and warmest months, temp seasonality, mean annual precip and rainfall, monthly min and max precip, seasonality in precip, mean summer rainfall, annual AET and PET, monthly min and max AET and PET, seasonality in AET and PET, AET/PET, PET-AET, NPP, ELEVR, ELEVX, ELEVM, area
(Qian et al. 2009)	Mean annual NDVI	Annual AET and range in AET, annual PET and range in PET, annual mean, min and range of temp, annual and range of precip, ELEVM, ELEVR, area
(Qian 2010)	NDVI	AET, mean annual temp, annual precip, ELEVR, area
(Rahbek and Graves 2001)	Ecosystem diversity	Mean annual temp, mean daily max, min and range in temp, annual precip, ELEVR, solar radiation, wind speed, cloud cover, frost frequency, wet-day frequency, mean vapor pressure, lat, area
(Rahbek et al. 2007)	Ecosystem diversity	Mean annual temp, Mean annual precip, NPP, ELEVR, species-energy (area x NPP), water-energy (temp x precip x NPP), temp-kinetics, area
(Ruggiero and Kitzberger 2004)	Annually-Integrated NDVI, interannual variability in NDVI, spatial variability in NDVI	Annual AET, annual PET, annual min temp, yearly temp amplitude, ELEVR, proportion of continental land
(Rodríguez et al. 2005)	Mean annual NDVI, land cover diversity	AET, PET, mean annual temp, annual precip, SD of mean monthly temp and precip (Jan –July), NPP, plant biomass, ELEVR
(Storch et al. 2006)	Mean annual NDVI	Mean annual AET, mean annual NPP, MDE
(Szabo et al. 2009)	Annual NPP, land cover diversity	Annual minimum and mean temperature and temperature range, annual precipitation, elevation range

(Tello and Stevens 2010)	Elevation, NPP, and spatial variation of both	Mean annual temperature, annual precipitation, and spatial variation of both, monthly SD of temperature and coefficient of variation of precipitation
(Tognelli and Kelt 2004)	Max annual NDVI, ecosystem diversity	Annual AET, Annual PET, Mean daily max, min, and range of temp, mean annual temp, annual precip, solar radiation, frost frequency, mean vapor pressure, mean wind speed, ELEVR

AET: Actual Evapotranspiration; CV = coefficient of variation; ELEVM: minimum elevation; ELEVR: elevation range; ELEVX: maximum elevation; EVI: Enhanced Vegetation Index; fPAR: fraction of Photosynthetically Active Radiation; GDD: Growing degree days; LAI: Leaf Area Index; lat = latitude; long = longitude; MDE: Mid-Domain Effect; NDVI: Normalized Difference Vegetation Index; NPP: Net Primary Productivity; PET: Potential Evapotranspiration; SD = standard deviation; Q1 to Q4 = geographic range size quartile

[a] Topographic data were included in the "other variables" category unless known to be solely from remote sensing data (e.g., the Shuttle Radar Topography Mission, SRTM).

Table 2.2. Remotely sensed datasets suitable for predicting broad-scale species richness at continental and global scales

Predictor	Sensor	Data / Product	Coverage	Sensor Spatial Resolution	Sensor Temporal Resolution	Product Update Frequency	Duration
Temperature	MODIS	Land Surface Temperature	Global	1 km, 5 km, 0.5°	Daily	Daily, 8 day, monthly	2000 – present
	VIIRS	Land Surface Temperature	Global	750 m	2 days	Unknown	2013 -
AET	MODIS	Evapo-transpiration	Global	1 km	Daily	8 day, monthly, annual	2000 – 2010; will be updated in future
Precipitation	TRMM	Various	Tropical & Subtropical	0.25°, 0.5°, 5°	3 hours	3 hours, daily, monthly	1997 - present
	GPMM	Various	Global	< 100 km	3 hours	3 hours, daily	2014 -
	MODIS	Snow Cover	Global	500 m, 0.05°	Daily	Daily, 8 day, monthly	2002 - present
	AVHRR	EASE-Grid Snow Water Equivalent	Global	25 km	Daily	Monthly	1978-2007
Primary productivity	AVHRR	Global Vegetation Index	Global	1 km, 8 km, 16 km	Daily	Weekly, 10 day, monthly	1989 – present
	MERIS	Global Vegetation Index	Global	300 m, 1.2 km	3 day	Monthly	2002 – 2012
	MODIS	EVI, NDVI, fPAR, NPP, GPP	Global	250 m, 1km, 0.25°	Daily	8 day, 16 day, monthly	2000/2002 – present
	VIIRS	NDVI	Global	750 m	2 days	Unknown	2013 –
		MERIS FAPAR – EU					

Predictor	Sensor	Data / Product	Coverage	Sensor Spatial Resolution	Sensor Temporal Resolution	Product Update Frequency	Duration
Topography	SRTM	DEM	60°N to 60°S	30-90 m	N/A	V1 – 2003 V2 – 2005 V3 – 2006 V4 – 2008	No ongoing acquisition
	Various	GTOPO30	Global	1 km	N/A	First released in 1996	No ongoing acquisition
	ASTER	GDEM	83°N to 83°S	30 m	N/A	V1 – 2009 V2 – 2011	2000 – present
	Primarily SRTM	GMTED2010	Global	30, 15 and 7.5 arc-seconds	N/A	2010	will be updated in future
	TanDEM-X	Global Digital Elevation Model	Global	12 m	N/A	2014	2010 – present
Land Cover / Habitat structure	SPOT- VGT	Various, including GLC2000	Global	1 km	1-2 days	Daily, 10 day	1998 – present
	AVHRR	Various	Global	1 km, 8 km, 1°	Daily	Various	1981 – present
	MODIS	Land Cover	Global	500 m	Daily	Yearly	2001 – 2010
		Vegetation continuous fields	Global	250 m	Daily	Yearly	2000 - 2010
	MERIS	GlobCover	Global	300 m, 1.2 km	3 day	2005/2006, 2009	2002 – 2012
	Landsat MSS/TM/ETM+	Land cover	Global	30 m (TM / ETM+), 79 m (MSS)	16 days (TM / ETM+), 18 days (MSS)	Various	1972 – present

Predictor	**Sensor**	**Data / Product**	**Coverage**	**Sensor Spatial Resolution**	**Sensor Temporal Resolution**	**Product Update Frequency**	**Duration**
	Sentinel-2	Various	Global	10 to 60 m	5 days	Unknown	2014 –
	GEDI	Various	50°N to 50°S	25 m		Unknown	Likely 2019 – 2024
	GLAS ICESat	Various	Global	~70 m	91 days	Unknown	2003 – 2010
	ALOS PALSAR	forest cover	Global	10m, 100 m	46 days	Unknown	2006 – 2011

ASTER: *Advanced Spaceborne Thermal Emission and Reflection* Radiometer; ALOS: Advanced Land Observing Satellite; AVHRR: *Advanced Very High Resolution Radiometer*; GLAS ICESat: Geoscience Laser Altimeter System, Ice Cloud and land Elevation Satellite; MERIS: MEdium Resolution Imaging Spectrometer; MODIS: Moderate Resolution Imaging Spectroradiometer; SPOT: Système Pour l'Observation de la Terre; SRTM: Shuttle Radar Topography Mission; TRMM: Tropical Rainfall Monitoring Mission; GPMM: Global Precipitation Monitoring Mission; VIIRS: Visible Infrared Imaging Radiometer Suite; TM: Thematic Mapper; ETM+: Enhanced Thematic Mapper; MSS: Multispectral Scanner System; GTOPO30: Global 30 Arc-Second topographic data; GLC2000: Global Land Cover product: NPP: Net Primary Productivity; NDVI: Normalized Difference Vegetation Index; EVI: Enhanced Vegetation Index; fPAR: fraction of Photosynthetically Active Radiation; GPP: Gross Primary Productivity; DEM: Digital Elevation Model; VIIRS: Visible Infrared Imaging Radiometer Suite

2.4.1 Ambient Energy

Continental to global scale models of species richness predominately use datasets such as WorldClim (Hijmans et al. 2005), which are based on spatial interpolations of in-situ measurements. In such regions, such as the Afrotropics, the availability and quality of interpolated climate datasets is limited (Hay and Lennon, 1999; Hijmans et al., 2005). A major strength of remotely sensed data is the spatially continuous nature of data, meaning no interpolation is needed among climate stations. Therefore, in areas of rugged topography or sparse *in-situ* measurements, there are opportunities to use spatially continuous remotely sensed measurements of precipitation and temperature that may lead to more accurate climate datasets in comparison to interpolated climate surfaces, particularly as the amount of remotely sensed information increases over time (Hijmans et al. 2005). However, it is important to note that the temporal record of many remotely sensed climatic datasets is limited to the past 15-30 years (Table 2; Yang et al., 2013), whereas interpolated climate data based on in-situ measurements are typically averaged over a longer time period (e.g., 50 to 60 years).

Examples of remotely sensed climate data products that may be of value include the Moderate Resolution Imaging Spectroradiometer (MODIS) Land Surface Temperature (LST) data, which has global coverage (Table 2.2). LST data are correlated to traditional near-surface air temperature data measured at weather stations, but LST provides additional, complimentary information because it varies by land cover type and thus uniquely measures the coupling of ecological and climatological effects (Mildrexler et al. 2011). Temperature variables are highly explanatory of species richness at high latitudes (e.g., Davieset al., 2007; Hawkins and DeVries, 2009; Hawkins, 2010; Hawkins et al., 2003; Qian, 2010), and we recommend the use of LST for representing temperature in models within the Nearctic and Palearctic biogeographic realms (Figure 2.2). Future remotely sensed climatic products are also expected from the new Visible

Infrared Imaging Radiometer Suite (VIIRS), onboard the Sumoi National Polar-orbiting Partnership satellite launched in October 2011, which is designed to allow for continuity of MODIS-like products such as LST (Justice et al. 2013).

2.4.2 Water and Water Energy

Another remotely sensed, spatially continuous climatic product that may be used in broad-scale species richness models is the Global Precipitation Measurement (GPM) mission, which will provide near real-time precipitation information across the globe (Hou et al. 2013). The GPM will use both radar and microwave satellites, and will be capable of measuring light, moderate and heavy rainfall, as well as snowfall. The GPM is modelled after the Tropical Rainfall Monitoring Mission (TRMM). Since its launch in 1997, TRMM has been capturing daily precipitation data for tropical and subtropical regions using a combination of radar, visible, infrared and microwave imaging (Kummerow et al. 1998). Finally, 8-day, monthly, and annual estimates of global evapotranspiration are available from 2000 to 2013 from MODIS (Mu et al. 2011). Given the importance of water availability for determining species richness within the Neotropic, Afrotropic, Indo-Malayan and Australasian biogeographic realms (Hawkins 2010; Qian 2010; Tognelli and Kelt 2004), we recommend the use of TRMM or GPM data, or MODIS AET data for modeling within these realms (Figure 2.2). Models covering high latitudes, such as within the Palearctic or Nearctic global biogeographic realm may benefit from snow cover data available from MODIS.

2.4.3 Primary Productivity

The ability of remote sensing data to capture primary productivity is well-documented and considered operational (Zhao et al. 2005). There are several approaches to estimate primary productivity, using physical models or vegetation indices. Regarding vegetation index-based approaches, NDVI is often applied. The NDVI is the ratio of reflected red energy, which plant pigments such as chlorophyll absorb, and near-infrared energy, which is reflected by plants, thus expressing relative vegetation greenness or with a correlative relationship to productivity (Berry et al. 2007). Although NDVI is the most commonly used vegetation index (Kerr and Ostrovsky 2003), it is sensitive to background and atmospheric conditions and tends to saturate in areas such as the tropics with high biomass (Huete et al. 2002). Where possible, we recommend the use of the fraction of photosynthetically available radiation (fPAR) absorbed by vegetation, or the Enhanced Vegetation Index (EVI) instead of the more commonly used NDVI (Table 2.2). The EVI is a ratio-based vegetation index similar to the NDVI; however, the EVI incorporates additional terms such as reflectance in the blue band and a canopy background adjustment factor (Huete et al. 2011). These additional terms means EVI it less sensitive than NDVI to background and atmospheric conditions, and also less likely to saturate in conditions of dense vegetation because it is more sensitive to vegetation structure and less to chlorophyll (Huete et al., 2002). Vegetation indices like the NDVI and EVI correlate to fPAR, a biophysical parameter of vegetation that depends on vegetation type and structure. Several studies have shown a relationship between species richness and fPAR (e.g., Coops et al. 2009a; Coops et al. 2009b). fPAR can be estimated from NDVI, but depending on background soil conditions, the relationship is not always linear (Huete et al. 2011). It is also possible to generate fPAR values directly via a non-empirical physical modeling approach based upon an understanding of available energy and the fraction absorbed by vegetation. The global fPAR product from MODIS is modelled using surface reflectance and biome type, the latter defining certain vegetation structural characteristics (Myneni et al. 2002). MODIS products such as fPAR are composited over 8 or 16 days in order to minimize atmospheric effects. The new Sumoi National Polar-orbiting satellite carries sensors that will offer continuity of vegetation and productivity related datasets from MODIS (Justice et al. 2013). Across the Palearctic realm, the European Union provides registered users with access to fPAR data up to the year 2012 from the MERIS sensor (http://fapar.jrc.ec.europa.eu).

2.4.4 Habitat Heterogeneity

While the temporal extent of global remotely sensed climatic products may be somewhat limited, many satellites have been collecting land surface reflectance data for considerably longer (>40 years; Table 2.2). Moreover, remote sensing data is associated with frequent, continuous acquisition. These characteristics, together with the ability to acquire data in remote regions, make remote sensing well-suited to measuring habitat heterogeneity. For instance, a model that incorporates remotely sensed land cover data could be used to regularly monitor potential anthropogenic influences on predicted patterns of species richness.

A number of different global land cover classifications have been produced over the past two decades using remote sensors such as MODIS, Advanced Very High Resolution Radiometer (AVHRR), and Système *Pour* l'*Observation de la Terre* (SPOT) (Bartholomé and Belward 2005; Friedl et al. 2010; Hansen et al. 2000; Loveland et al. 2000). Unfortunately, inconsistent data sources, methodologies, and spatial or thematic resolutions make these data products difficult to compare (Herold et al. 2008). Further, these global datasets poorly distinguish landscapes with a mix of vegetation types and densities, and small and relatively rare features such as wetlands and urban areas (Herold et al. 2008). Landsat imagery, with 30 m spatial resolution, may be preferred in this regard. The Landsat series of sensors have been in continuous operation since 1972, making it the longest running record of any similar earth observation satellite (Wulderet al. 2012a). However, until recently, the use of Landsat was uncommon for large area mapping, relative to sensors like MODIS, because of the considerable effort and cost required for collecting and processing large numbers of individual images associated with the higher spatial resolution of Landsat (Townshend et al. 2012). As of 2008, all of the USGS's new and archived Landsat imagery were made freely available to the public (Wulderet al. 2012a). This change of policy combined with improved computing power has made it now feasible to use Landsat to create global land cover and land cover change products (Chen et al. 2015; Gong et al. 2013; Hansen et al. 2013; Townshend et al. 2012). Also, advanced algorithms for cloud and haze removal (e.g., Masek et al., 2006; Zhu and Woodcock, 2012), and for selecting the best available image or pixel from among multiple, spatially overlapping images acquired at different dates (e.g., Griffiths et al., 2013; Roy et al., 2010; White et al., 2014), enhance the utility of Landsat and other optical data in persistently cloudy tropical regions.

Other remotely sensed data that may be used to characterize habitat heterogeneity include RADAR (RAdio Detection And Ranging) data, which emits and receives cloud-penetrating microwave energy, and LIDAR (LIght Detection And Ranging) data, which emits and receives optical energy. As discussed in a recent review by Bergen et al. (2009), LIDAR and RADAR systems can capture information regarding the structural complexity of vegetation, and recent spaceborne sensors provide broad scale data to inform on vertical vegetation characteristics important for global biodiversity conservation. A recent product that could potentially be useful for species richness models within the Neotropic realms for example, is a RADAR-based, circa 2007, pan-tropical forest cover map produced using PALSAR data from the ALOS-1 satellite (Woods Hole Reserach Centre 2012). The map distinguishes dense forest, open forest, flooded forests, sparse vegetation, and non-vegetated surfaces, and could be updated in future after the upcoming launch of the ALOS-2 satellite.

Also of potential interest for species richness modeling is space-borne LIDAR. The Geoscience Laser Altimeter System (GLAS), on-board the Ice, Cloud and land Elevation Satellite (ICESat) was used by Lefsky (2010) and Simard et al. (2011) to produce global gridded maps of vegetation height. Bolton et al. (2013) investigate these global products and offer some cautions regarding regional representation and the use of global vegetation height products. The ICESat mission ended in 2010, with the last laser data acquisition in October 2009. ICESat-2 is scheduled for launch in mid-2016, although different sensor technology means the potential and limits of the data for biodiversity applications remains to be determined. While limited to the orbit of the International Space Station, the recently announced Global Ecosystem Dynamics Investigation (GEDI) will provide LIDAR data over much of the non-boreal globe beginning in 2019. GEDI is designed to capture 3D vegetation structure with a spatial resolution of approximately 25 m. However, with coverage planned for 50°S to 50°N, the GEDI may not be applicable for global, nor continental-scale species richness models in realms such as the Nearctic and Palearctic.

As further global-scale operational products of LIDAR and RADAR data are developed, their utility for broad-scale biodiversity models can be further investigated. For now, data from various remote sensors can be integrated to maximize the spatial, spectral and temporal resolution for a particular application. For example, LIDAR data may be used as a sampling tool within a larger area captured by moderate to low resolution optical or radar imagery, rather than aiming for wall-to-wall coverage of the more costly, highly detailed, and spatially limited, types of data (Bergen et al. 2009; Wulder et al. 2012b).

Studies reviewed generally used one of two sources of topographic information for modelling species richness: the U.S. Geological Survey’s (USGS) Global 30 Arc Second

Elevation Data Set (GTOPO30), or WorldClim (Hijmans et al. 2005), which is based on data from the Shuttle Radar Topography Mission (SRTM) and the GTOPO30. The SRTM data are offered at a spatial resolution of 30 m in the United States and 90 m for the rest of the world, between approximately 60°N and 60°S. In the WorldClim dataset, the SRTM data were aggregated to 1 km spatial resolution. The GTOPO30 was produced in 1996 from a variety of compiled data sources and covers the entire globe at a spatial resolution of 1 km. Recently, the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) was produced by the USGS and the National Geospatial-Intelligence Agency to replace the GTOPO30. It incorporates elevation data from the SRTM that were not available when the GTOPO30 was produced in 1996 (Danielson and Gesch 2011). Another free, global topographic dataset is the *Advanced Spaceborne Thermal Emission and Reflection* Radiometer (ASTER) Global Digital Elevation Model (GDEM2). The GDEM2 provides global topographic data for almost the entire globe (approximately 83°N to 83°S) and at a higher spatial resolution (30 m) than both the SRTM and GTOPO30 data. However, in some areas at least, the vertical accuracy of ASTER GDEM2 is lower than that SRTM products (Athmania and Achour 2014; Gesch et al. 2012). As well, there are some artifacts in the ASTER v2 product in and around water bodies, and also a fair amount of noise (fine-scale variability) (Tachikawa et al. 2011). The recently-released (2015) WorldDEM products with data from the TanDEM-X Mission, provide elevation data with (soon-to-be) complete global coverage with 12 m spatial resolution; however, these data are not available free of charge.

2.5 Conclusion

Continental and global scale maps of species richness enable global biodiversity conservation goals to be set, assessed, and prioritized. One approach of producing these maps is to model richness as a function of environmental variables including climate, primary productivity and habitat heterogeneity. Currently, remote sensing is not a major source of data used in these models. For instance, the predominant source of climate and topographic variables are interpolated surfaces from *in-situ* measures and contour maps, respectively. These may be of low quality in remote or topographically complex areas. Remote sensing can also provide spatially explicit, quantitative and repeatable estimates of land surface temperature, precipitation, evapotranspiration, primary productivity, topography and land cover, and these are readily available for modelling species richness at broad scales. As well, technological advances are allowing land and vegetation cover to be mapped with remotely sensed data at ever-increasing spatial resolutions over large spatial extents. The high spatial and temporal resolution of many remotely sensed datasets will be valuable for assessing finer-scale spatial heterogeneity in species richness gradients, and also for assessing threats to richness hotspots from global change and anthropogenic disturbance. Further, while each global region may vary in terms of the most important data inputs, and national interests may vary in terms of the desired use and representation of the end models, remotely sensed products that are systematically collected at global extents and freely available to all, allow for some standardization of models from region to region. Ongoing challenges and uncertainties associated with the integration of remotely sensed data in species richness models include the limited temporal extent of many of the climate related datasets, and often a need for advanced image processing skills. We predict the use of remote sensing in broad scale species richness models will only increase and become more important as the archive of imagery grows, and as non-specialists become more aware of the capabilities of remote sensing for biodiversity monitoring through efforts such as this study.

2.6 References

Athmania, D., & Achour, H. (2014). External Validation of the ASTER GDEM2, GMTED2010 and CGIAR-CSI- SRTM v4.1 Free Access Digital Elevation Models (DEMs) in Tunisia and Algeria. *Remote Sensing*, *6*(5), 4600–4620. doi:10.3390/rs6054600

Ballesteros-Mejia, L., Kitching, I. J., Jetz, W., Nagel, P., & Beck, J. (2013). Mapping the biodiversity of tropical insects: species richness and inventory completeness of African sphingid moths. *Global Ecology and Biogeography*, *22*(5), 586–595. doi:10.1111/geb.12039

Bartholomé, E., & Belward, A. S. (2005). GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*, *26*(9), 1959–1977. doi:10.1080/01431160412331291297

Bergen, K. M., Goetz, S. J., Dubayah, R. O., Henebry, G. M., Hunsaker, C. T., Imhoff, M. L., et al. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *Journal of Geophysical Research*, *114*, G00E06. doi:10.1029/2008JG000883

Berry, S., Mackey, B., & Brown, T. (2007). Potential applications of remotely sensed vegetation greeness to habitat analysis and the conservation of dispersive fauna. *Pacific Conservation Biology*, *13*, 120–127.

Bolton, D. K., Coops, N. C., & Wulder, M. A. (2013). Investigating the agreement between global canopy height maps and airborne Lidar derived height estimates over Canada. *Canadian Journal of Remote Sensing*, *39*(S1), 139–151. doi:10.5589/m13-036

Bradley, B. A., & Fleishman, E. (2008). Can remote sensing of land cover improve species distribution modelling? *Journal of Biogeography*, *35*(7), 1158–1159. doi:10.1111/j.1365-2699.2008.01928.xBuckley, L. B., & Jetz, W. (2007). Environmental and historical constraints on global patterns of amphibian richness. *Proceedings of the Royal Society B: Biological Sciences, 274*(1614), 1167–73. doi:10.1098/rspb.2006.0436

Carnicer, J., & Díaz-Delgado, R. (2008). Geographic differences between functional groups in patterns of bird species richness in North America. *Acta Oecologica*, *33*(2), 253–264. doi:10.1016/j.actao.2007.12.001

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., et al. (2015). Global land cover mapping at 30m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, *103*, 7–27. doi:10.1016/j.isprsjprs.2014.09.002

Coops, N. C., Waring, R. H., Wulder, M. A., Pidgeon, A. M., & Radeloff, V. C. (2009a). Bird diversity: a predictable function of satellite-derived estimates of seasonal variation in canopy light absorbance across the United States. *Journal of Biogeography*, *36*(5), 905–918. doi:10.1111/j.1365-2699.2008.02053.x

Coops, N. C., Wulder, M. A., & Iwanicka, D. (2009). Exploring the relative importance of satellite-derived descriptors of production, topography and land cover for predicting breeding bird species richness over Ontario, Canada. *Remote Sensing of Environment*, *113*(3), 668–679. doi:10.1016/j.rse.2008.11.012

Cord, A. F., Meentemeyer, R. K., Leitão, P. J., & Václavík, T. (2013). Modelling species distributions with remote sensing data: bridging disciplinary perspectives. *Journal of Biogeography*, *40*(12), 2226–2227. doi:10.1111/jbi.12199Currie, D. J., (1991). Energy and large-scale patterns of animal-and plant-species richness. *The American Naturalist*. 137(1), 27–49.

Currie, D. J., Mittelbach, G. G., Cornell, H. V., Field, R., Guegan, J.F., Hawkins, B. A., et al. (2004). Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters*, 7(12), 1121–1134. doi: 10.1111/j.1461-0248.2004.00671.x

Danielson, J. J., & Gesch, D. B. (2011). *Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010): U.S. Geological Survey Open-File Report 2011-1073*. U.S. Geological Survey, Reston, Virginia.

Davies, R. G., Orme, C. D. L., Storch, D., Olson, V. A., Thomas, G., Ross, S., et al. (2007). Topography, energy and the global distribution of bird species richness. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1606), 103–10. doi:10.1098/rspb.2006.0061

Davies, R. G., Orme, C. D. L., Webster, A. J., Jones, K. E., Blackburn, T. M., & Gaston, K. J. (2007). Environmental predictors of global parrot (Aves: Psittaciformes) species richness and phylogenetic diversity. *Global Ecology and Biogeography*, *16*(2), 220–233. doi:10.1111/j.1466-8238.2007.00282.x

Diniz-Filho, J. A. F., & Bini, L. M. (2005). Modelling geographical patterns in species richness using eigenvector-based spatial filters. *Global Ecology and Biogeography*, *14*(2), 177–185. doi:10.1111/j.1466-822X.2005.00147.x

Diniz-Filho, J. A. F., Ceccarelli, S., Hasperué, W., & Rabinovich, J. (2013). Geographical patterns of Triatominae (Heteroptera: Reduviidae) richness and distribution in the Western Hemisphere. *Insect Conservation and Diversity*, *6*(6), 704–714. doi:10.1111/icad.12025

Diniz-Filho, J. A. F., Rangel, T. F. L. V. B., & Hawkins, B. A. (2004). A test of multiple hypotheses for the species richness gradient of South American owls. *Oecologia*, *140*(4), 633–8. doi:10.1007/s00442-004-1577-4

Duro, D. C., Coops, N. C., Wulder, M. A., & Han, T. (2007). Development of a large area biodiversity monitoring system driven by remote sensing. *Progress in Physical Geography, 31*(3), 235–260. doi:10.1177/0309133307079054

Evans, K.L., Warren, P.H., Gaston, K.J., (2005). Species-energy relationships at the macroecological scale: a review of the mechanisms. *Biological reviews of the Cambridge Philosophical Society*, 80(1), 1–25. doi: 10.1017/S1464793104006517

Ferrier, S., Powell, G. V. N., Richardson, K. S., Manion, G., Overton, J. M., Allnutt, T. F., et al. (2004). Mapping more of terrestrial biodiversity for global conservation assessment. *Bioscience*, 54(12), 1101–1109.

Field, R., Hawkins, B. A., Cornell, H. V., Currie, D. J., Diniz-Filho, J. A. F., Guégan, J.-F. et al. (2008). Spatial species-richness gradients across scales: a meta-analysis. *Journal of Biogeography*, 36(1), 132–147. doi: 10.1111/j.1365-2699.2008.01963.x

Fløjgaard, C., Normand, S., Skov, F., & Svenning, J.-C. (2011). Deconstructing the mammal species richness pattern in Europe - towards an understanding of the relative importance of climate, biogeographic history, habitat heterogeneity and humans. *Global Ecology and Biogeography, 20*(2), 218–230. doi:10.1111/j.1466-8238.2010.00604.x

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*(1), 168–182. doi:10.1016/j.rse.2009.08.016

Gesch, D., Oimoen, M., Zhang, Z., Meyer, D., & Danielson, J. (2012). Validation of the Aster Global Digital Elevation Model Version 2 Over the Conterminous United States. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XXXIX-B4*, 281–286.

Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., et al. (2013). Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing*, *34*(7), 2607–2654. doi:10.1080/01431161.2012.748992

Griffiths, P., Linden, S. Van Der, Kuemmerle, T., & Hostert, P. (2013). A pixel-based Landsat compositing algorithm for large area land cover mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1–14. doi:10.1109/JSTARS.2012.2228167

H-Acevedo, D., & Currie, D. J. (2003). Does climate determine broad-scale patterns of species richness? A test of the causal link by natural experiment. *Global Ecology and Biogeography*, *12*(6), 461–473. doi:10.1046/j.1466-822X.2003.00058.x

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, *21*(6-7), 1331–1364. doi:10.1080/014311600210209

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., et al. (2013). High-resolution global maps of 21st-Century forest cover change. *Science*, *342*(6160), 850–853. doi:10.1126/science.1244693

Hawkins, B. A. (2010). Multiregional comparison of the ecological and phylogenetic structure of butterfly species richness gradients. *Journal of Biogeography*, *37*(4), 647–656. doi:10.1111/j.1365-2699.2009.02250.x

Hawkins, B. A., & DeVries, P. J. (2009). Tropical niche conservatism and the species richness gradient of North American butterflies. *Journal of Biogeography*, *36*(9), 1698–1711. doi:10.1111/j.1365-2699.2009.02119.x

Hawkins, B. A., Diniz-Filho, J. A. F., Jaramillo, C. A., & Soeller, S. A. (2007). Climate, niche conservatism, and the global bird diversity gradient. *The American naturalist*, *170 Suppl*, S16–27. doi:10.1086/519009

Hawkins, B. A., Diniz-Filho, J. A. F., & Soeller, S. A. (2005). Water links the historical and contemporary components of the Australian bird diversity gradient. *Journal of Biogeography*, *32*(6), 1035–1042. doi:10.1111/j.1365-2699.2004.01238.x

Hawkins, B. A., Field, R., Cornell, H. V., Currie, D. J., Guégan, J. F., Kaufman, D. M., et al. (2003). Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, *84*(12), 3105–3117. doi:http://dx.doi.org/10.1890/03-8006

Hawkins, B. A., & Porter, E. (2003a). Water–energy balance and the geographic pattern of species richness of western Palearctic butterflies. *Ecological Entomology*, *28*, 678–686.

Hawkins, B. A., & Porter, E. E. (2003b). Relative influences of current and historical factors on mammal and bird diversity patterns in deglaciated North America. *Global Ecology and Biogeography*, *12*(6), 475–481. doi:10.1046/j.1466-822X.2003.00060.x

He, K. S., Bradley, B. A., Cord, A. F., Rocchini, D., Tuanmu, M., Schmidtlein, S., et al. (2015). Will remote sensing shape the next generation of species distribution models? *Remote Sensing in Ecology and Conservation*, *1*(1), 4–18. doi:10.1002/rse2.7Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A., & Schmullius, C. (2008). Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment*, *112*(5), 2538–2556. doi:10.1016/j.rse.2007.11.013

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. doi:10.1002/joc.1276

Hortal, J., Rodríguez, J., Nieto-Díaz, M., & Lobo, J. M. (2008). Regional and environmental effects on the species richness of mammal assemblages. *Journal of Biogeography, 35*(7), 1202–1214. doi:10.1111/j.1365-2699.2007.01850.x

Hou, A., Kakar, R., Neeck, S., Azarbarzin, A., Kummerow, C., Kojima, M., et al. (2013). The Global Precipitation Measurement (GPM) Mission. *Bulletin of the American Meterological Society*.

Huang, Y., Dai, Q., Chen, Y., Wan, H., Li, J., & Wang, Y. (2011). Lizard species richness patterns in China and its environmental associations. *Biodiversity and Conservation, 20*(7), 1399–1414. doi:10.1007/s10531-011-0033-0

Huete, A. R., Didan, K., Miura, T., & Rodriguez, E. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment, 83*, 195–213. doi:10.1016/S0034-4257(02)00096-2

Huete, A. R., Didan, K., van Leeuwen, W., Miura, T., & Glenn, E. (2011). MODIS Vegetation Indices. In B. Ramachandran (Ed.), *Land Remote Sensing and Global Environmental Change, Remote Sensing and Digital Image Processing* (pp. 579–602). Springer. doi:10.1007/978-1-4419-6749-7

Hurlbert, A. H., & Haskell, J. P. (2003). The effect of energy and seasonality on avian species richness and community composition. *The American Naturalist*, *161*(1), 83–97. doi:10.1086/345459

Jetz, W., & Rahbek, C. (2002). Geographic range size and determinants of avian species richness. *Science*, *297*(5586), 1548–1551. doi:10.1126/science.1072779

Jetz, W., Rahbek, C., Colwell, R. K., 2004. The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecology Letters*. 7, 1180–1191. doi: 10.1111/j.1461-0248.2004.00678.x

Jetz, W., McPherson, J. M., Guralnick, R. P. (2012). Integrating biodiversity distribution knowledge: toward a global map of life. *Trends in Ecology & Evolution*, 27(3), 151–159. doi: 10.1016/j.tree.2011.09.007

Justice, C. O., Román, M. O., Csiszar, I., Vermote, E. F., Wolfe, R. E., Hook, S. J., et al. (2013). Land and cryosphere products from Suomi NPP VIIRS: Overview and status. *Journal of Geophysical Research: Atmospheres*, *118*(17), 9753–9765. doi:10.1002/jgrd.50771

Keil, P., Simova, I., & Hawkins, B. A. (2008). Water-energy and the geographical species richness pattern of European and North African dragonflies (Odonata). *Insect Conservation and Diversity*, *1*, 142–150.

Kerr, J. T., & Ostrovsky, M. (2003). From space to species: ecological applications for remote sensing. *Trends in Ecology & Evolution*, *18*(6), 299–305. doi:10.1016/S0169-5347(03)00071-5

Kerr, J. T., Perring, M., & Currie, D. J. (2006). The missing Madagascan mid-domain effect. *Ecology Letters*, *9*(2), 149–159. doi:10.1111/j.1461-0248.2005.00860.x

Kerr, J. T., Southwood, T. R., & Cihlar, J. (2001). Remotely sensed habitat diversity predicts butterfly species richness and community similarity in Canada. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(20), 11365–11370. doi:10.1073/pnas.201398398

Kissling, W. D., Sekercioglu, C. H., & Jetz, W. (2012). Bird dietary guild richness across latitudes, environments and biogeographic regions. *Global Ecology and Biogeography*, *21*(3), 328–340. doi:10.1111/j.1466-8238.2011.00679.x

Kreft, H., & Jetz, W. (2007). Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(14), 5925–30. doi:10.1073/pnas.0608361104

Kummerow, C., Barnes, W., Kozu, T., Shiue, J., & Simpson, J. (1998). The Tropical Rainfall Measuring Mission (TRMM) Sensor Package. *Journal of Atmospheric and Oceanic Technology*, *15*(3), 809–817. doi:10.1175/1520-0426(1998)015<0809:TTRMMT>2.0.CO;2

Lefsky, M. A. (2010). A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophysical Research Letters*, *37*(15), L15401. doi:10.1029/2010GL043622

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., & Merchant, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, *21*(6), 1303–1330.

Luo, Z., Tang, S., Li, C., Fang, H., Hu, H., Yang, J., et al. (2012). Environmental effects on vertebrate species richness: testing the energy, environmental stability and habitat heterogeneity hypotheses. *PloS ONE*, *7*(4), e35514. doi:10.1371/journal.pone.0035514

Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., et al. (2006). A Landsat surface reflectance dataset for North America, 1990-2000. *IEEE Geoscience and Remote Sensing Letters*, *3*(1), 68–72.

McPherson, J. M., & Jetz, W. (2007). Type and spatial structure of distribution data and the perceived determinants of geographical gradients in ecology: the species richness of African birds. *Global Ecology and Biogeography*, *16*(5), 657–667. doi:10.1111/j.1466-8238.2007.00318.x

Mildrexler, D. J., Zhao, M., & Running, S. W. (2011). A global comparison between station air temperatures and MODIS land surface temperatures reveals the cooling role of forests. *Journal of Geophysical Research*, *116*(G3), G03025. doi:10.1029/2010JG001486

Mittelbach, G. G., Steiner, C. F., Scheiner, S. M., Gross, K. L., Reynolds, H. L., Waide, R. B., et al. (2001). What is the observed relationship between species richness and productivity? *Ecology*, *82*(9), 2381–2396. http://dx.doi.org/10.1890/0012-9658(2001)082[2381:WITORB]2.0.CO;2

Morisette, J.T., Nickeson, J.E., Davis, P., Wang, Y., Tian, Y., Woodcock, C.E., Shabanov, N., Hansen, M., Cohen, W.B., Oetter, D.R., Kennedy, R.E. (2003). High spatial resolution satellite observations for validation of MODIS land products: IKONOS observations acquired under the NASA Scientific Data Purchase. *Remote Sensing of Environment*, 88 (1-2), 100–110. doi:10.1016/j.rse.2003.04.003.

Mu, Q., Zhao, M., & Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, *115*(8), 1781–1800. doi:10.1016/j.rse.2011.02.019

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A, Kent, J., (2000). Biodiversity hotspots for conservation priorities. *Nature* 403, 853–8. doi:10.1038/35002501

Myneni, R., Hoffman, S., Knyazikhin, Y., Privette, J., Glassy, J., Tian, Y., et al. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment*, *83*, 214–231. doi:10.1016/S0034-4257(02)00074-3

Nagendra, H. (2001). Using remote sensing to assess biodiversity. *International Journal of Remote Sensing*, *22*(12), 2377–2400. Nagendra, H., Lucas, R., Honrado, J. P., Jongman, R. H. G., Tarantino, C., Adamo, M., & Mairota, P. (2013). Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecological Indicators*, *33*, 45–59. doi:10.1016/j.ecolind.2012.09.014

Qian, H. (2010). Environment–richness relationships for mammals, birds, reptiles, and amphibians at global and regional scales. *Ecological Research*, *25*(3), 629–637. doi:10.1007/s11284-010-0695-1

Qian, H., Wang, S., Li, Y., & Wang, X. (2009). Breeding bird diversity in relation to environmental gradients in China. *Acta Oecologica*, *35*(6), 819–823. doi:10.1016/j.actao.2009.08.009

Qian, H., Wang, X., Wang, S., & Li, Y. (2007). Environmental determinants of amphibian and reptile species richness in China. *Ecography*, *30*(4), 471–482. doi:10.1111/j.2007.0906-7590.05025.x

Rahbek, C., Gotelli, N. J., Colwell, R. K., Entsminger, G. L., Rangel, T. F. L. V. B., & Graves, G. R. (2007). Predicting continental-scale patterns of bird species richness with spatially explicit models. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1607), 165–174. doi:10.1098/rspb.2006.3700

Rahbek, C., & Graves, G. R. (2001). Multiscale assessment of patterns of avian species richness. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(8), 4534–4539. doi:10.1073/pnas.071034898

Rodríguez, M. Á., Belmontes, J. A., & Hawkins, B. A. (2005). Energy, water and large-scale patterns of reptile and amphibian species richness in Europe. *Acta Oecologica*, *28*(1), 65–70. doi:10.1016/j.actao.2005.02.006

Roy, D. P., Ju, J., Kline, K., Scaramuzza, P. L., Kovalskyy, V., Hansen, M., et al. (2010). Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, *114*(1), 35–49. doi:10.1016/j.rse.2009.08.011

Ruggiero, A., & Kitzberger, T. (2004). Environmental correlates of mammal species richness in South America: effects of spatial structure, taxonomy and geographic range. *Ecography*, *27*, 401–416. doi:10.1111/j.0906-7590.2004.03801.x

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research*, *116*(G4), G04021. doi:10.1029/2011JG001708

Storch, D., Davies, R. G., Zajícek, S., Orme, C. D. L., Olson, V. A., Thomas, G. H., et al. (2006). Energy, range dynamics and global species richness patterns: reconciling mid-domain effects and environmental determinants of avian diversity. *Ecology Letters*, *9*(12), 1308–20. doi:10.1111/j.1461-0248.2006.00984.x

Szabo, N. D., Algar, A. C., & Kerr, J. T. (2009). Reconciling topographic and climatic effects on widespread and range-restricted species richness. *Global Ecology and Biogeography*, *18*(6), 735–744. doi:10.1111/j.1466-8238.2009.00479.x

Tachikawa, T., Kaku, M., Iwasaki, A., Gesch, D., Oimoen, M., Zhang, Z., et al. (2011). *ASTER Global Digital Elevation Model Version 2 – Summary of Validation Results*. (D. Meyer, Ed.).

Tello, J. S., & Stevens, R. D. (2010). Multiple environmental determinants of regional species richness and effects of geographic range size. *Ecography*, *33*(4), 796–808. doi:10.1111/j.1600-0587.2009.05991.x

Tognelli, M., & Kelt, D. (2004). Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. *Ecography*, *27*(4), 427–436.

Townshend, J. R. G., Masek, J. G., Huang, C., Vermote, E. F., Gao, F., Channan, S., et al. (2012). Global characterization and monitoring of forest cover using Landsat data: opportunities and challenges. *International Journal of Digital Earth*, *5*(5), 373–397. doi:10.1080/17538947.2012.713190

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., & Steininger, M. (2003). Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution*, *18*(6), 306–314. doi:10.1016/S0169-5347(03)00070-3

Waide, R., Willig, M., Steiner, C., Mittelbach, G.G., Gough, L., Dodson, S., Juday, G., Parmenter, R. (1999). The relationship between productivity and species richness. *Annual review of Ecology and Systematics*, 30, 257–300. doi: 10.1146/annurev.ecolsys.30.1.257

White, J. C., Wulder, M. A., Hobart, G. W., Luther, J. E., Hermosilla, T., Griffiths, P., et al. (2014). Pixel-based image compositing for large-area dense time series applications and science. *Canadian Journal of Remote Sensing*, *43*(3), 192–212. doi:10.1080/07038992.2014.945827

Woods Hole Reserach Centre. (2012). *Pan-tropical Forest Cover*. *Pan-tropical Forest Cover*. http://www.whrc.org/mapping/pantropical/alos.html

Wulder M. A., White J. C., Goward S. N., Masek J. G., Irons J. R., Herold M., et al. (2008). Landsat continuity: Issues and opportunities for land cover monitoring. *Remote Sensing of Environment,* 112(3), 955–969. doi: 10.1016/j.rse.2007.07.004

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., & Woodcock, C. E. (2012a). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, *122*, 2–10. doi:10.1016/j.rse.2012.01.010

Wulder, M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., et al. (2012b). Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment*, *121*, 196–209. doi:10.1016/j.rse.2012.02.001

Yang, J., Gong, P., Fu, R., Zhang, M., Chen, J., Liang, S., et al. (2013). The role of satellite remote sensing in climate change studies. *Nature Climate Change*, *3*(10), 875–883. doi:10.1038/nclimate1908

Zhao, M., Heinsch, F. A., Nemani, R. R., & Running, S. W. (2005). Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment*, *95*(2), 164–176. doi:10.1016/j.rse.2004.12.011

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, *118*, 83–94. doi:10.1016/j.rse.2011.10.028

3. Mapping dominant tree species over large forested areas using Landsat best-available-pixel image composites

Abstract

Remotely sensed image composites that are pixel-, rather than, scene-based, are increasingly feasible over large areas and fine spatial resolutions. For large jurisdictions that utilize remotely sensed imagery for ecosystem mapping and monitoring, pixel-based composites enable a wider range of applications, at higher quality. The goal of this study was to model spatial distributions of six tree species over a large forested area of Saskatchewan, Canada (>39 million ha) at 30 m spatial resolution using a multi-year Best-Available-Pixel (BAP) Landsat composite. We tested the influence of the BAP composite on the resultant maps by comparing species composition and configuration for areas where imagery was from a single sensor, year, and day of year, to areas with variable composite characteristics. Model error rates ranged from 9 % to 24 %, Area-Under-the-Curve values approaching 1, and met ecological expectations. The BAP composite was found to have little effect on model outcomes, with composition and configuration values in non-reference areas being similar for all species but one, which had an unexpected configuration. Moreover, sensor, year, and day of year were similar for reference and non-reference blocks for all species. Results indicate Landsat BAP image composites are useful for generating large-area maps of tree species distributions.

3.1 Introduction

Effective forest management requires knowledge of the spatial distribution of tree species composition and abundance. Species information is used to assess risks and impacts associated with a variety of natural or anthropogenic disturbances, including fires, insect and invasive plant infestations, and resource extraction. Species composition and abundance, together with other metrics such as species richness, species endemism, and rarity are also important metrics of biodiversity that can be used to guide conservation planning (Fleishman et al. 2006) and ecosystem service assessment (Kremen 2005). In addition, climate change may alter the distribution of tree species in the future (Coops and Waring 2010; Hamann and Wang 2006; Pfeifer-Meister et al. 2013; Thuiller et al. 2005), and knowledge of the current distribution is the first step in attempting to understand, monitor, and as possible, manage those changes.

Despite the importance and various needs for tree species distribution data, the availability of these data is limited. Some projects may benefit from the tree observations and vegetation plot data compiled and shared via online databases such as the Global Biodiversity Information Facility (gbif.org) and the Global Index of Vegetation-Plot Databases (givd.info). However, these data remain spatially incomplete and biased towards easily accessible or protected areas (García Márquez et al. 2012; Hortal et al. 2007). Expert range maps (e.g., Little 1971) provide a general indication of where species occur, but overestimate the true distribution of a species (Jetz et al. 2007; McPherson and Jetz 2007). Strategic-level forest inventories are typically undertaken in areas that have the capacity to support commercial timber production. Outside these areas, forest inventory data may be available, generally with less spatial and attributional detail, according to the forest monitoring needs in the area. In Canada, for example, forest inventories are common in the more intensively managed southern forests (Falkowski et al. 2009), while only the sample-based National Forest Inventory aims to systematically characterize forest resources outside of managed forest areas (Gillis et al. 2005; Wulder et al. 2004a).

The paucity of detailed inventory data makes satellite remote sensing a necessary source of information from which species distributions can be mapped or modelled over large areas. High spatial resolution imagery can offer opportunities for mapping individual tree structure and composition (Wulderet al. 2004b), but the spatial image extents are limited (e.g., 10 by 10 km) requiring many images (with variable view angles and illumination conditions) to map a given area, leading to high cost. Imagery with larger extents are often of interest for mapping larger areas; however, these data come with pixel sizes that subsume many individual objects and landscape features, diminishing the variance between pixels, and limiting the capacity to map high levels of categorical detail, such as tree species composition. Issues such as atmospheric contamination (i.e., clouds, haze) further limit scene availability and confound mapping efforts. Several recent advances address some of these shortcomings and offer increased capacity to use remotely sensed data for predictive species mapping over large areas at higher spatial resolutions. Specifically, as of 2008, satellite imagery from the Landsat series of sensors, extending from 1972 to present, are freely available to the public (Woodcock et al. 2008; Wulderet al. 2012). Free and open access to analysis-ready Landsat data has enabled considerable innovative capacity (Wulder and Coops 2014). Combined with improved computing power that facilitates large-area image compositing approaches (Roy et al. 2010) as well as the applications-focused best-available-pixel (BAP) approaches of Griffiths et al. (2013) and White et al. (2014), spatially exhaustive coverage of large areas at a spatial resolution of 30 m in a systematic and transparent fashion is now possible. For instance, compositing approaches can be based upon use of the best available observation for each pixel, with "best" being defined according to a set of scores for characteristics such as year, day of year (DOY), distance to cloud and cloud shadow, and sensor (Griffiths et al. 2013). Regional composites of medium spatial resolution imagery can be expected to become increasingly common (Griffiths et al. 2014). Detailed, efficient (large-area) maps of species distributions are a likely product of these compositing approaches, so long as the models are robust to some composite-imposed spectral variability. The goal of this research is to generate spatially detailed (30 m) distribution maps for six tree species over a large forested area of Canada using a multi-year Best-Available-Pixel

(BAP) Landsat composite. A specific objective was to evaluated the impact of composite characteristics (i.e., DOY, year, and sensor) on model outcomes, which was assessed by the spatial pattern (composition and configuration) of the predictions.

3.2 Methods

3.2.1 Study area

The study area is approximately 39 million ha, and comprises the three forested ecozones of Saskatchewan, Canada. From North to South, these are the Taiga Shield, the Boreal Shield, and the Boreal Plains (Ecological Stratification Working Group 1996) (Figure 3.1). The Boreal Plains is adjacent to the Prairie ecozone further south, and consists of rolling uplands and plains with a mixture of deciduous and coniferous vegetation species (McLaughlan et al. 2010). The Boreal Shield and Taiga Shield are characterized by a harsh climate and poorer soils, a greater proportion of coniferous tree species, and a lower diversity of plant species (McLaughlan et al. 2010; Pastor et al. 1996). A variety of provincial forest inventory data exist in the managed forest area of the province (Figure 3.1), with varying spatial scales, levels of attribution, and temporal frequencies (Gillis et al. 2005; Saskatchewan Environment - Forest Service 2004; Saskatchewan Ministry of Environment 2009).

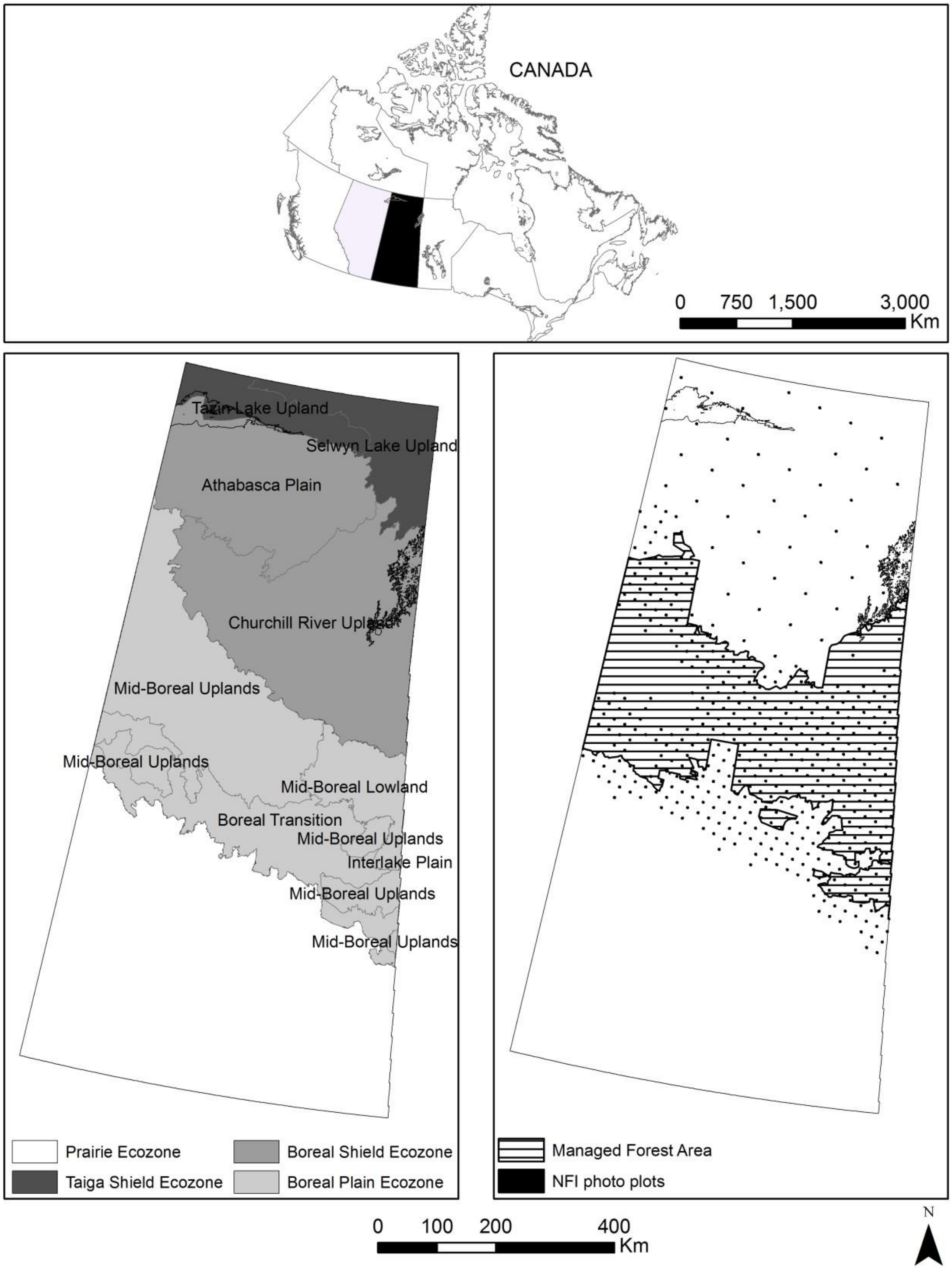


Figure 3.1. Study area. (a) The province of Saskatchewan (highlighted in black), in central Canada. (b) The forested ecozones and ecoregions of Saskatchewan. (c) The distribution of Canada's National Forest Inventory 2 km × 2 km photo plots across Saskatchewan, including inside and outside of the Managed Forest Area. These inventory data were the source of the training data used to model tree species distributions.

3.2.2 Tree species distribution data

Tree species distribution data were acquired from Canada's National Forest Inventory (NFI). The NFI consists of a grid of permanent sample plots distributed across the country, the majority of which are 2 km x 2 km "photo plots" (i.e., derived from air photo interpretations), within which multiple polygons are delineated indicating species composition and relative abundance. To reduce uncertainty at polygon edges, we removed 30 m (one pixel) from the inner edge of each polygon to ensure agreement between the inventory data and our predictor data (Verbyla and Hammond 1995). To ensure adequate sample size and quality, we did not model species that occurred relatively infrequently or that occurred exclusively in polygons of heterogeneous composition (≤90% of one species). In addition, polygons that were observed to have burned or been harvested since they were inventoried were removed, as were polygons that were very small in size (less than one pixel). Thus of the eleven tree species identified in the NFI photo plot data, we modelled the six most common (Table 3.1): black spruce (*Picea mariana*), trembling aspen (*Populus tremuloides*), jack pine (*Pinus banksiana*), white birch (*Betula papyrifera*), tamarack (*Larix laricina*), and white spruce (*Picea glauca*). Species that were present in insufficient numbers for modeling were *Abies balsamea, Acer negundo, Fraxinus pennsylvanica, Pinus contorta,* and *Populus balsamifera.* Our six target species may be found across all three forested ecozones in the province (McLaughlan et al. 2010).

Table 3.1. Species modelled, area and number of National Forest Inventory Photo Plot (NFI PP) polygons dominated by each

English Name	**Latin Name**	**Dominant (present at ≥90% abundance)[a]**		**Non-Dominant (present at ≤10% abundance)[a]**
		Number of polygons	Area (km^2)	Number of polygons
White birch	*Betula papyrifera*	29	0.71	5404
Tamarack	*Larix laricina*	459	27.11	4974
White spruce	*Picea glauca*	49	0.73	5348
Black spruce	*Picea mariana*	2057	55.81	3376
Jack pine	*Pinus banksiana*	856	38.63	4577
Trembling aspen	*Populus tremuloides*	1983	88.48	3450

[a] Count follows exclusion of very small polygons (less than 900 m^2) and those that were burned or harvested in the time since the polygon was delineated and attributed.

3.2.3 Image composite data

We used a *multi-year* BAP surface reflectance composite as the source of spectral information in our distribution models. A detailed description of this and other compositing methods is provided in White et al. (2014). Briefly, candidate pixel observations were scored according to sensor (Landsat TM or ETM+), year, DOY, distance to clouds or cloud shadows, and haze, and the pixels with the highest score used to populate the final image composite. Our target was Landsat 5 TM imagery from August 1, 2010; however, candidate pixels included all observations acquired ± 30 days of August 1, 2009, 2010, and 2011, from Landsat 5 TM or 7 ETM+, as required to provide complete, cloud-free coverage of the study area. In Table 3.2 we show the number of unique images considered and selected for the final BAP composite for Saskatchewan as a whole (encompassing our study area as well as the Prairie Ecoregion). After scoring, 5% of pixel observations in the final composite of our study area were acquired from 2009 imagery, 69% from 2010 imagery, and 2% from 2011 imagery. The remaining 24% of pixels had BAP observations for both 2009 and 2011, and a proxy value was generated by taking the average of the 2009 and 2011 observations (see White et al., 2014). Pixels with proxy values were excluded from the analyses because there were no logical corresponding DOY or sensor values with which to assess relationships. In total, 95% of pixel observations were from Landsat 5 TM, and 5% from Landsat 7 ETM+. Almost 30% of pixel observations came within 7 days of the target DOY (August 1), with the remaining pixels acquired within 30 days of August 1. In the context of this study, which was designed to model tree species distributions, it is important to note that the majority of pixels (by area) in the composite came from imagery acquired in the 2010 target year and, furthermore that tree species distributions tend to change slowly over longer time horizons. As such, the multi-year image composite used in this study is appropriate for modelling species distributions.

Table 3.2. Number of Landsat images used in Best-Available-Pixel composite for Saskatchewan, Canada.

	Candidate*			**Composite**		
Year	**TM**	**ETM+**	**Total**	**TM**	**ETM+**	**Total**
2009	379	328	707	345	188	533
2010	362	367	729	342	284	626
2011	406	398	804	333	218	551

*Candidate images are ± 30 days of August 1 with less than 70% cloud cover.

In order to reduce undesirable variability in the spectral reflectance values of the predictor variables used in our models (see Loveland and Merchant 1991), we calculated several spectral indices from the image composite data, and used these to exclude non-vegetated areas, and to the extent possible, non-forested areas from our analyses (Figure 3.2). Specifically, the Normalized Difference Vegetation Index (NDVI), and the Tasseled Cap (Crist and Cicone 1984; Kauth and Thomas 1976) Greenness (TCG), Brightness (TCB), and Wetness (TCW) indices were used to remove water, bare ground, urban areas, and sparsely vegetated areas. As well, the Normalized Burn Ratio (NBR) (Key and Benson 2006), was used to identify and remove areas that had experienced fire in recent years, with the threshold NBR value chosen for this analysis (0.15) validated by data from the Canadian National Fire Database (Natural Resources Canada 2010). An unsupervised classification of the Landsat spectral bands for the remaining pixels served to remove additional areas subsequently identified as cultivated land.

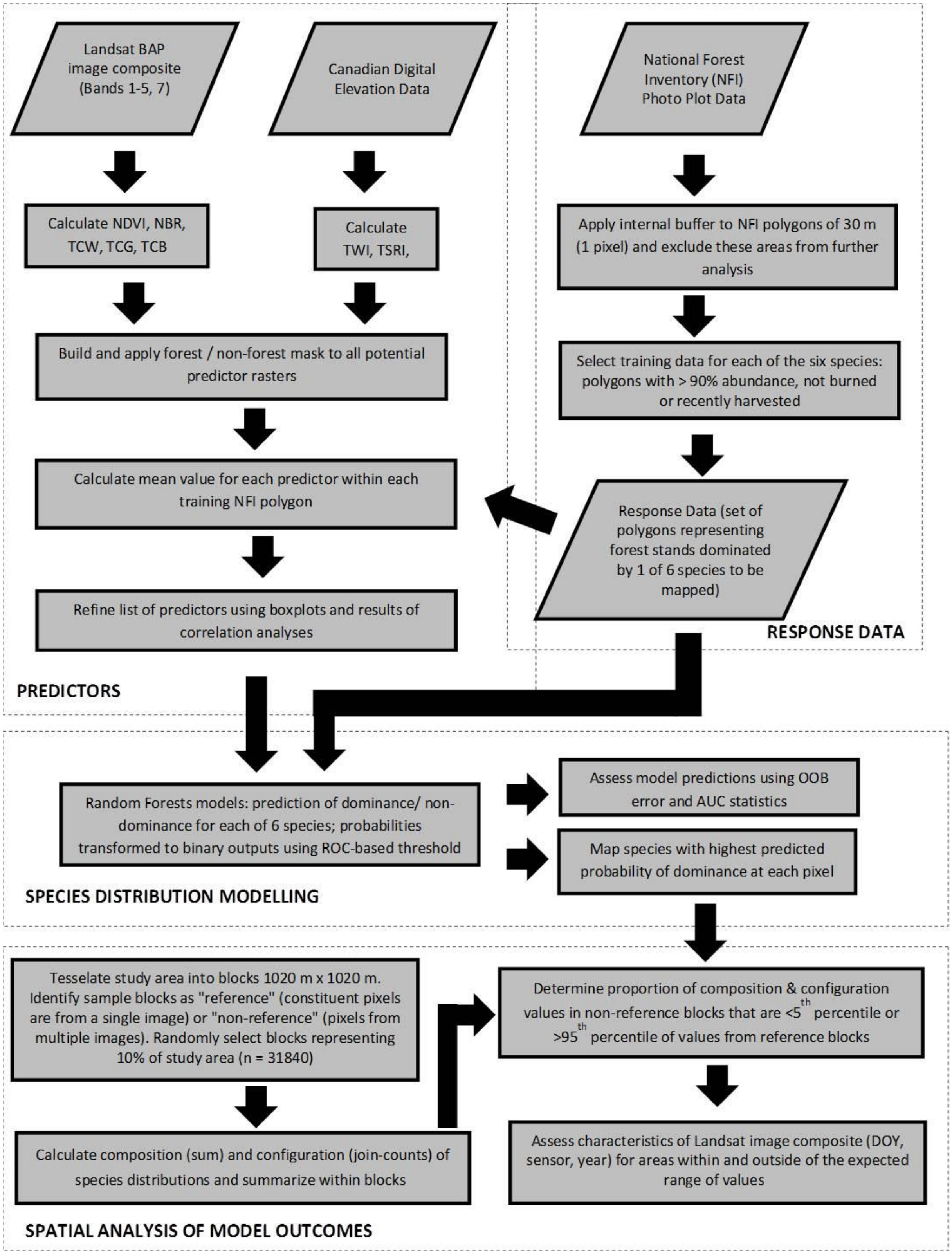


Figure 3.2. Flowchart of the data and methods followed to model species distributions and assess the impact of BAP composites of the resultant maps.

3.2.4 Topographic data

Topographic data were acquired from the freely available Canadian Digital Elevation Data (http://www.geobase.ca/geobase/en/data/cded/). These elevation data are derived from provincial and national topographic data sources, and are provided as a 1:50,000 digital elevation model (DEM). The DEM, which has a native spatial resolution of approximately 23 m (0.75 arc seconds), was resampled to match the 30 m spatial resolution of our image composite using bilinear resampling. From the 30 m DEM we calculated slope (in degrees), the Topographic Solar Radiation Index (TRASP; a transformed measure of aspect), and the Topographic Wetness Index (TWI) (Table 3.3).

Table 3.3. Predictor variables used in the Random Forests models of individual tree species distributions.

Name	**Description**	**Rationale**
Brightness	Tasseled Cap components (Crist and Cicone 1984; Crist 1985)	TC brightness can differentiate between soil and vegetated surfaces (Crist et al. 1986) and can help differentiate among successional stages / stand age (Song et al. 2007).
Greenness		Greenness in relates to biomass and vegetation vigor and is highly correlated to the Normalized Difference Vegetation Index (NDVI), which is useful for general land cover classification (Defries et al. 1995; Running et al. 1994)
Wetness		TC Wetness correlated to structural complexity (M. J. Hansen et al. 2001), perhaps particularly for successional stages / stand age (Wulder, Skakun, et al. 2004c).
Topographic Wetness Index (TWI)	Model of potential surface moisture, based on topographic position (Beven and Kirkby 1979): *Ln(specific catchment area / tan(slope in radians))*	Soil moisture directly affects plant growth.
Topographic Solar Radiation Aspect Index (TRASP)	Values range from 0 to 1, with 0 indicating cool, NE slopes, and 1 indicating warm SW slopes (Roberts and Cooper 1989).	Solar radiation affects soil moisture and heat load, and directly affects plant growth.

3.2.5 Species distribution modeling

Species distribution modelling was conducted using Random Forests™ (RF) in R 3.1 (Breiman 2001). We chose RF, a type of decision tree, because it can accommodate non-normal responses and non-linear relationships, and automatically account for interactions among predictors (De'ath and Fabricus 2000; Elith et al. 2008; Hawkins 2012). Decision trees involve a sequence of binary splits at values of the predictor variables that result in the maximum differentiation of values of the response variable (in this case, species dominance or non-dominance at a given location). In RF, many (500 to 2000) single trees are developed, each constructed from a different bootstrapped sample of the training data and a randomly selected subset of the predictor variables (Prasad et al. 2006). The predictions are averaged over all trees to generate an overall probability while minimizing the chances of over-fitting to the training data (Franklin 2009; Prasad et al. 2006). Ensemble tree methods have been found to perform well relative to most other predictive methods across many regions, and species, including plants (Elith et al. 2006; Guisanet al. 2007a; Prasad et al. 2006).

The use of RF for species distribution modeling involved both fitting and prediction stages. First, each model was fit using the species observation data from Table 3.1 and the mean of each predictor variable. The predictor variables used in the modelling were selected from among the multiple, aforementioned topographic and spectral variables after conducting a Spearman's rank correlation analysis and assessing variable utility through boxplots for each predictor variable across all species. Specifically, we selected the following three spectral and two topographic indices for use as inputs, all with correlations less than ±0.15: the Tasseled Cap TCG, TCB, and TCW, the TRASP and the TWI (Table 3.3). We used 500 decision trees, with a random subset of two of the explanatory variables chosen for input for each of these individual trees. Because of the unbalanced number of observations of dominance and non-dominance) in the training data for a given species (Table 3.1), we used a “down-sampling” approach, specifying that each model should use all samples from the least common class (dominance), and an equal number of samples from the more common class (non-dominance) (Chen et al. 2004). A separate RF model was generated for each species, and thus the number of samples varied across models.

Each output RF model contained probabilities of dominance ranging from 0 to 1, which are classified by default into the binary classes dominance or non-dominance, if probabilities are ≥0.5 or <0.5, respectively. However, often a threshold probability other than 0.5 is preferred (Nenzén and Araújo 2011), particularly for rare species (Freeman and Moisen 2008). To choose appropriate threshold probabilities for our models, we generated a Receiver Operating Characteristic (ROC) plot for each species using the auc.roc.plot() function in the PresenceAbsence library in R (v3.1.2). An ROC plot shows how the rate of true positives (y-axis) versus false positives (x-axis) of a model vary for all threshold probabilities between 0 and 1. A perfect classification would pass through the upper left corner of the plot (100% true positives and 0% false positives). A threshold value that achieves the minimal distance between this place of perfect classification and the curve, is an appropriate value to transform continuous outputs to binary classifications (Liu et al. 2005).

Finally, each of the six models were re-run using values of the predictor variables for locations (pixels) where species observations were absent, and the predictions of dominance and non-dominance output as continuous raster surfaces using the determined thresholds. A secondary goal was to create a forest composition map with all species combined. A composite species map was generated by evaluating the individual probabilities of dominance resulting from each of the six RF models were compared at each location (pixel) and that species with the highest overall probability was selected as the appropriate classification value for that location.

3.2.6 Model evaluation

Model performance was assessed using the *out-of-bag* (OOB) error generated internally by the RF method eliminating the need for a separate cross-validation (Breiman 2001). Specifically, the ability of the classifier to correctly predict observed values was assessed, where discrete class predictions were based on species-specific probability thresholds determined through the ROC analysis described above. We also examined the Area Under the Curve (AUC) associated with an ROC plot, as calculated by back-predicting on our observed data (essentially but not exactly the same data used for training because of the subsampling and consensus approach used in RandomForests). The AUC ranges from 0.5 to 1, and indicates the proportion of times that the model discriminates between our two outcomes better than random (Jiménez-Valverde 2012), or more specifically, the proportion of times a randomly chosen instance of dominance has a value larger than that for a randomly chosen instance of non-dominance (Fielding and Bell 1997). Thus higher AUC values indicate better models.

Finally, we compared our resultant tree distribution models with previous studies, general knowledge of the species' ranges, and trends of dominance per ecozone and ecoregion. In particular, we calculated the areal extent of each dominant species as predicted in our overall forest composition map per ecozone and ecoregion and ranked these in descending order. We repeated this calculation using homogenous polygons from NFI photo plot data to assess agreement between model outputs and the training data over a broader scale. Similarly, we generated regional summaries of species dominance by combining predictions of relative basal area for the same species from a recent study by Beaudoin et al. (2014).

3.2.7 Assessing the effects of the image composite on spatial patterns of the models

To assess the impact that compositing had on predicting species distributions, we tested the hypothesis that the spatial pattern of predictions was similar for reference sample blocks and non-reference sample blocks. Reference blocks had pixel observations derived from a single year, a single DOY, and a single sensor, while non-reference blocks had pixel observations from multiple years, DOYs, and sensors. Spatial patterns can be quantified by a combination of composition and configuration. Whereas composition is *aspatial* and refers to the variety and (relative) abundance of different features (e.g., tree species), configuration refers both to the spatial characteristics of individual patches such as size and shape, as well as spatial relationships among neighbouring patches or neighbouring cells (Gustafson 1998). While composition indicates what is present at any given location, configuration metrics provide a context to local conditions, and permits study of how spatial patterns are an expression of process (Fahrig 2005; Turner 1989).

The analysis was undertaken within sample blocks measuring 1020 m x 1020 m, distributed over a random 10% of the study area (for a total of 31,840 samples). An extent of 1020 m x 1020 m was chosen to ensure coverage of the data gaps (of 1 to 14 pixels in size, see Goward et al., 2010; Storey et al., 2005) resulting from Landsat 7 ETM+ Scan-Line Correction failure, while also being a number within which 30 m Landsat pixels could be equally divided. Three image composite characteristics were evaluated: acquisition year (2009, 2010, or 2011), sensor (TM or ETM+), and the number of days from the target DOY of August 1 (ranging from 0 to 30).

For each species, composition was quantified within each 1020 m x 1020 m block as the sum of pixels with predicted dominance for that species. All blocks had dominance of at least one species. Configuration was measured for each species using join counts within each 1020 m x 1020 m area. A join count test can be used to assess spatial autocorrelation in categorical, especially binary, variables, such as dominance/non-dominance (Boots 2006). Using a join count, the spatial configuration of a species can be quantified as clustered or dispersed, relative to complete spatial randomness (O'Sullivan and Unwin 2010). For binary data, the two categories are normally referred to as either "Black" (B) or "White" (W) (here, dominance or non-dominance, respectively). For this analysis, we were interested only in the J_{BB} (dominance-dominance) join-count statistic:

$$J_{BB} = \frac{1}{2}\left(\sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \delta_{ij} x_i \, x_j\right)$$

where *i* and *j* are the two sampling units being compared, x_i is the value of the sampling unit (1 or 0), and δij is the adjacency of *i* and *j* (1 when they are adjacent, 0 when they are not). The expected values of joins are then calculated based on the proportion of each category and number of total joins in the study (which depends on how connectivity is defined), and the observed and expected values are then compared to assess the null hypothesis of complete spatial randomness (Fortin and Dale 2005). We computed a join count for each species, using the Rook's case definition of contiguity (four neighbours). Each species had to be predicted within at least two 30 m cells within each 1020 m x 1020 m block to be included in the analysis. Missing data within each 1020 m x 1020 m block were reclassified as zeroes to allow calculation of the join-counts, while remaining statistically conservative.

To assess whether the use of an image composite affected our modelled species distributions, we compared values of composition and configuration between reference (n = 23,581) and non-reference (n=8259) sample blocks. We calculated the frequency distribution of the composition and configuration values for each of the six species from the reference sample blocks, and extracted the 5th and 95th percentiles for each. The number of composition and configuration values in the non-reference sample blocks that fell below the 5th or above the 95th percentiles was then calculated. Blocks with these unexpected values were then mapped and a summary of their sensor type, year and DOY characteristics were extracted.

3.3 Results

3.3.1 Species distribution modelling

The ROC threshold optimization method resulted in a threshold of 0.7 for all species except for *Picea mariana* and *Populus tremuloides,* for which the optimum threshold was 0.6 (Table 3.4). These thresholds were used to map the distribution of each individual tree species (Figure 3.3). Combining the individual species maps creates an overall map of forest composition (Figure 3.4). *Pinus banksiana* was predicted to dominate over the largest spatial extent (8.9 million hectares in total), particularly at mid-to-high latitudes (Figure 3.4). *Populus tremuloides* was predicted to dominate in the southern extreme of the Boreal Plain ecozone, in the Aspen Parkland ecoregion, but is also found across the province even in the far north, with predicted dominance covering 7.3 million ha in total. *Picea mariana* was predicted to be the next most widespread species, dominating over 6.6 million ha, particularly at mid-latitudes, but being also widespread in the north. *Larix laricina* was predicted to be dominant at low to mid latitudes in the Boreal transition ecoregion and in lowland areas and known wet areas such as the Saskatchewan River delta, straddling the eastern border of Saskatchewan, covering 3.4 million ha in total. *Picea glauca* and *Betula papyrifera* were predicted to dominate over a much smaller area (approximately 1.3 million ha and 824,000 ha respectively).

Table 3.4. Threshold used to translate probabilities of species dominance to a binary dominant or non-dominant variable. The threshold selected minimized the distance on a plot of the ROC (Receiver Operating Characteristic) curve between the upper left corner of the plot and the curve.

Species	**Threshold used to create binary map**[a]	**OOB error rate**[b] **(for class "dominant")**	**Area Under the Curve (AUC)**[c]
Betula papyrifera	0.7	0.24	0.99
Larix laricina	0.7	0.12	0.99
Picea glauca	0.7	0.24	0.99
Picea mariana	0.6	0.11	1
Pinus banksiana	0.7	0.14	1
Populus tremuloides	0.6	0.09	1

[a] According to the criteria of minimizing the distance on a plot of the ROC (Receiver Operating Characteristic) curve between the upper left corner of the plot and the curve.

[b] The Out of Bag (OOB) error indicates the total number of misclassified data points from within the out-of-bag sample (Breiman 2001).

[c] Values close to 1 indicate good model fit

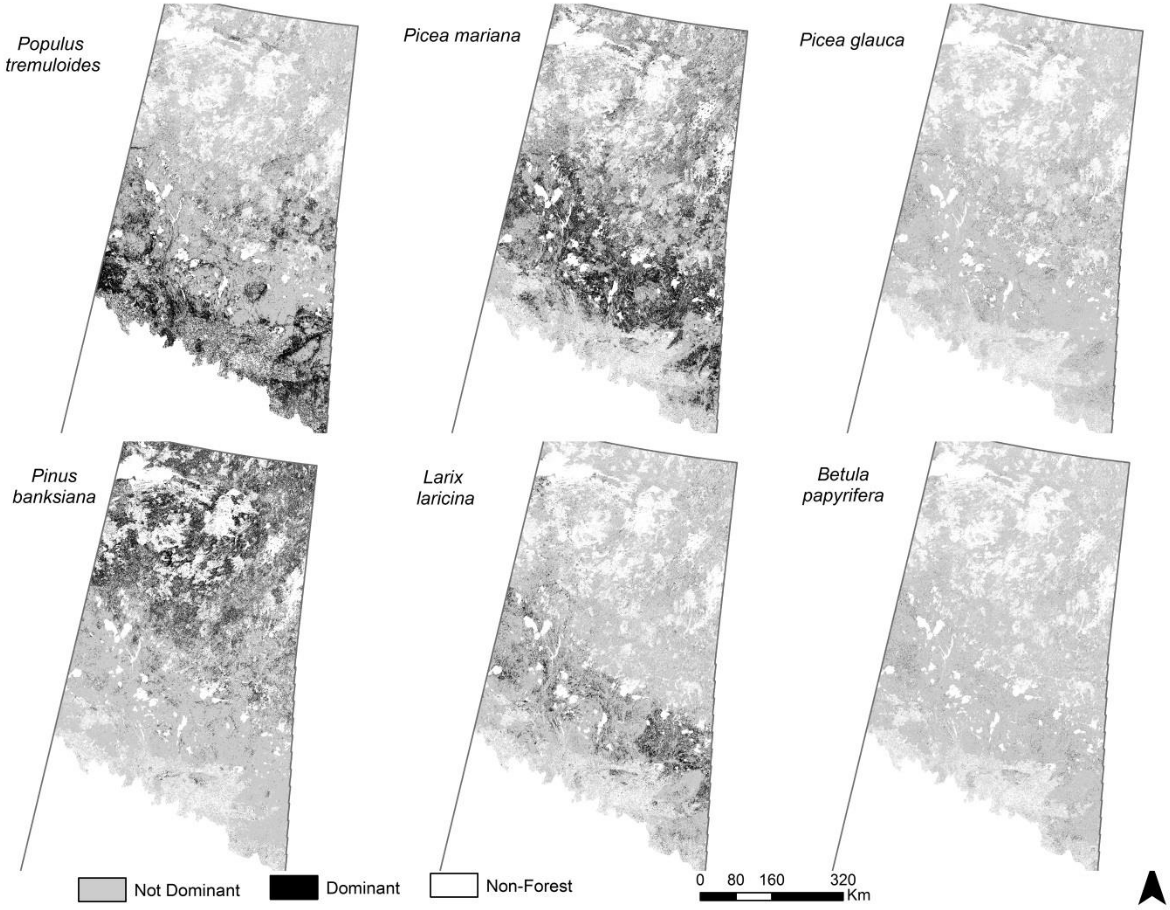


Figure 3.3. Individual models of tree species dominance in Saskatchewan.

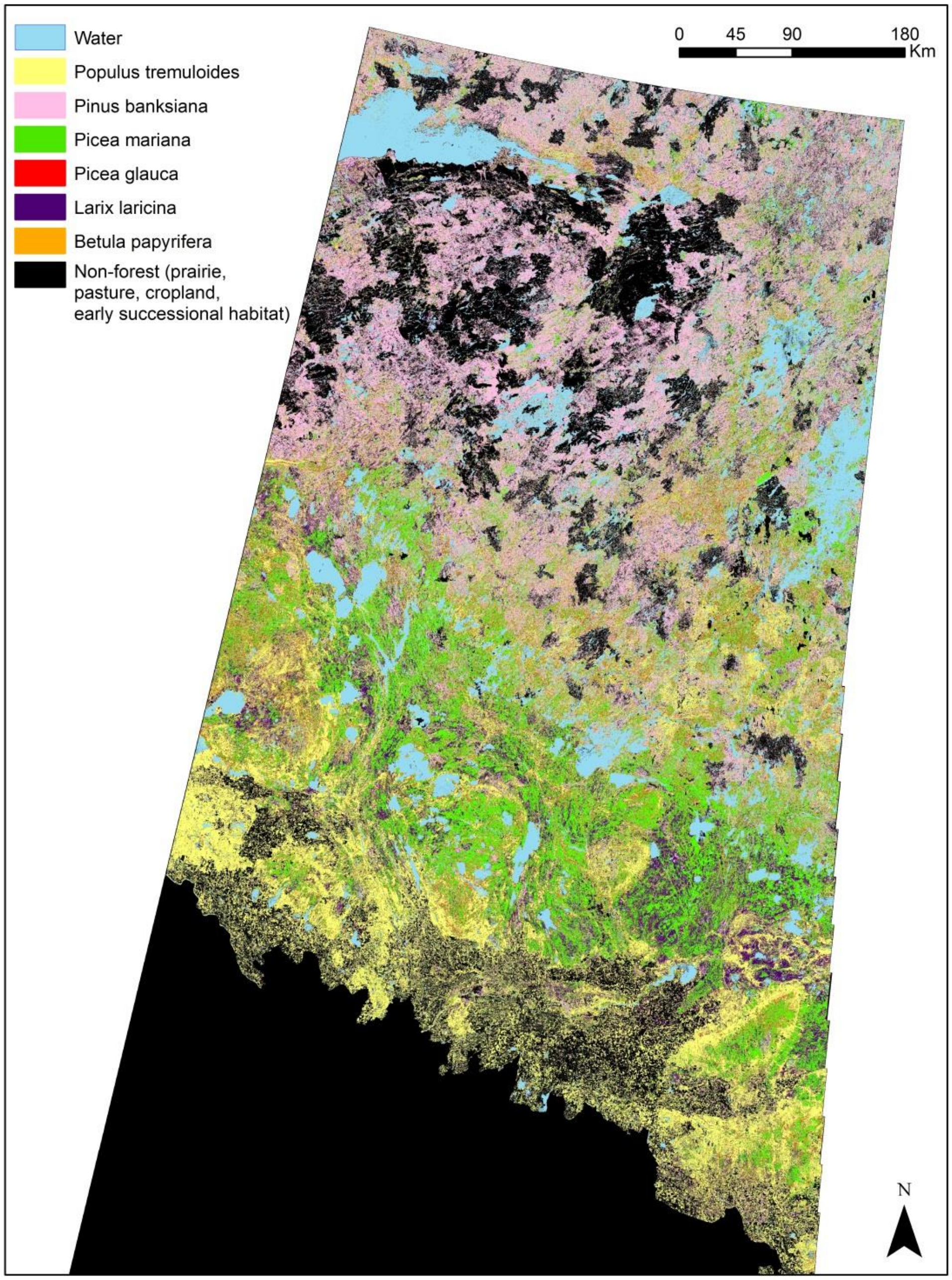


Figure 3.4. Forest composition map showing tree species with highest predicted probability of dominance at each location.

3.3.2 Model evaluation

The ability of our models to correctly classify the dominance or non-dominance of individual species varied from species to species (Table 3.4). Overall, OOB error rates were less than 25%, indicating reasonable model fit to the training data. Error rates were much lower for species with sufficient sample sizes. Specifically, at the selected thresholds, *Populus tremuloides* (with one of the highest sample sizes) had the lowest OOB error at 9%. *Betula papyrifera* and *Picea glauca* (with the two lowest sample sizes) had the highest OOB error rates at 24% each. AUC values were very high for all models (0.99 to 1), indicating good model performance.

Trends in the relative areal extent of species dominance for the province's forested ecozones and ecoregions predicted in this study are generally comparable to those in the NFI photo plot database, as well as to those of species occurrence from Beaudoin et al. (2014) (Table 3.5). For example, all three studies/datasets indicated that both the Boreal Plain Ecozone, and Boreal Transition Ecoregion, are dominated by *Populus tremuloides*. All three datasets also suggest *Pinus banksiana* is dominant in the Athabasca Plain ecoregion. Some differences are also apparent. For instance, the current study predicts *Pinus banksiana* to be dominant across the largest proportion of the Taiga Shield, whereas the other datasets indicate *Picea mariana* is most dominant.

Table 3.5. Spatial Extent of Dominant Tree Species in Saskatchewan's Forested Ecozones and Ecoregions.

		Most Spatially Expansive Species		
Ecozone	**Ecoregion**	**NFI photo plot data[1]**	**Beaudoin et al. 2014[2]**	**This study**
Taiga Shield		*Picea mariana* (70% of ecoregion)	*Picea mariana* (80% of ecoregion)	*Pinus banksiana* (61% of ecoregion)
	Tazin Lake Upland	*Picea mariana*	*Picea mariana*	*Pinus banksiana*
	Selwyn Lake Upland	*Picea mariana*	*Picea mariana*	*Pinus banksiana*
Boreal Shield		*Pinus banksiana* (48% of ecoregion)	*Picea mariana* (49% of ecoregion)	*Pinus banksiana* (46% of ecoregion)
	Athabasca Plain	*Pinus banksiana*	*Pinus banksiana*	*Pinus banksiana*
	Churchill River Upland	*Picea mariana*	*Picea mariana*	*Pinus banksiana*
Boreal Plain		*Populus tremuloides* (48% of ecoregion)	*Populus tremuloides* (48% of ecoregion)	*Populus tremuloides* (40% of ecoregion)
	Mid-Boreal Uplands	*Populus tremuloides*	*Picea mariana*	*Picea mariana*
	Boreal Transition	*Populus tremuloides*	*Populus tremuloides*	*Populus tremuloides*
	Mid-Boreal Lowlands	*Larix laricina*	*Picea mariana*	*Larix laricina*

[1] Calculations included only homogenous polygons (those where ≥ 90% of the polygon is comprised of a single species).

[2] As in the generation of our forest composition map, each individual species distribution map from Beaudoin et al. 2014 was combined and assigned the value of the species with the highest predicted relative basal area. Calculations in this table are based on the combined map.

3.3.3 Assessing the effects of the image composite on spatial patterns of the models

The proportion of composition and configuration values within the non-reference sample blocks falling outside the expected distribution (5th to 95th percentiles of values of the reference blocks) was fairly low. Depending on the species, 3.7% to 9.6% of the blocks had unexpected composition values, while 7.1% to 16.5% had unexpected configuration values (Table 3.6). Given that we set the critical value of the statistical comparison to 0.10, we would expect around 10% of blocks to have unexpected composition and configuration. Only the configuration of *Populus tremuloides* had a higher than statistically expected number of unexpected blocks (16.5%). The spatial distribution of the sample blocks with unexpected composition and configuration values appears random (Figure 3.5). Further, the DOY, year and sensor characteristics in these expected and unexpected regions were found to be very similar. For instance, all of these non-reference sample blocks, regardless of whether they had expected or unexpected values of species composition and configuration, were comprised primarily of imagery from 2010, with the difference in proportions of 2010 imagery between expected and unexpected blocks ranging from ~2% to 9%, depending on the species. Likewise, all blocks contained imagery primarily from Landsat 5; blocks with unexpected values of composition and configuration differed in terms of sensor composition by no more than 9% from blocks with expected values. Mean DOY differed by no more than two days for blocks with expected and unexpected samples of species composition and configuration.

Table 3.6. Proportion of composition and configuration values within non-reference sample blocks (n=8259)* that are <5th or >95th percentile of values within reference sample blocks (n=23,581).

	Betula papyrifera	*Larix laricina*	*Picea glauca*	*Picea mariana*	*Pinus banksiana*	*Populus tremuloides*
True sample size*****	6177	7361	5588	7966	7446	7675
% unexpected Composition	3.71	9.96	5.71	9.63	7.13	8.33
True sample size*****	4469	6517	4274	7682	6677	7481
% unexpected Configuration	7.14	9.56	7.25	8.69	8.62	16.53

*note that the non-reference sample size was less than 8259 for each species because for each species in turn, blocks had to contain at least one pixel of presence (dominance) for analysis of composition, and at least two pixels of presence (dominance) for analysis of configuration.

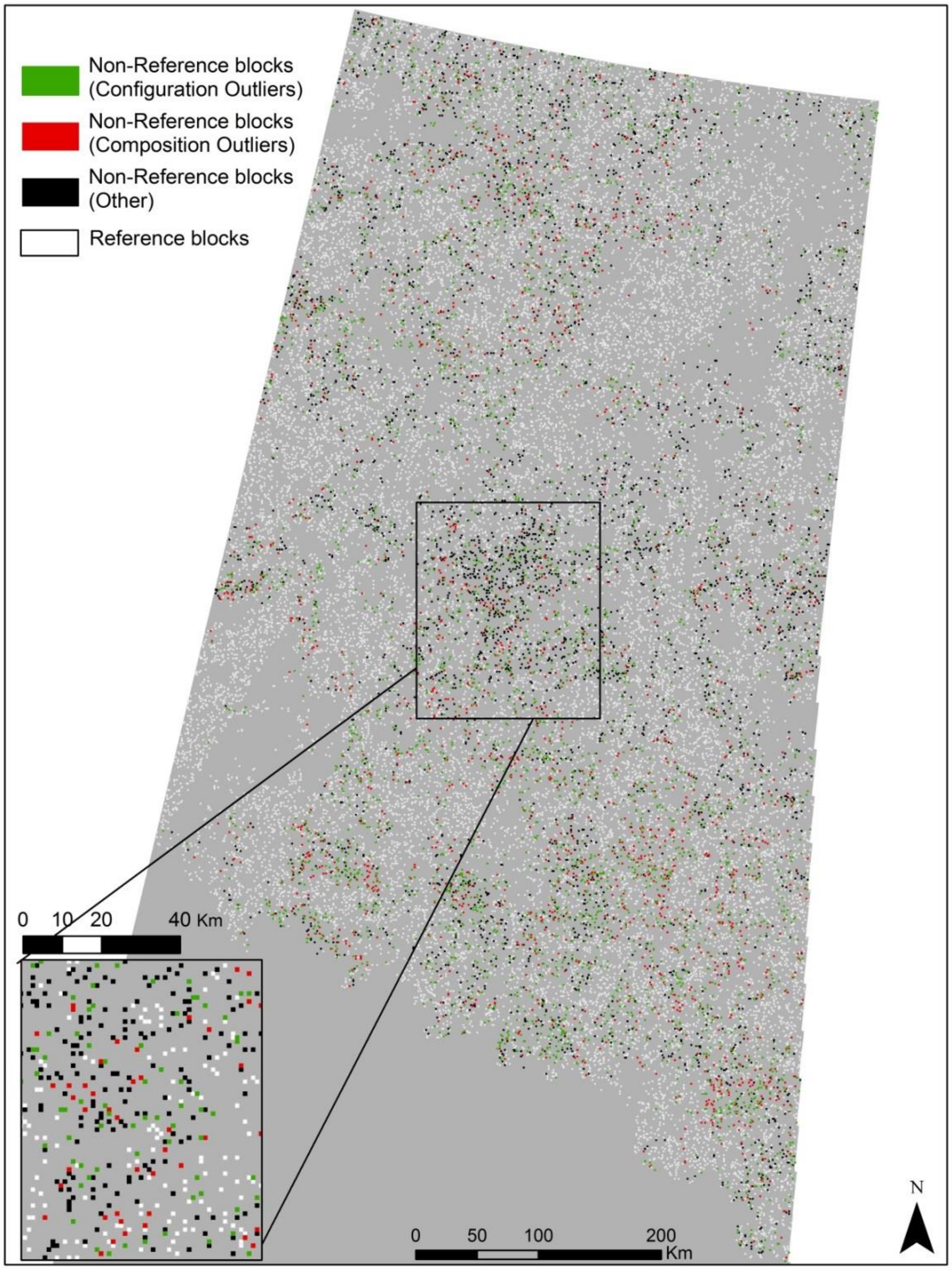


Figure 3.5. Nonreference sample blocks (1020 m × 1020 m) are those comprising a mixture of sensor types, image years, or image DOY. Blocks are highlighted where species composition and configuration values were outside the 5th–95th percentile of values.

3.4 Discussion

The use of satellite imagery for vegetation mapping is a desirable supplement to ground or photo-based inventories because of the large spatial extents that can be covered by satellite imagery, as well as the associated automated, repeated acquisition. Tree species distribution mapping based on satellite imagery involves the detection or classification of separate spectral reflectance signatures for each species (Bradter et al. 2011); however, tree species classification is difficult, as many vegetation species have overlapping spectral reflectance characteristics in the wavebands collected by typical multispectral sensors (Immitzer et al. 2012; van Aardt and Wynne 2001). Hyperspectral imagery, which collects reflectance in many, narrow wavebands, is often needed to map species composition to a high, or even modest, degree of accuracy (e.g., Buddenbaum et al., 2005; Ustin and Xiao, 2001). However, this type of imagery is not currently cost effective for inventorying large regions because, like high spatial resolution data, hyperspectral imagery are associated with small spatial extents, requiring multiple scenes or, more likely, airborne collections to represent a given area, thereby increasing data costs and processing overhead. In this study we mapped the probable distribution of dominance of six tree species across a large region using freely available moderate spatial resolution multispectral imagery. Our model error rates (~10-25%) were typical for forest species distribution modeling using this type of imagery (e.g., Evans and Cushman, 2009).

Some error and uncertainty in our distribution models can be attributed to limited sample size. Sample size of the species data has been shown to affect the accuracy of predictive models in previous research (Stockwell and Peterson 2002; Wisz et al. 2008). Although machine learning and ensemble methods like Random Forests can perform relatively well with small to moderate sample sizes, especially when absence information is available in addition to presence information (Elith et al. 2006; Guisan2007b), estimates of error in this study were nonetheless highest for the species with the lowest sample sizes (Table 3.1). Model accuracies may have also been affected by characteristics of the individual species modelled. Specifically, wide-ranging species are typically more challenging to model than species with more particular niches (Guisanet al. 2007a; McPherson et al. 2004). The six tree species modelled in this study are all wide-ranging species, and are generally tolerant of a range of soils and parent materials (Farrar 1995). That our training data indicated where a species was and was not dominant was therefore likely particularly important. Indeed, distribution models are particularly robust when reliable absence data are available in addition to presence data (Brotons et al. 2004).

Another potential source of uncertainty in large area mapping and modelling relates to the remotely sensed data itself. In multi-temporal image analysis, differences in atmospheric conditions, and variability in phenology, sun angle and view angle of imagery (Song and Woodcock 2003) can lead to some uncertainty. Moreover, relationships between species and image spectral reflectances will vary seasonally according to species phenology (Maeda et al. 2014). In this study, we explored the use of a multi-year BAP composite to generate a series of distribution models for the six most common tree species in the forested area of Saskatchewan. We found that mean acquisition year, DOY, and sensor were similar regardless of the level of local complexity found in the composite. In other words, the variability of image characteristics across the BAP composite was actually very small, which was achievable due to the vast archive of open-source Landsat imagery (White and Wulder 2013), the rules used for compositing, and the pre-processing applied that converted the data to surface reflectance (White et al. 2014).

Composition and configuration of the predicted species are important characteristics to consider in species distribution modelling because spatial pattern is an expression of underlying spatial processes (Nelson and Boots 2008). Spatial pattern analysis is used to assess model error as it allows patterns in error or uncertainty to be detected and enables departures from random noise to be determined (Wulder et al. 2007). The approach used in this study allows mapping and detection of statistical departures in patterns of species distributions generated from non-reference imagery (Nelson and Boots 2005). We found that in all but one instance (*Populus tremuloides*) the composition and configuration of species distributions was not different among sample blocks with variable composite characteristics. This species in particular may stand out from the rest simply due to its overall prominence across the study area. Nonetheless, as the majority of our models were unaffected by the compositing, we are confident that the compositing rules used relating to sensor type, target DOY and cloud contamination minimized illumination and phenological differences sufficiently across space. Overall, results indicate that the relation between predicted species distributions and important environmental processes are represented, rather than species spatial pattern being the result of data artifacts in the composite. Overall, our individual maps of the probable distribution of tree species dominance across Saskatchewan meet expected trends as captured through two other independent data sources. For instance, broad-leafed species *Populus tremuloides* and *Betula papyrifera* were predicted to be dominant over a greater spatial extent in the south, relative to the north. The species predicted to be most widespread were *Picea mariana*, *Pinus banksiana,* and *Populus tremuloides*. These three species are typically dominant over a large number of ecosites in the province (McLaughlan et al. 2010). The primary difference among the three datasets is this study's predicted dominance of *Pinus banksiana* in the Taiga Shield, versus the dominance of *Picea mariana* in this region in the other datasets. We note that there are very few NFI plots in this ecozone, however, and that the differences between our results and those of Beaudoin et al.

(2014) likely stem from the use of different remotely sensed data (MODIS) with a different spatial resolution (250 m), and a different modelling approach (kNN).

3.5 Conclusion

Regional-scale, spatially comprehensive maps of forest composition have traditionally been limited by the mismatch between desired, versus available, spatial extent and spatial resolution of data. In this analysis, we have demonstrated the capacity to use 30 m Landsat data to map detailed tree species distributions over large areas, by capitalizing on archived, multi-temporal imagery composited using the BAP approach (White et al. 2014). The variability introduced by the BAP compositing was found to be minimal, and resulted in mostly insignificant differences in this large-area mapping application. Future applications will benefit from analyses of the effects of BAP compositing in other geographic regions and for features of interest other than tree species. The potential of the BAP approach to provide source data for developing species distribution maps over large areas will continue to increase with Landsat continuity and the launch of complementary satellites such as Sentinel-2 in 2015 (Drusch et al. 2012; Roy et al. 2014). The constellations of these new satellites have been designed such that, taken together, the majority of the Earth will be able to be imaged twice weekly (Wulder and Coops 2014) at a 30 m spatial resolution. Spatially continuous maps of tree species distributions over large areas will be useful for a variety of information needs, including forest management, carbon modeling, ecosystem service assessment, and conservation planning.

3.6 References

Beaudoin, A., Bernier, P. Y., Guindon, L., Villemaire, P., Guo, X. J., Stinson, G., et al. (2014). Mapping attributes of Canada's forests at moderate resolution through k NN and MODIS imagery. *Canadian Journal of Forest Research*, *44*(5), 521–532. doi:10.1139/cjfr-2013-0401

Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, *24*(1), 43–69. doi:10.1080/02626667909491834

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

Bradter, U., Thom, T. J., Altringham, J. D., Kunin, W. E., & Benton, T. G. (2011). Prediction of National Vegetation Classification communities in the British uplands using environmental data at multiple spatial scales, aerial images and the classifier random forest. *Journal of Applied Ecology*, *48*(4), 1057–1065. doi:10.1111/j.1365-2664.2011.02010.x

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5–32.

Brotons, L., Thuiller, W., Araújo, M., & Hirzel, A. (2004). Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, *27*, 437–448.

Buddenbaum, H., Schlerf, M., & Hill, J. (2005). Classification of coniferous tree species and age classes using hyperspectral data and geostatistical methods. *International Journal of Remote Sensing*, *26*(24), 5453–5465. doi:10.1080/01431160500285076

Chen, C., Liaw, A., & Breiman, L. (2004). *Using random forest to learn imbalanced data*. Technical Report 666.

Coops, N. C., & Waring, R. H. (2010). A process-based approach to estimate lodgepole pine (Pinus contorta Dougl.) distribution in the Pacific Northwest under climate change. *Climatic Change*, *105*(1-2), 313–328. doi:10.1007/s10584-010-9861-2

Crist, E. P. (1985). A TM tasseled cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, *17*, 301–306.

Crist, E. P., & Cicone, R. C. (1984). A Physically-based transformation of Thematic Mapper data - the TM Tasseled Cap. *IEEE Transactions on Geoscience and Remote Sensing*, *GE-22*(3), 256–263. doi:10.1109/TGRS.1984.350619

Crist, E. P., Laurin, R., & Cicone, R. C. (1986). Vegetation and soils information contained in transformed Thematic Mapper data. In *Proceedings of IGARSS' 86 Symposium* (pp. 1465–70). Paris, France: ESA SP-254. European Space Agency.

De’ath, G., & Fabricus, K. E. (2000). Classification and Regression Trees: a powerful yet simple technique for ecological data analysis. *Ecology, 81*(11), 3178–3192.

Defries, R., Hansen, M., & Townshend, J. (1995). Global discrimination of land cover types from metrics derived from AVHRR Pathfinder data. *Remote Sensing of Environment*, *54*(3), 209–222.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., et al. (2012). Sentinel-2: ESA’s optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, *120*, 25–36. doi:10.1016/j.rse.2011.11.026

Ecological Stratification Working Group. (1996). *A National Ecological Framework for Canada*. Ottawa/Hull. Agriculture and Agri-Food Canada, Research Branch, Centre for Land and Biological Resources Research, and Environment Canada, State of the Environment Directorate, Ecozone Analysis Branch.

Elith, J., Graham, C., Anderson, R., Dudik, M., Ferrier, S., Guisan, A., et al. (2006). Novel methods improve prediction of species’ distributions from occurrence data. *Ecography*, *29*(2), 129–151.

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *The Journal of Animal Ecology*, *77*(4), 802–13. doi:10.1111/j.1365-2656.2008.01390.x

Evans, J. S., & Cushman, S. A. (2009). Gradient modeling of conifer species using random forests. *Landscape Ecology*, *24*(5), 673–683. doi:10.1007/s10980-009-9341-0

Fahrig, L. (2005). When is a landscape perspective important? In J. Wiens & M. Moss (Eds.), *Issues and Perspectives in Landscape Ecology* (pp. 3–10). Cambridge University Press, Cambridge UK.

Falkowski, M. J., Wulder, M. A., White, J. C., & Gillis, M. D. (2009). Supporting large-area, sample-based forest inventories with very high spatial resolution satellite imagery. *Progress in Physical Geography*, *33*(3), 403–423. doi:10.1177/0309133309342643

Farrar, J. (1995). *Trees in Canada*. Fitzhenry and Whiteside Ltd.

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*(1), 38–49. doi:10.1017/S0376892997000088

Fleishman, E., Noss, R., & Noon, B. (2006). Utility and limitations of species richness metrics for conservation planning. *Ecological Indicators*, *6*(3), 543–553. doi:10.1016/j.ecolind.2005.07.005

Fortin, M.-J., & Dale, M. (2005). *Spatial analysis: A guide for ecologists*. Cambridge University Press.

Franklin, J. (2009). *Mapping species distributions: spatial inference and prediction*. Cambridge University Press, Cambridge UK.

Freeman, E. A., & Moisen, G. G. (2008). A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling*, *217*(1-2), 48–58. doi:10.1016/j.ecolmodel.2008.05.015

García Márquez, J. R., Dormann, C. F., Sommer, J. H., Schmidt, M., Thiombiano, A., Da, S. S., et al. (2012). A methodological framework to quantify the spatial quality of biological databases. *Biodiversity and Ecology*, *4*, 25–39. doi:10.7809/b-e.00057.25

Gillis, M. D., Omule, A. Y., & Brierley, T. (2005). Monitoring Canada's forests: The National Forest Inventory. *The Forestry Chronicle*, *81*(2), 214–221. doi:10.5558/tfc81214-2

Goward, S., Williams, D., Arvidson, T., & Irons, J. (2010). The future of Landsat-class remote sensing. In B. Ramachandran, C. O. Justice, & M. Abrams (Eds.), *Land Remote Sensing and Global Environmental Change: NASA's Earth Observing System and the Science of ASTER and MODIS*. Springer.

Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V. C., Abrudan, I. V., Lieskovsky, J., et al. (2014). Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites. *Remote Sensing of Environment*, *151*, 72–88. doi:10.1016/j.rse.2013.04.022

Griffiths, P., Linden, S. Van Der, Kuemmerle, T., & Hostert, P. (2013). A pixel-based Landsat compositing algorithm for large area land cover mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1–14. doi:10.1109/JSTARS.2012.2228167

Guisan, A., Zimmermann, N. E., Elith, J., Graham, C. H., Phillips, S. J., & Peterson, A. T. (2007a). What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? *Ecological Monographs*, *77*(4), 615–630.

Guisan, A., Graham, C., Elith, J., & Huettmann, F. (2007b). Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions*, *13*(3), 332–340. doi:10.1111/j.1472-4642.2007.00342.xGustafson, E. J. (1998). Quantifying landscape spatial pattern: What is the state of the art ? *Ecosystems*, *1*(2), 143–156.

Hamann, A., & Wang, T. (2006). Potential effects of climate change on ecosystem and tree species distribution in British Columbia. *Ecology*, *87*(11), 2773–2786.

Hansen, M. J., Franklin, S. E., Woudsma, C., & Peterson, M. (2001). Forest Structure Classification in the North Columbia Mountains Using the Landsat TM Tasseled Cap Wetness Component. *Canadian Journal of Remote Sensing*, *27*(1), 20–32. doi:10.1080/07038992.2001.10854916

Hawkins, B. A. (2012). Eight (and a half) deadly sins of spatial analysis. *Journal of Biogeography*, *39*(1), 1–9. doi:10.1111/j.1365-2699.2011.02637.x

Hortal, J., Lobo, J. M., & Jiménez-Valverde, A. (2007). Limitations of biodiversity databases: case study on seed-plant diversity in Tenerife, Canary Islands. *Conservation Biology*, *21*(3), 853–863. doi:10.1111/j.1523-1739.2007.00686.x

Immitzer, M., Atzberger, C., & Koukal, T. (2012). Tree species classification with Random Forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote Sensing*, *4*(12), 2661–2693. doi:10.3390/rs4092661

Jetz, W., Sekercioglu, C. H., & Watson, J. E. M. (2007). Ecological correlates and conservation implications of overestimating species geographic ranges. *Conservation Biology*, *22*(1), 110–119. doi:10.1111/j.1523-1739.2007.00847.x

Jiménez-Valverde, A. (2012). Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, *21*(4), 498–507. doi:10.1111/j.1466-8238.2011.00683.x

Kauth, R., & Thomas, G. (1976). The tasselled cap - a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In *LARS Symposia, Paper 159*.

Key, C. H., & Benson, N. C. (2006). Landscape Assessment (LA) Sampling and Analysis Methods. In *FIREMON: Fire Effects Monitoring and Inventory System*. Rocky Mountain Research Station, Ogden, UT: USDA Forest Service Gen. Tech. Rep. RMRS-GTR-164-CD: LA 1-51.

Kremen, C. (2005). Managing ecosystem services: what do we need to know about their ecology? *Ecology letters*, *8*(5), 468–79. doi:10.1111/j.1461-0248.2005.00751.x

Little, E.L. Jr. (1971). *Atlas of United States trees, volume 1, conifers and important hardwoods: U.S. Department of Agriculture Miscellaneous Publication 1146*.

Liu, C., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions, *3*(December 2004), 385–393.

Loveland, T. R., Merchant, J. W., Ohlen, D. O., & Brown, J. F. (1991). Development of a Land-Cover Characteristics Database for the Conterminous U.S . *Photogrammertic Engineering and Remote Sensing*, *57*(1), 1453–1463.

Maeda, E. E., Heiskanen, J., Thijs, K. W., & Pellikka, P. K. E. (2014). Season-dependence of remote sensing indicators of tree species diversity. *Remote Sensing Letters*, *5*(5), 404–412. doi:10.1080/2150704X.2014.912767

McLaughlan, M. S., Wright, R. A., & Jiricka, R. D. (2010). Field guide to the Ecosites of Saskatchewan's Provincial Forests. Prince Albert, Saskatchewan: Saskatchewan Ministry of Environment, Forest Service. http://www.environment.gov.sk.ca/forests

McPherson, J. M., & Jetz, W. (2007). Type and spatial structure of distribution data and the perceived determinants of geographical gradients in ecology: the species richness of African birds. *Global Ecology and Biogeography*, *16*(5), 657–667. doi:10.1111/j.1466-8238.2007.00318.x

McPherson, J. M., Jetz, W., & Rogers, D. (2004). The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *Journal of Applied Ecologyecology*, *41*, 811–823.

Natural Resources Canada. (2010). Canadian National Fire Database - agency fire data. Canadian Forest Service, Northern Forestry Centre, Edmonton, AB. http://cwfis.cfs.nrcan.gc.ca/ha/nfdb

Nelson, T. A., & Boots, B. (2005). Identifying insect infestation hot spots: an approach using conditional spatial randomization. *Journal of Geographical Systems*, *7*(2-3), 291–311.

Nelson, T. A., & Boots, B. (2008). Detecting spatial hot spots in landscape ecology. *Ecography*, *31*(5), 556–566. doi:10.1111/j.0906-7590.2008.05548.x

Nenzén, H. K., & Araújo, M. B. (2011). Choice of threshold alters projections of species range shifts under climate change. *Ecological Modelling*, *222*(18), 3346–3354. doi:10.1016/j.ecolmodel.2011.07.011

O'Sullivan, D., & Unwin, D. (2010). Area Objects and Spatial Autocorrelation. In *Geographic Information Analysis* (2nd ed.). Hoboken, New Jersey: John Wiley & Sons.

Pastor, J., Mladenoff, D. J., Haila, Y., Bryant, J., & Payette, S. (1996). Biodiversity and Ecosystem Processes in Boreal Regions. In H. A. Mooney, J. H. Cushman, E. Medina, O. E. Sala, & E.-D. Schulze (Eds.), *Functional Roles of Biodiversity: A Global Perspective*. Wiley.

Pfeifer-Meister, L., Bridgham, S. D., Little, C. J., Reynolds, L. L., Goklany, M. E., & Johnson, B. R. (2013). Pushing the limit: experimental evidence of climate effects on plant range distributions. *Ecology*, *94*(10), 2131–2137.

Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer Classification and Regression Tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, *9*(2), 181–199. doi:10.1007/s10021-005-0054-1

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.

Roy, D. P., Ju, J., Kline, K., Scaramuzza, P. L., Kovalskyy, V., Hansen, M., et al. (2010). Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, *114*(1), 35–49. doi:10.1016/j.rse.2009.08.011

Roy, D. P., Wulder, M. A., Loveland, T. R., C.E., W., Allen, R. G., Anderson, M. C., et al. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, *145*, 154–172. doi:10.1016/j.rse.2014.02.001

Running, S., Loveland, T., & Pierce, L. (1994). A vegetation classification logic based on remote sensing for use in global biogeochemical models. *Ambio*, *23*(1), 77–81.

Saskatchewan Environment - Forest Service. (2004). *Saskatchewan Forest Vegetation Inventory - Forest Planning Manual, Version 4.0*.

Saskatchewan Ministry of Environment. (2009). *Saskatchewan's 2009 State of the Environment Report*.

Song, C., Schroeder, T., & Cohen, W. (2007). Predicting temperate conifer forest successional stage distributions with multitemporal Landsat Thematic Mapper imagery. *Remote Sensing of Environment*, *106*(2), 228–237. doi:10.1016/j.rse.2006.08.008

Song, C., & Woodcock, C. (2003). Monitoring forest succession with multitemporal Landsat images: factors of uncertainty. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(11), 2557–2567.

Stockwell, D. R., & Peterson, A. T. (2002). Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, *148*(1), 1–13. doi:10.1016/S0304-3800(01)00388-X

Storey, J., Scaramuzza, P., Schmidt, G., & Barsi, J. (2005). Landsat 7 Scan Line Corrector-Off Gap-Filled Product Development. In *Proceedings from 16th William T. Pecora Memorial Symposium: Global Priorities in Land Remote Sensing*. Oct 23-27, 2005, Sioux Falls, South Dakota.

Thuiller, W., Lavorel, S., Araújo, M. B., Sykes, M. T., & Prentice, I. C. (2005). Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(23), 8245–50. doi:10.1073/pnas.0409902102

Turner, M. (1989). Landscape ecology: the effect of pattern on process. *Annual review of ecology and systematics*, *20*, 171–197.

Ustin, S. L., & Xiao, Q. F. (2001). Mapping successional boreal forests in interior central Alaska. *International Journal of Remote Sensing*, *22*(9), 1779–1797. doi:10.1080/01431160118269

Van Aardt, J. A. N., & Wynne, R. H. (2001). Spectral Separability among Six Southern Tree Species. *Photogrammetric Engineering and Remote Sensing*, *67*(12), 1367–1375.

Verbyla, D. L., & Hammond, T. O. (1995). Conservative bias in classification accuracy assessment due to pixel- by-pixel comparison of classified images with reference grids. *International Journal of Remote Sensing*, *16*(3), 581–587.

White, J. C., & Wulder, M. A. (2013). The Landsat observation record of Canada: 1972 – 2012. *Canadian Journal of Remote Sensing*, *39*(6), 455–467.

White, J. C., Wulder, M. A., Hobart, G. W., Luther, J. E., Hermosilla, T., Griffiths, P., et al. (2014). Pixel-based image compositing for large-area dense time series applications and science. *Canadian Journal of Remote Sensing*, *43*(3), 192–212. doi:10.1080/07038992.2014.945827

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., & Guisan, A. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, *14*(5), 763–773. doi:10.1111/j.1472-4642.2008.00482.x

Woodcock, C., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., et al. (2008). Free access to Landsat imagery. *Science*, *320*, 1011–1012.

Wulder, M. A., & Coops, N. C. (2014). Make Earth observations open access. *Nature*, *513*, 30–31.

Wulder, M. A., Kurz, W. A., & Gillis, M. (2004a). National level forest monitoring and modeling in Canada. *Progress in Planning*, *61*(4), 365–381. doi:http://dx.doi.org/10.1016/S0305-9006(03)00069-2

Wulder, M. A., Hall, R. J., Coops, N. C., & Franklin, S. E. (2004b). High spatial resolution remotely sensed data for ecosystem characterization. *BioScience*, *54*(6), 511–521. doi:10.1641/0006-3568(2004)054[0511:HSRRSD]2.0.CO;2

Wulder, M. A., Skakun, R. S., Kurz, W. A., & White, J. C. (2004c). Estimating time since forest harvest using segmented Landsat ETM+ imagery. *Remote Sensing of Environment*, *93*(1-2), 179–187. doi:10.1016/j.rse.2004.07.009

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., & Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, *122*, 2–10. doi:10.1016/j.rse.2012.01.010

Wulder, M. A., White, J. C., Coops, N. C., Nelson, T. A., & Boots, B. (2007). Using local spatial autocorrelation to compare outputs from a forest growth model. *Ecological Modelling*, *209*(2-4), 264–276. doi:10.1016/j.ecolmodel.2007.06.033

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

Abstract

Traditionally, forest inventory and ecosystem mapping at local to regional scales relies on manual interpretation of aerial photographs, based on standardized, expert-driven classification schemes that provide the information needed for ecosystem management. These current approaches 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 based on aerial photographs. The clusters describe ecosystems ranging from open shrublands to dense, productive forest and include a riparian zone and many wetter and wetland ecosystems. The 18 clusters provide complementary, spatially-explicit information for characterizing the landscape as a complex mosaic of units defined by topography and vegetation structure. This study demonstrates that using various types of high spatial resolution remotely sensed data in a quantitative classification for mapping ecosystems at local scales may offer valuable information regarding the distribution of wetlands, and vegetation distribution and height.

4.1 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). Boundaries between ecosystems are, in reality, most often gradual and fuzzy; attributes such as species composition overlap along environmental gradients (Whittaker 1967, Kent 1997). The science of delineating regions (e.g., ecological or biogeographical) in geographical space is referred to as *regionalization* (Loveland and Merchant 2004; Olstad 2012) and is important for the understanding and management of the natural world (Mackey et al. 2007; McMahon et al. 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.

Global biodiversity monitoring and conservation planning uses region-based maps at the broadest scales, with boundaries drawn based on global climate, geology, and evolutionary history (e.g., Olson et al. 2001). 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 and Csillag 2002; Leathwick et al. 2003; Sayre et al. 2009). At the finest scale, 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 et al. 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., Resources Inventory Committee 1998; Canadian Forest Service 2001; Resource Information Management Branch 2005) 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 high spatial resolution 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 et al. 2007; Morgan et al. 2010).

Remote sensing data can capture various structural, compositional, and functional aspects of ecosystems. For instance, structural classes such as conifer versus broadleaf forest, and young versus old forest are generally well distinguished with multispectral imagery (e.g., Johansen et al. 2007; Valeria et al. 2014). High spatial resolution multispectral imagery can distinguish tree species with high accuracy in some cases (e.g., Immitzer et al. 2012), although the task is notoriously difficult due to the overlapping and highly variable nature of the spectral signatures of many vegetation species (Leckie et al. 2005; Wulderet al. 2004). Vegetation indices, such as the Normalized Difference Vegetation Index (NDVI) (Rouse et al. 1974), correlate with patterns of net primary productivity (Goward et al. 1985). NDVI facilitates the separation of vegetated from non-vegetated areas and the classification of vegetation structural classes (e.g., Running et al. 1994; Defries et al. 1995; Gillespie et al. 2006; Johansen et al. 2007) and of wetlands (e.g., Dechka et al. 2002; Barron et al. 2014; White et al. 2015).

Light Detection And Ranging (LiDAR) technology captures highly accurate, direct measurements of three-dimensional vegetation structure (Lim et al. 2003; van Leeuwen and Nieuwenhuis 2010) and thus can be an effective tool in forest inventory and management

(Wulderet al. 2008). 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). Structural attributes can affect or be indicative of differences in forest composition (Heinzel and Koch 2011; Holmgren and Persson 2004; Kennel et al. 2013) and forest age (Lefsky et al. 1999), as well as of functions related to forest hydrology (Roth et al. 2007; Varhola and Coops 2013) and carbon sequestration (Asner et al. 2012). Researchers have highlighted the importance of LiDAR for mapping wetlands. The high spatial resolution of LiDAR-derived DEMs can lead to improved wetland delineation compared to use of aerial photography, especially in areas of low topographic variation (Hogg and Holland 2008; Maxa and Bolstad 2009). However, in a contrasting study from Minnesota, Knight et al. (2013) found that the high resolution LiDAR topographic data did not significantly improve results over other sources of topographic data.

Beyond mapping individual components of ecosystems (structure, function, or composition), the variety of attributes available from remote sensing can be combined to enhance our ability to quantitatively map and describe complex ecological units. Combining various types of remotely sensed data in a quantitative (statistical) regionalization, researchers have effectively captured environmental domains over large regions (e.g., Handcock and Csillag 2002; Coops et al. 2009; Fitterer et al. 2012; 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 and Hoffman 2004; Kupfer et al. 2012; Long et al. 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. However, research using vegetation structural data from LiDAR 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 these 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.

4.2 Methods

4.2.1 Study Area

Calvert and Hecate Islands (total 37,433 ha) are remote islands on the central coast of British Columbia, Canada (Figure 4.1). 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 are classified as the Mountain Hemlock zone, Wet Hypermaritime Subzone (MHwh). Gridded climate data for the region (Wang et al. 2012) indicate that average (1990–2012) annual temperature is 7.6°C for the CWHvh2 and 4.8°C for the MHwh. 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 coastal forests (Banner et al. 2005).

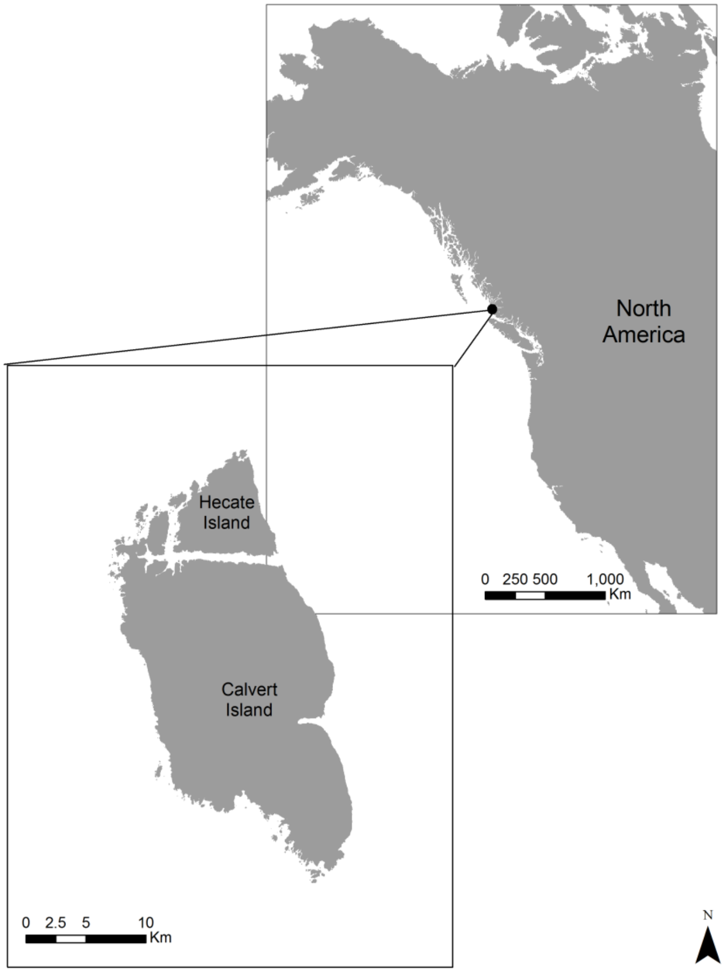


Figure 4.1. Study area in coastal British Columbia, Canada

4.2.2 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, and one that is relevant to forest management and landscape planning. Airborne LiDAR data were acquired 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 ±26°, 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^2 and a standard deviation of 1.07 pt/m^2. A small number of data voids in the LiDAR coverage were present and were excluded from the analysis.

LiDAR data were separated into ground and non-ground classes using an automated classifier (Terrasolid TerraScan) followed by manual refinement (Merrick Advanced Remote Sensing - MARS). A Digital Elevation Model (DEM) was extracted from a Triangular Irregular Network (TIN) of ground-classified points in LAStools. The DEM was then used to calculate the following terrain indices: slope (%), the Topographic Radiation ASPect (TRASP) (Roberts and 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 4.1). Elevation, slope, aspect, topographic position, and a topographic wetness index are useful inputs in automated ecosystem classification and predictive vegetation mapping because these variables directly and indirectly influence plant growth and 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., MacMillan et al. 2007;

Chastain and Struckhoff 2008; Dobrowski et al. 2008; Fraser et al. 2012), and standard ecosystem classification in British Columbia at the local scale relies heavily on topographic concepts (Banner et al. 1996).

Based on previous research (e.g., Magnussen and Boudewyn 1998; Næsset 2002; Hopkinson and Chasmer 2009), non-ground LiDAR points were converted to i) a canopy height model representing maximum canopy height and ii) a variety of area-based canopy statistics representing height and density (i.e., vertical and horizontal heterogeneity). Area-based canopy statistics utilized in this analysis were available only for areas with vegetation greater than 2 m in height (we refer to these as forested areas), and included average height (mean height of points above 2 m), coefficient of variation of height (2nd central moment about the mean), and gap fraction (proportion of laser points below 2 m).

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[1] , as well as the NDVI layer (values < 0.2) to remove water and un-vegetated pixels.

---

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

Table 4.1. 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	2nd 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 (Weiss 2001; Tagil and Jenness 2008; De Reu et al. 2013).
Topographic Radiation ASPect (TRASP)	TRASP = (1 - cos((3.1416/180)(aspect-30)))/2 where aspect is in degrees (Roberts and Cooper 1989).	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 Wetness Index (TWI)	TWI = ln(specific catchment area / tan(slope)). Calculated using the D-infinity algorithm (Tarboton 1997).	Higher values are considered wetter, lower values drier.

4.2.3 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 et al. 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 influence by climate, not edaphic features). *Variants* may be used to refine the classification even further, by distinguishing regional 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 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 these local site series (each given a unique two letter and two digit code). 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). Among other information, these field surveys indicated BEC site series. After removing data points for which the site series attribution was at all uncertain, a total of 56 field plots and 110 visual, helicopter-based survey plots were available as reference data for this analysis. 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).

4.2.4 Unsupervised regionalization

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 4.2). 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.

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 and Harabasz 1974) and the Average Silhouette Width (ASW) (Table 4.3). 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 x 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.

Table 4.2. 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

4.2.5 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

$$D = 1 - [\mathrm{sum}(n/N)^2]$$

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 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 and 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.

4.3 Results

4.3.1 Unsupervised regionalization

We identified 12 forested clusters (Clusters 1–12) and six non-forested clusters (Clusters 13–18) (Figure 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 4.3), 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.

Table 4.3. 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)[a]	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)[a]	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)[b]	BIC identifies the model (cluster solution) that would most likely produce the observed data, penalizing for complexity (number of clusters * number of observations). Optimal number of clusters maximizes this index (i.e., BIC will be higher for smaller numbers of clusters).	2	2

[a] ASW and CH were calculated in R based on *k*-means clustering.
[b] BIC was calculated in SPSS based on Two-Step clustering.

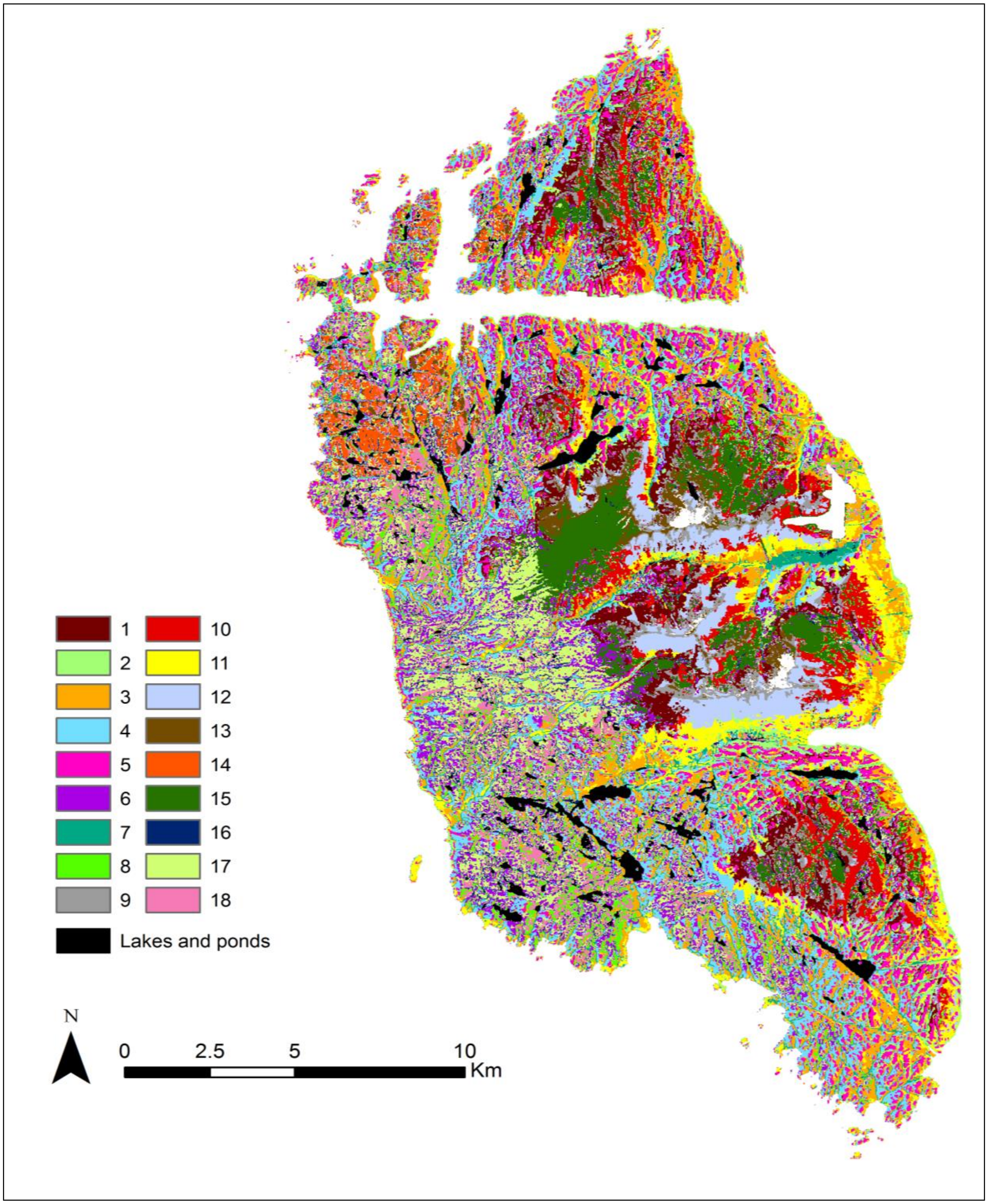


Figure 4.2. Eighteen clusters representing a range of forested (Clusters 1 to 12) and non-forested (Clusters 13 to 18) terrestrial ecosystems on Calvert and Hecate Islands. White areas within the land mass are data voids.

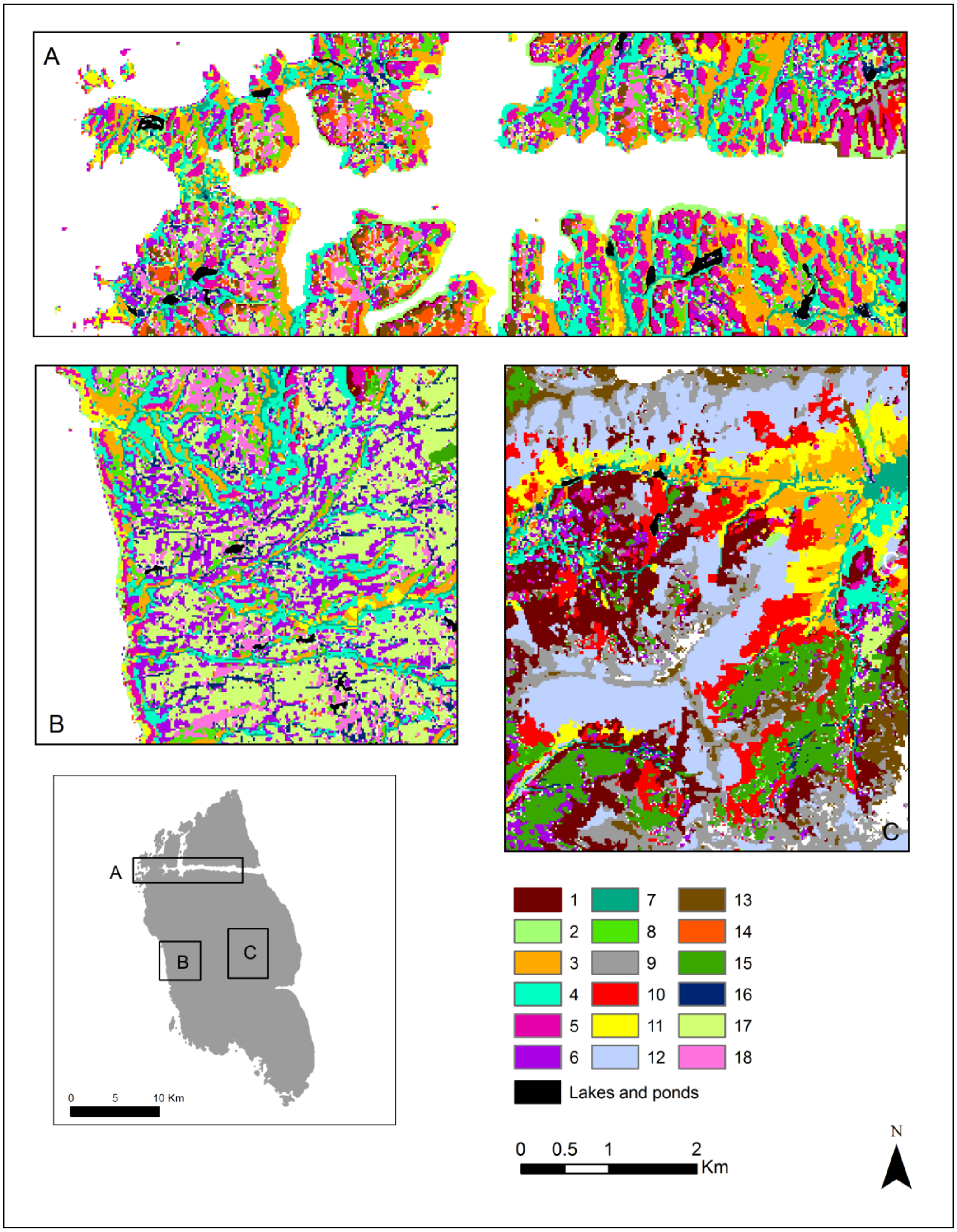


Figure 4.3. Zoomed in view of 18 clusters on Calvert and Hecate Islands. Clusters 1 to 12 are forested, while Clusters 13 to 18 are non-forested.

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 4.4, Figures 4.4 and 4.5). 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 = 85m), 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 (Figure 4.5).

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).

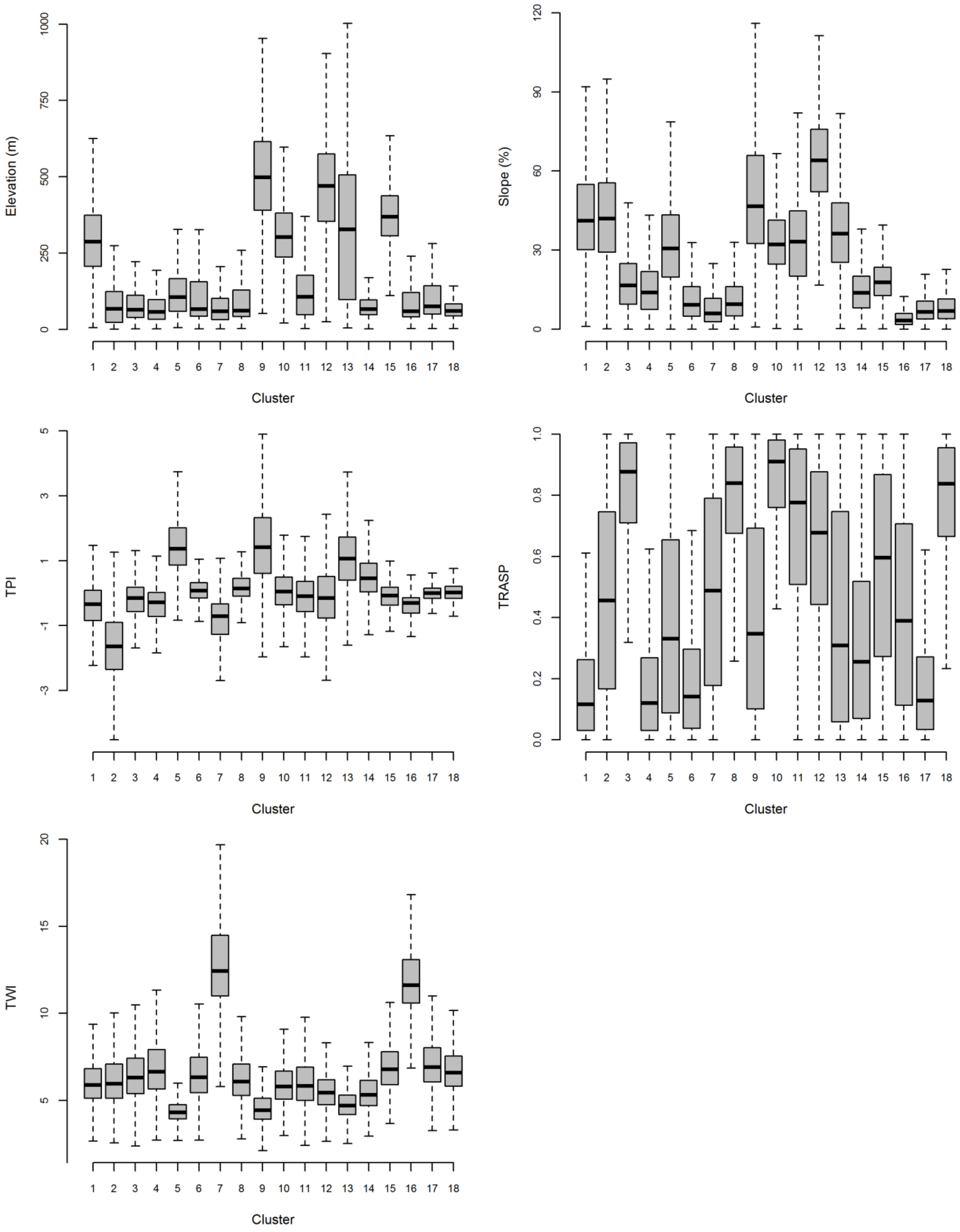


Figure 4.4. Distribution of clusters across each LiDAR-derived terrain index. Clusters 1 to 12 are forested, while Clusters 13 to 18 are non-forested.

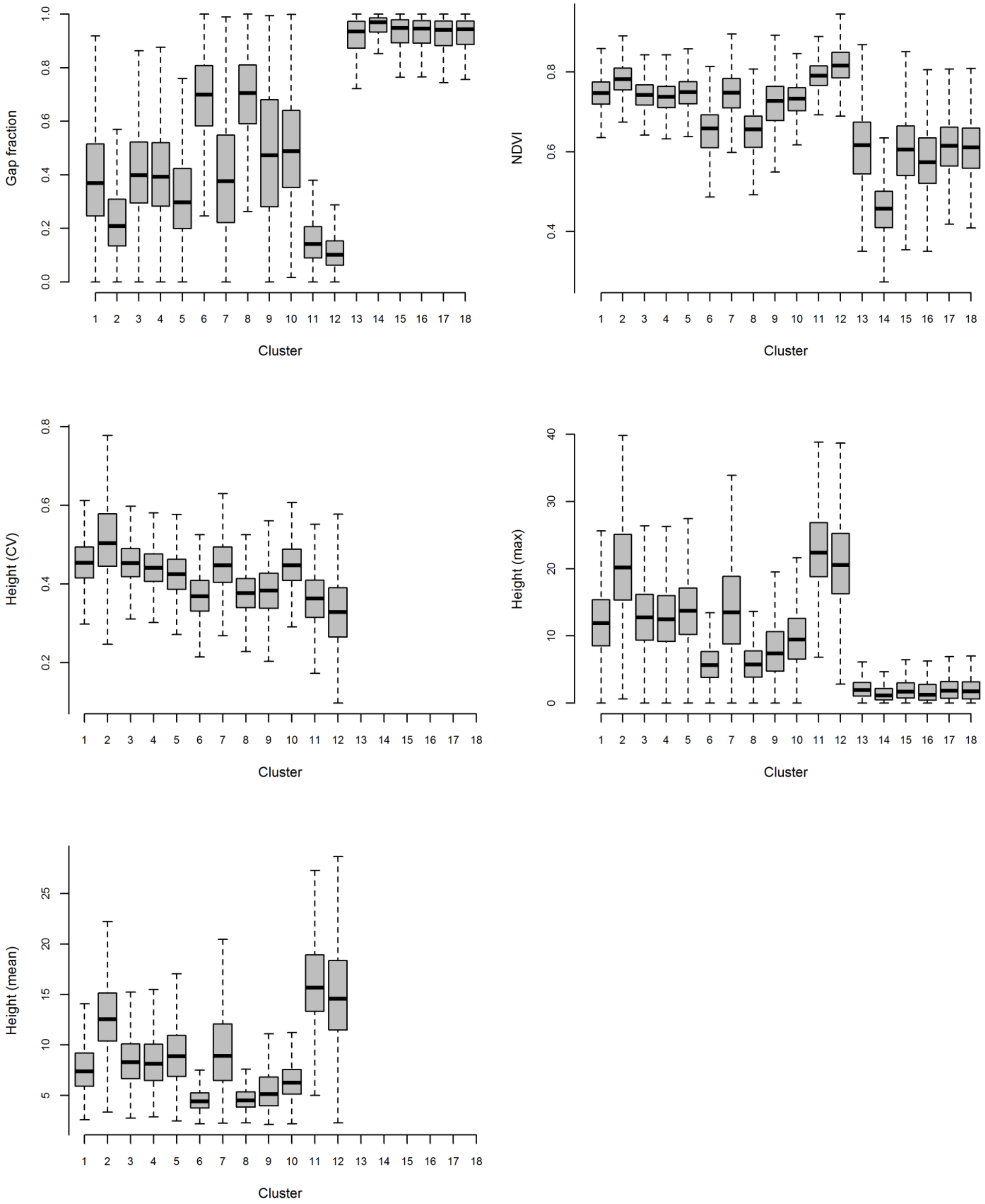


Figure 4.5. Distribution of clusters across LiDAR and RapidEye vegetation data. Clusters 1 to 12 are forested, while Clusters 13 to 18 are non-forested.

Table 4.4. Mean values of remotely sensed inputs and descriptors of 18 clusters. Clusters 1 through 12 are forested; 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	1,936	292	0.39	0.46	12.08	7.72	0.75	44	-0.41	0.16	6.07
2	1,160	86	0.24	0.52	20.30	13.00	0.78	43	-1.65	0.47	6.29
3	3,903	81	0.42	0.46	12.80	8.48	0.74	18	-0.20	0.83	6.52
4	3,696	73	0.41	0.44	12.65	8.41	0.74	15	-0.38	0.16	6.93
5	2,900	119	0.32	0.42	13.74	9.04	0.75	33	1.52	0.39	4.39
6	2,908	113	0.69	0.37	5.99	4.73	0.64	12	0.08	0.18	6.67
7	1,473	85	0.40	0.45	14.29	9.77	0.74	8	-0.88	0.49	12.89
8	1,921	100	0.69	0.38	5.88	4.76	0.65	11	0.20	0.80	6.32
9	1,020	506	0.48	0.38	7.97	5.73	0.71	51	1.55	0.41	4.59
10	2,039	311	0.50	0.45	9.67	6.51	0.73	34	0.07	0.86	5.96
11	2,206	120	0.16	0.36	23.38	16.61	0.79	33	-0.10	0.69	6.13
12	1,326	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	1,902	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	2,449	105	0.93	n/a	1.93	n/a	0.60	8	-0.02	0.16	7.20
18	1,674	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

4.3.2 Map Comparisons

The clusters contained information that is 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 4.5, Supplementary Figure 4.S1) and each TEM site series contained multiple clusters (Supplementary Figure 4.S2). 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 (Figure 4.6). 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 (Figure 4.6).

Areas mapped as rivers (RI) in the TEM were entirely captured by Cluster 7 (Figures 4.7, 4.S2), 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 (MHwh) biogeoclimatic subzone (e.g., MHwh/01(MB), MHwh/06(MD), MHwh/02(MM), MHwh 03(MR), MHwh/00(MS), and MHwh/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 (Figure 4.6). 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 4.6, Figure 4.7). 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 % to 27% of the islands.

Table 4.5. Comparison of 18 clusters to existing Terrestrial Ecosystem Mapping data

Cluster	**Dominant TEM unit[a,b] (% 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)

[a] 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.

[b] See pie charts in Fig S1 for further details.

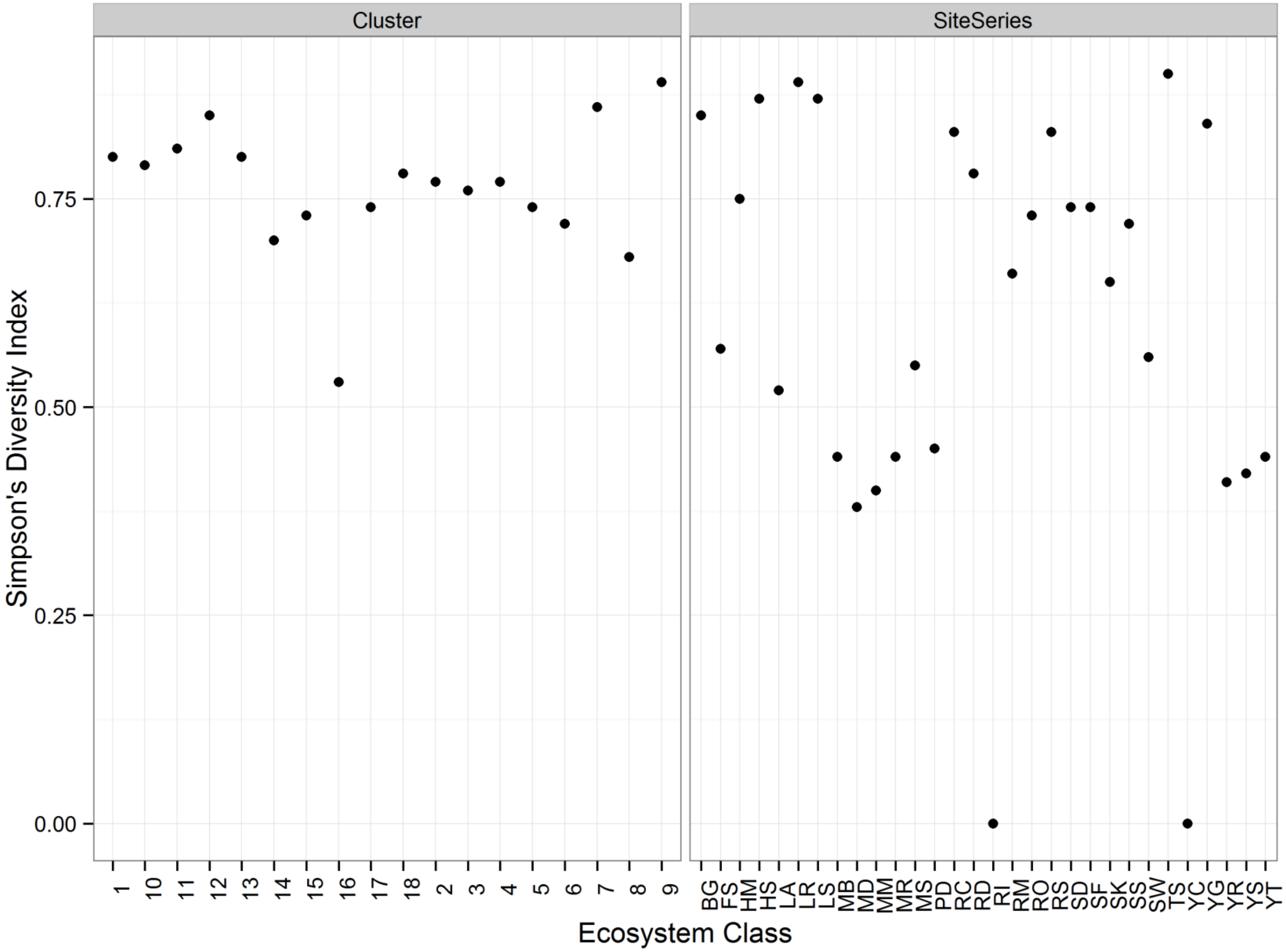


Figure 4.6. Diversity of TEM site series (with respect to cluster composition) and diversity of clusters (with respect to TEM site series). Lower values of the Simpson's Diversity Index represent greater similarity between TEM and cluster types

Table 4.6. 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**[a,b]	**Percentage of study area (clusters)**	**Percentage of study area (TEM)**
Shrub/herb upland	13	MHwh/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) MHwh/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) MHwh/06(MD) MHwh/08(YS) MHwh/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) MHwh/00(YR) MHwh/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) MHwh/01(MB) MHwh/03(MR) MHwh/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

---

[a] MHwh = 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 and Klinka 1994; MacKenzie and Moran 2004).

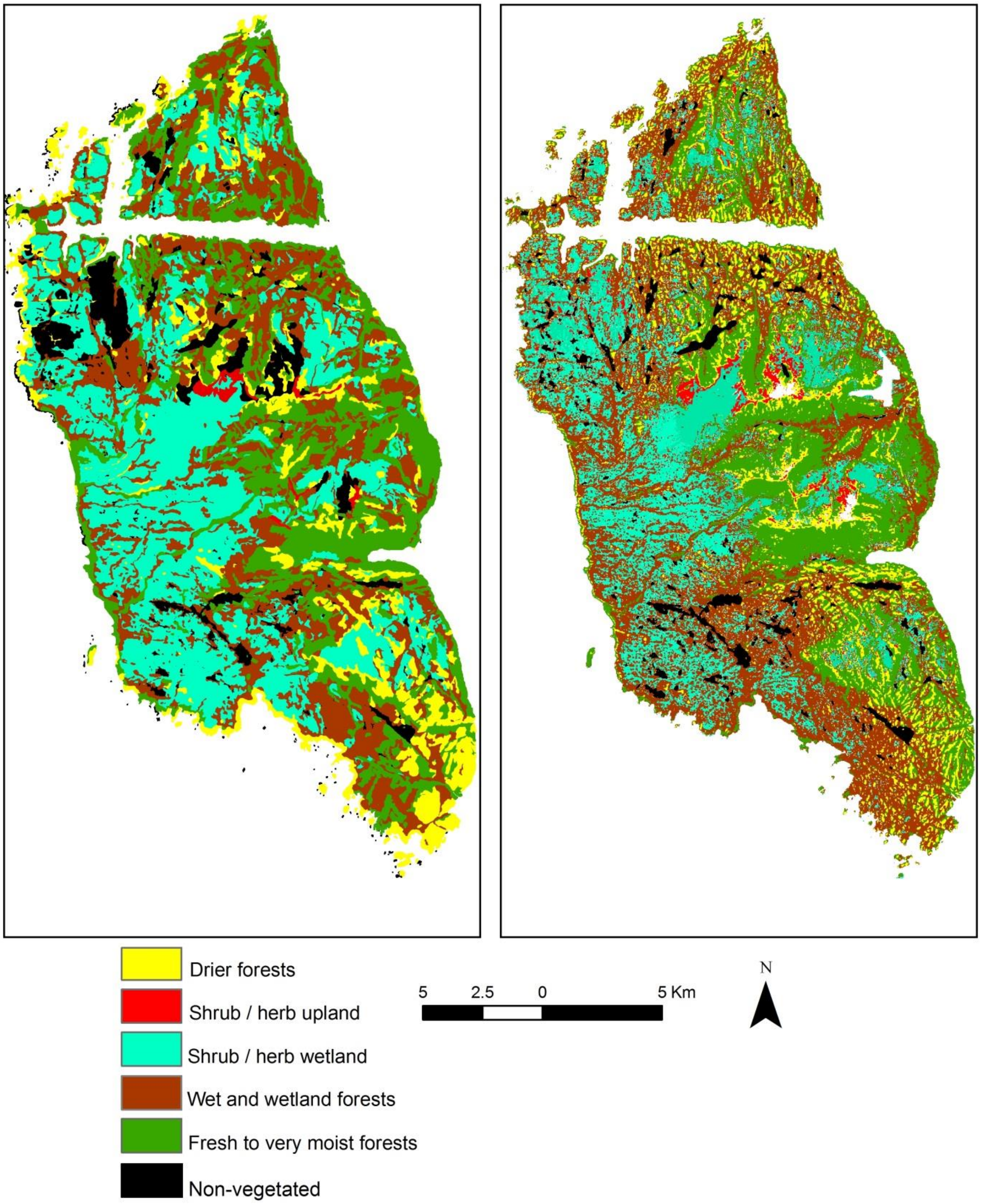


Figure 4.7. Generalized ecosystem classes as depicted by the expert-based classification (Terrestrial Ecosystem Map) in the left panel, and the unsupervised regionalization on the right

4.4 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 provide information that complements or augments that which is found in 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, 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 4.4, Figures 4.4 and 4.5). 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 et al. 2007; Diggins and Catterlin 2014; Fekedulegn et al. 2003; Holland and Steyn 1975), as well as to soil microbial communities (Carletti et al. 2008).

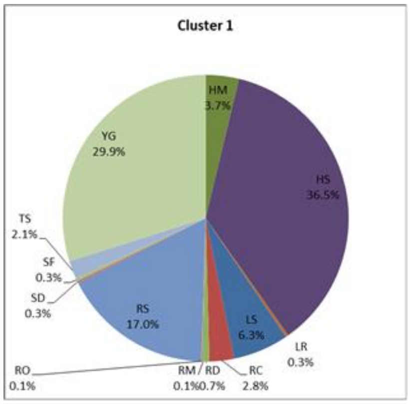
Another unique and promising aspect of the quantitative regionalization demonstrated in this research relates to the delineation of wetland (defined here to include wet forest) ecosystems. Wetlands provide a range of important ecological processes and ecosystem services, including critical habitat, water 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 and 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 et al. 2002; Ozesmi and Bauer 2002) or relative to a standard minimum mapping unit, such as the 0.5 to 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 association in the field (such as acidity/alkalinity and magnitude of lateral and vertical water flow (MacKenzie and Moran 2004) are not apparent in multispectral imagery, particularly under heavy vegetation cover (Dechka et al. 2002; Rosenqvist et al. 2007). In this study, we predicted the location of wetlands using terrain indices generated from the high spatial resolution LiDAR data, combined with multi-spectral imagery from RapidEye (helpful in detecting water and wet soil). 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 (Figures 4.7, and 4.S1). 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., Figures 4.2, 4.3, 4.7 and 4.S2). Future research could examine the potential of using even higher spatial resolution (e.g., < 5m) 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.

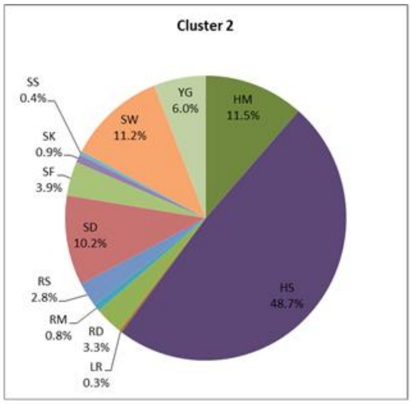
A final key and uniquely informative aspect of the quantitative regionalization approach is the ability to quantify and map spatial variation in forest height and productivity at 20 m spatial resolution. Forest inventories generally do measure stand height, basal area, canopy cover, and so on, however 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.

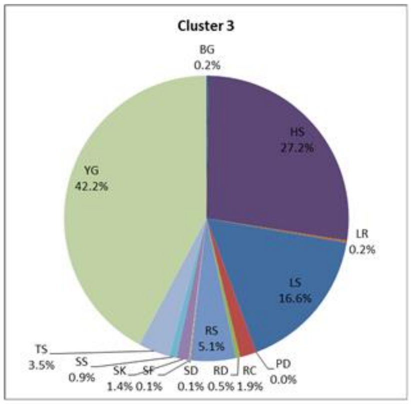
Understanding the unique information contained within an expert-driven compared to a quantitative remotely sensed ecological classification has been the subject of several contemporary studies (e.g., Thomas et al. 2002). There are several reasons for differences between an expert-driven and a quantitative 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 (Schmidtlein et al. 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 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 et al. 2002; Vierling et al. 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 useful to incorporate hyperspectral remotely sensed data, which is well suited to mapping vegetation composition (e.g., Jones et al. 2010).

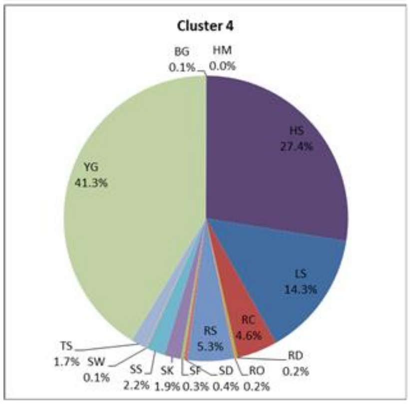
4.5 Conclusion

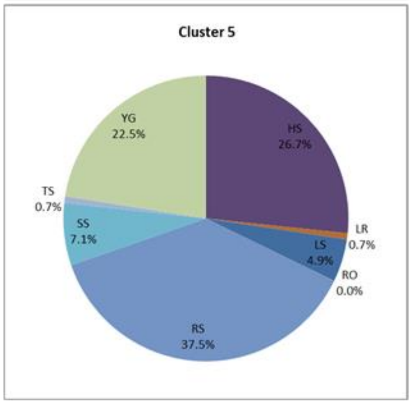
This study has demonstrated that LiDAR and high spatial resolution multi-spectral imagery can be combined using a quantitative regionalization to map a variety of ecosystem types across a heterogeneous landscape. Given the quantitative nature of our approach, it is transparent and easily repeatable, 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 is 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 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.

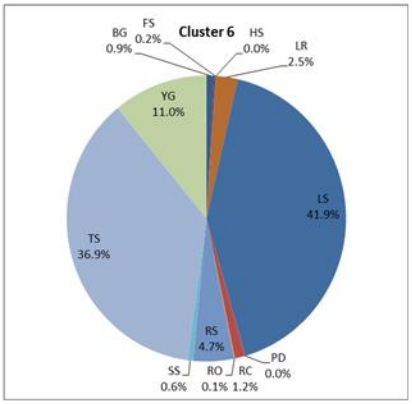

Cluster 1

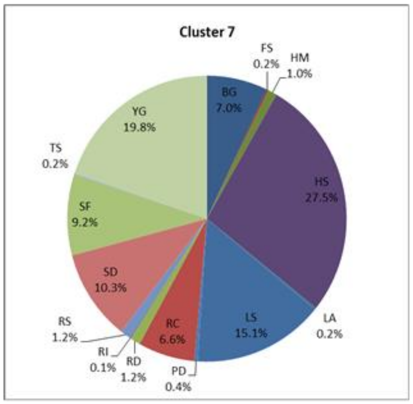

Cluster 2

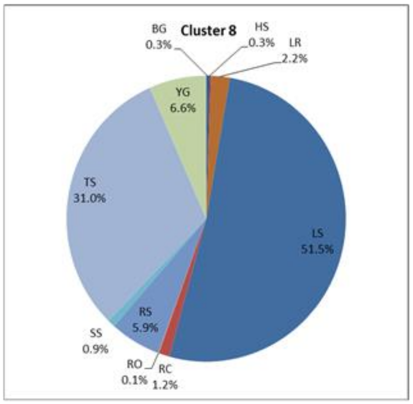

Cluster 3

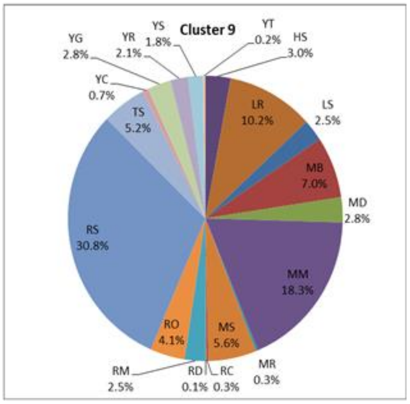

Cluster 4

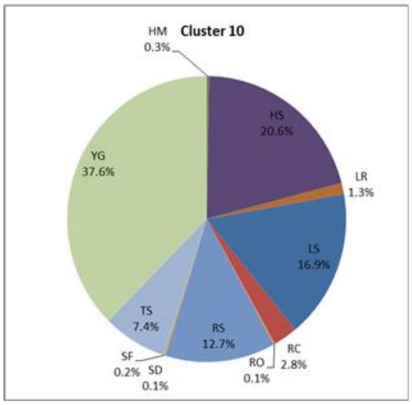

Cluster 5

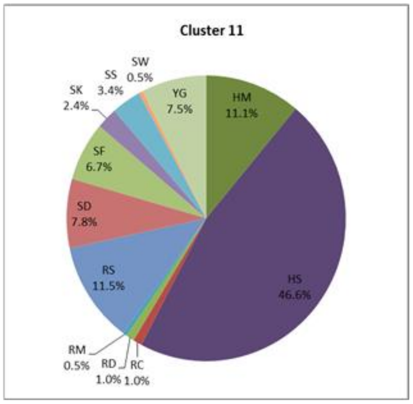

Cluster 6

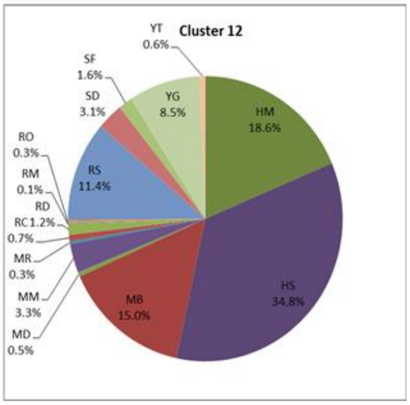

Cluster 7


Cluster 8


Cluster 9


Cluster 10


Cluster 11


Cluster 12

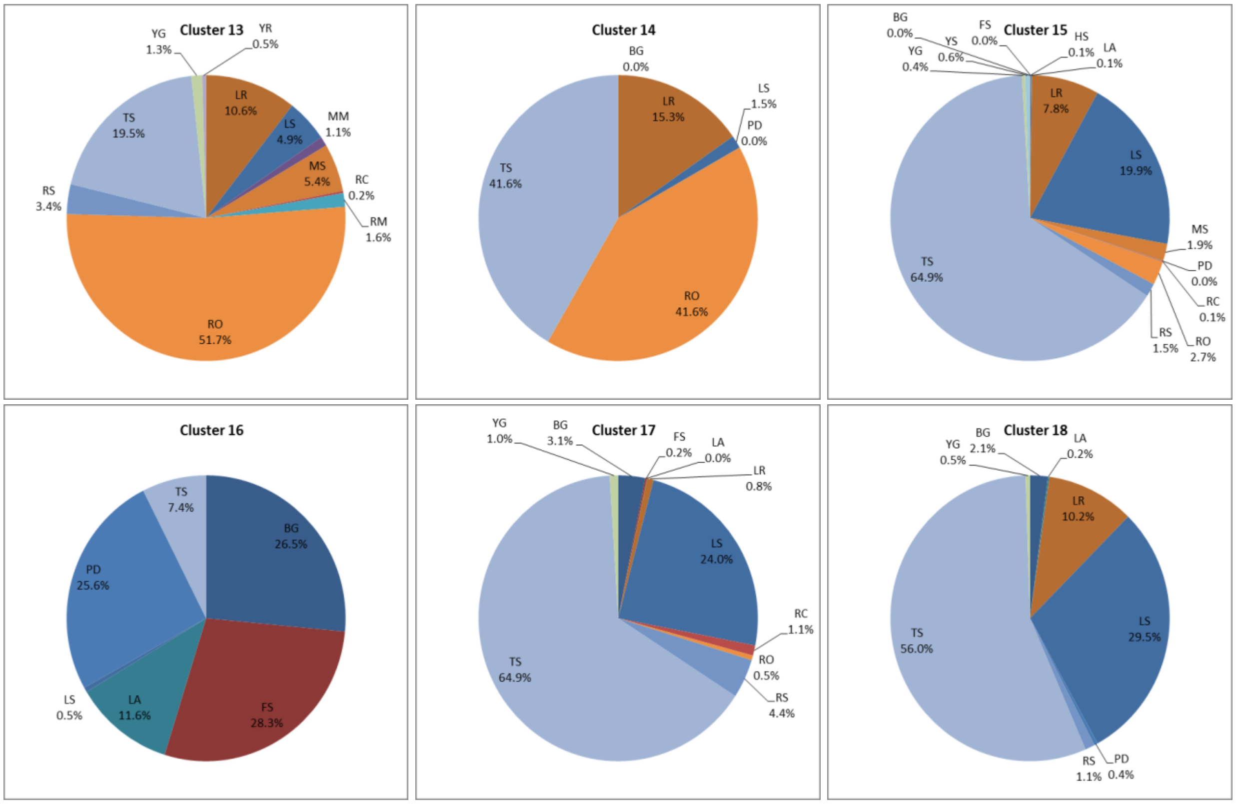
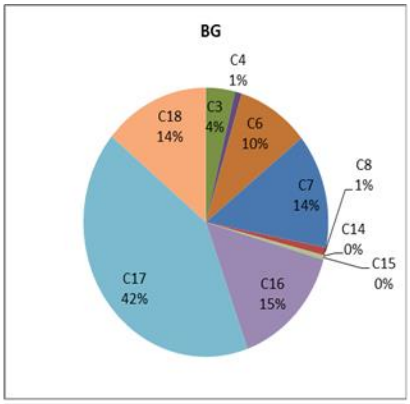
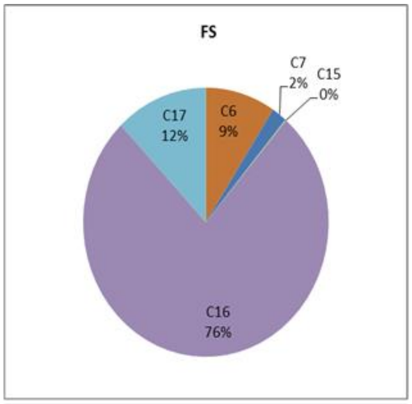
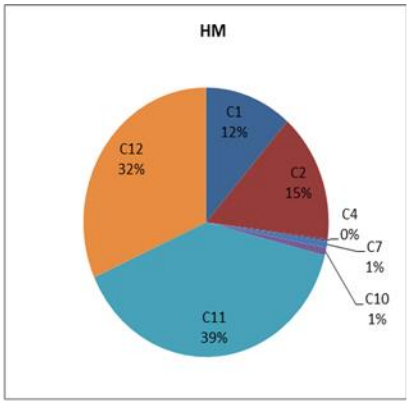
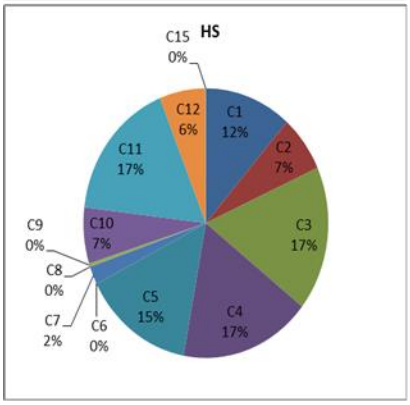


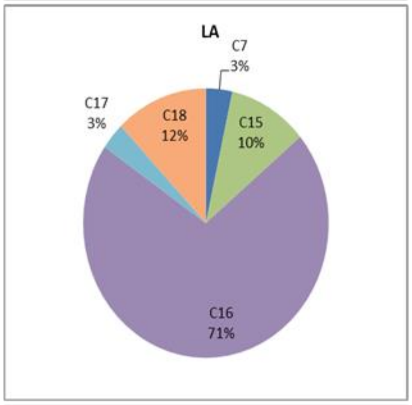
Figure 4.S1 Composition of each cluster in terms of provincial map units (TEM site series). Rare TEM site series and very small polygons were excluded

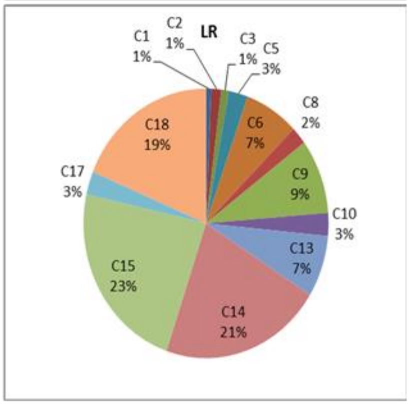

BG
C4
1%
C3
4%
C6
10%
C8
1%
C7
14%
C14
0%
C15
0%
C18
14%
C17
42%
C16
15%

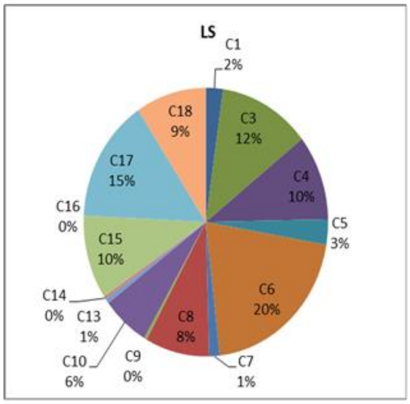

FS
C7
2%
C15
0%
C17
12%
C6
9%
C16
76%

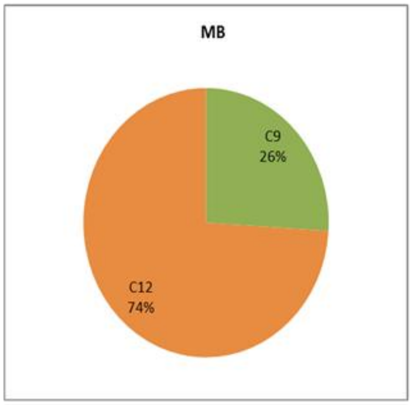

HM
C1
12%
C12
32%
C2
15%
C4
0%
C7
1%
C10
1%
C11
39%

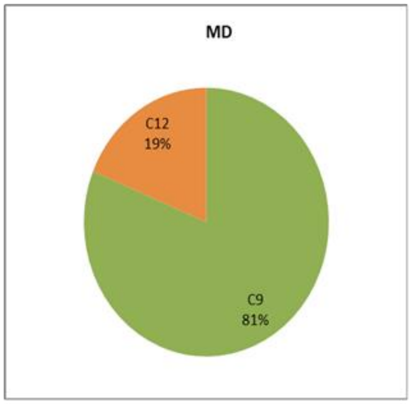

HS
C15
0%
C12
6%
C1
12%
C11
17%
C2
7%
C3
17%
C9
0%
C10
7%
C8
0%
C5
15%
C4
17%
C7
2%
C6
0%

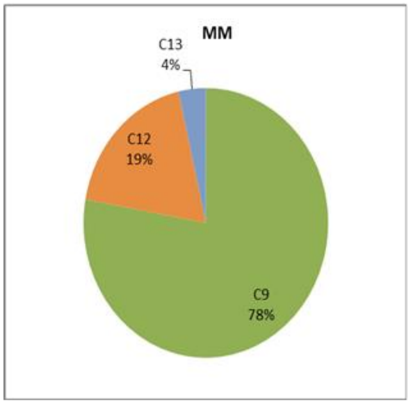

LA
C7
3%
C17
3%
C18
12%
C15
10%
C16
71%

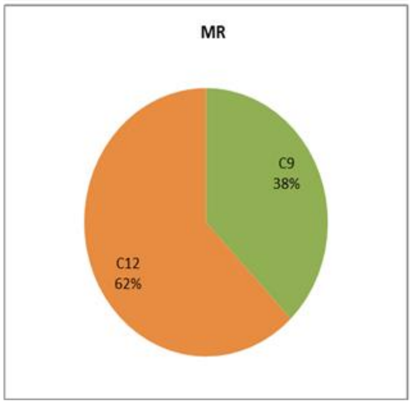

C1
1%
C2
1%
LR
C3
1%
C5
3%
C8
2%
C18
19%
C6
7%
C9
9%
C17
3%
C10
3%
C13
7%
C15
23%
C14
21%

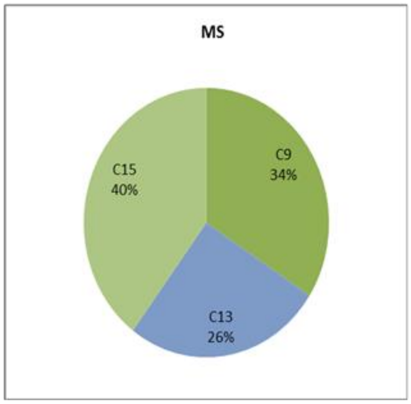

LS
C1
2%
C18
9%
C3
12%
C17
15%
C4
10%
C16
0%
C5
3%
C15
10%
C6
20%
C14
0%
C13
1%
C8
8%
C10
6%
C9
0%
C7
1%

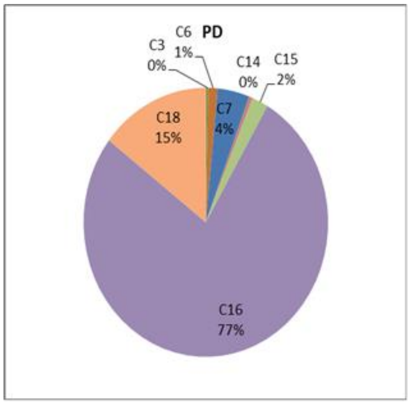

MB
C9
26%
C12
74%

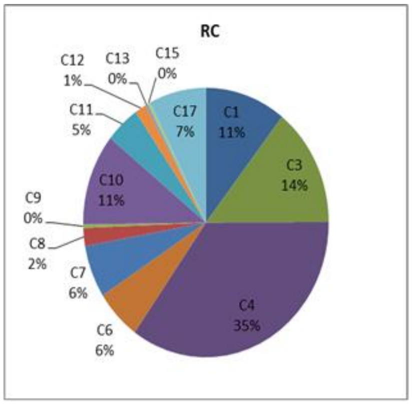

MD
C12
19%
C9
81%

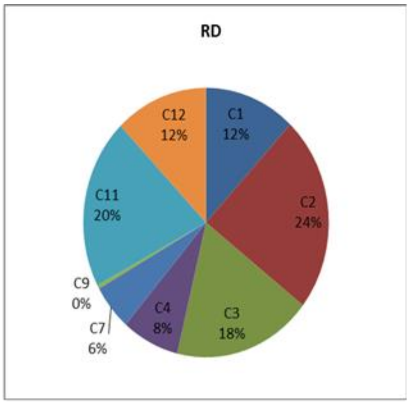

C13
4%
MM
C12
19%
C9
78%

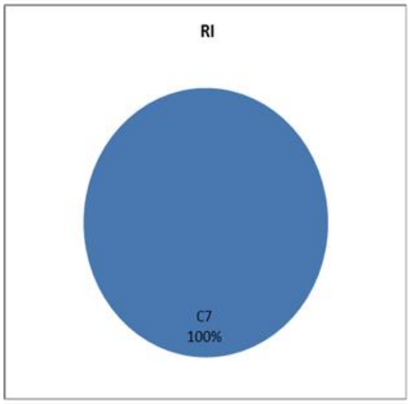

MR
C9
38%
C12
62%

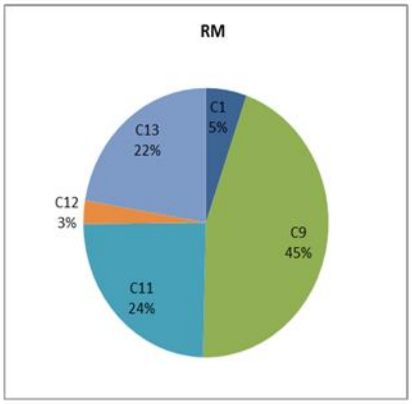

MS
C15
40%
C9
34%
C13
26%

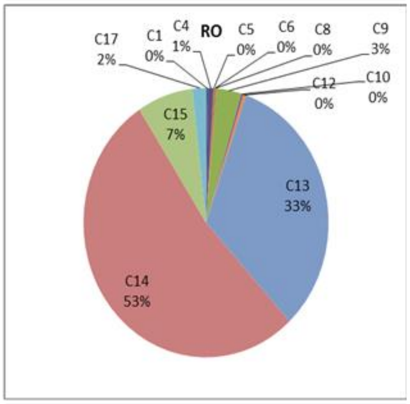

PD
C6 1%
C3 0%
C14 0%
C15 2%
C18 15%
C7 4%
C16 77%

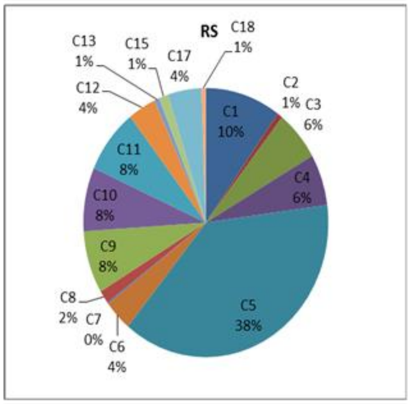

RC
C12 1%
C13 0%
C15 0%
C11 5%
C17 7%
C1 11%
C3 14%
C10 11%
C9 0%
C8 2%
C7 6%
C6 6%
C4 35%

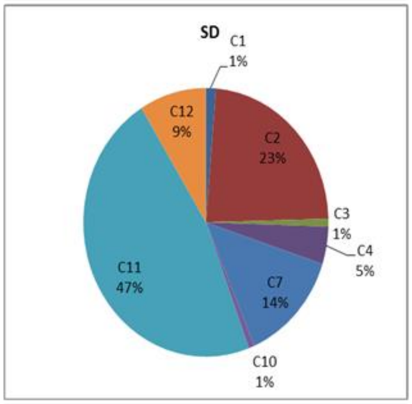

RD
C12 12%
C1 12%
C11 20%
C2 24%
C9 0%
C7 6%
C4 8%
C3 18%

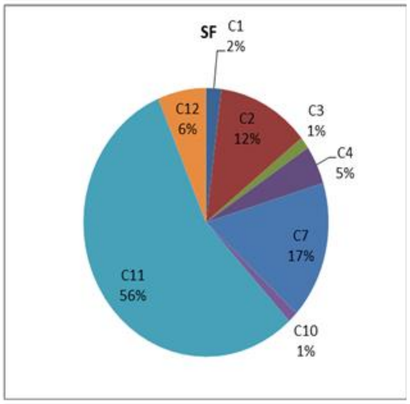

RI
C7 100%

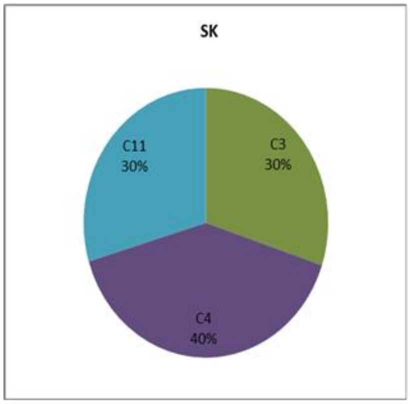

RM
C13 22%
C1 5%
C12 3%
C9 45%
C11 24%

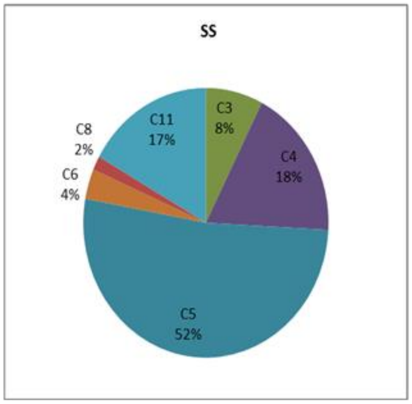

RO
C17 2%
C1 0%
C4 1%
C5 0%
C6 0%
C8 0%
C9 3%
C12 0%
C10 0%
C15 7%
C13 33%
C14 53%


RS
C13 1%
C15 1%
C17 4%
C18 1%
C12 4%
C2 1%
C3 6%
C1 10%
C11 8%
C4 6%
C10 8%
C9 8%
C8 2%
C7 0%
C6 4%
C5 38%


SD
C1 1%
C12 9%
C2 23%
C3 1%
C4 5%
C11 47%
C7 14%
C10 1%


SF
C1 2%
C12 6%
C2 12%
C3 1%
C4 5%
C7 17%
C11 56%
C10 1%


SK
C11 30%
C3 30%
C4 40%


SS
C8 2%
C11 17%
C3 8%
C6 4%
C4 18%
C5 52%


SW
C4 30%
C11 70%

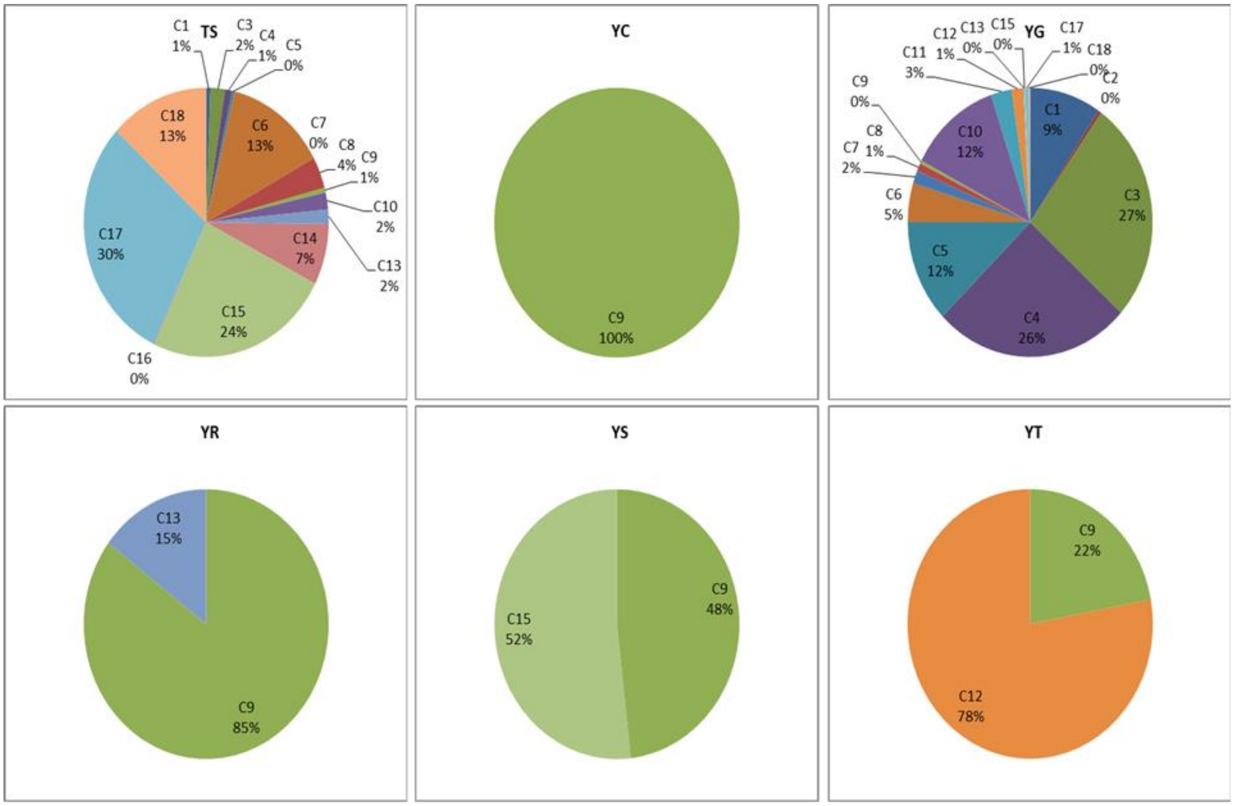


Figure 4.S2 Composition of each TEM site series in terms of the 18 clusters. Rare TEM site series and very small polygons excluded.

4.5 References

Asner, G. P., Mascaro, J., Muller-Landau, H. C., Vieilledent, G., Vaudry, R., Rasamoelina, M., et al. (2012). A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia*, *168*(4), 1147–60. doi:10.1007/s00442-011-2165-z

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. doi: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., 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. doi: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. doi:10.1002/hyp.9609

BC Ministry of Forests and Range and BC Ministry of Environment. (2010). *Field Manual for Describing Terrestiral Ecosystems. 2nd edition. Land Management Handbook 25*. Victoria, BC.

Brinson, M. (1993). Changes in the functioning of wetlands along environmental gradients. *Wetlands*, *13*(2), 65–74. doi:10.1007/BF03160866

Bryan, B. A. (2006). Synergistic techniques for better understanding and classifying the environmental structure of landscapes. *Environmental Management*, *37*(1), 126–140. doi: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., Pizzeghello, D., Concheri, G., Zanella, A., Nardi, S., & Squartini, A. (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. doi: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. doi: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. doi: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. doi:10.1016/S0168-1699(02)00108-4

Coops, N. C., Hilker, T., Wulder, M. A., St-Onge, B., Newnham, G., Siggins, A., & Trofymow, J. A. (2007). Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees*, *21*(3), 295–310. doi: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. doi:10.1016/j.ecoinf.2008.09.005

De Reu, J., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., et al. (2013). Application of the topographic position index to heterogeneous landscapes. *Geomorphology*, *186*, 39–49. doi:10.1016/j.geomorph.2012.12.015

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. doi:10.5589/m02-064

Defries, R., Hansen, M., & Townshend, J. (1995). Global discrimination of land cover types from metrics derived from AVHRR Pathfinder data. *Remote Sensing of Environment*, *54*(3), 209–222.

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. doi:http://dx.doi.org/10.1656/045.021.0301

Dobrowski, S. Z., Safford, H. D., Cheng, Y. Ben, & 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. doi:10.3170/2008-7-18560

Emili, L. (2003). *Hydrochemical characteristics of hypermaritime dorest-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. doi: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). Springer Netherlands. doi:10.1007/978-94-007-5784-4

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. doi:10.1016/j.ecolind.2012.02.024

Franklin, J. (1995). Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography*, *19*(4), 474–499. doi:10.1177/030913339501900403

Franklin, J. (2013). Mapping vegetation from landscape to regional scales. In E. van der Maarel & J. Franklin (Eds.), *Vegetation Ecology* (2nd editio.). 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. doi:10.1016/j.jag.2011.08.013

Gillespie, T., Zutta, B. R., Early, M. K., & Saatchi, S. (2006). Predicting and quantifying the structure of tropical dry forests in South Florida and the Neotropics using spaceborne imagery. *Global Ecology and Biogeography*, *15*(3), 225–236. doi:10.1111/j.1466-822X.2005.00203.x

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. doi: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.

Handcock, 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*, S39–60. doi:10.1007/s00267-003-1084-0

Heinzel, J., & Koch, B. (2011). Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation*, *13*(1), 152–160. doi:10.1016/j.jag.2010.09.010

Hogg, A. R., & Holland, J. (2008). An evaluation of DEMs derived from LiDAR and photogrammetry for wetland mapping. *Forestry Chronicle*, *84*(6), 840–849. doi: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. doi:10.1080/02827580410019472

Holmgren, J., & Persson, Å. (2004). Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, *90*(4), 415–423. doi:10.1016/S0034-4257(03)00140-8

Hopkinson, C., & Chasmer, L. (2009). Testing LiDAR models of fractional cover across multiple forest ecozones. *Remote Sensing of Environment*, *113*(1), 275–288. doi:10.1016/j.rse.2008.09.012

Immitzer, M., Atzberger, C., & Koukal, T. (2012). Tree species classification with Random Forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote Sensing*, *4*(12), 2661–2693. doi:10.3390/rs4092661

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. doi: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. doi:10.1016/j.rse.2010.07.002

Kennel, P., Tramon, M., Barbier, N., & Vincent, G. (2013). Canopy height model characteristics derived from airbone laser scanning and its effectiveness in discriminating various tropical moist forest types. *International Journal of Remote Sensing*, *34*(24), 8917–8935. doi:10.1080/01431161.2013.858846

Kent, M., Gill, W., Weaver, R., & Armitage, R. (1997). Landscape and plant community boundaries in biogeography. *Progress in Physical Geography*, *21*(3), 315–353. doi:10.1177/030913339702100301

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. doi: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. doi: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. doi:10.1111/j.1523-1739.2003.00469.x

Leckie, D., Tinis, S., Nelson, T., Burnett, C., Gougeon, F., Cloney, E., & Paradine, D. (2005). Issues in species classification of trees in old growth conifer stands. *Canadian Journal of Remote Sensing*, *31*(2), 175–190. doi:10.5589/m05-004

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. doi: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. doi:10.1641/0006-3568(2002)052[0019:LRSFES]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. doi:10.1191/0309133303pp360ra

Long, J., Nelson, T. A., & Wulder, M. A. (2010). Regionalization of Landscape Pattern Indices Using Multivariate Cluster Analysis. *Environmental Management*, *46*(1), 134–142. doi: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. doi:10.1007/s00267-003-5181-x

MacKenzie, W., & Moran, J. (2004). *Wetlands of British Columbia: A Guide to Identification*. Victoria, BC: BC Ministry of Forests Research Branch. Land Management Handbook 52.

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. doi: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. doi:10.1016/j.geoderma.2007.04.027

Magnussen, S., & Boudewyn, 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. doi:10.1139/x98-078

Maxa, M., & Bolstad, P. (2009). Mapping northern wetlands with high resolution satellite images and LiDAR. *Wetlands*, *29*(1), 248–260. doi:10.1672/08-91.1

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. doi: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. doi:10.1525/bio.2010.60.1.9

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. doi:10.1016/S0034-4257(01)00290-5

Olson, D. M., Dinerstein, E., Wikramanaya, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., et al. (2001). Terrestrial ecoregions of the world: a new map of life on Earth. *BioScience*, *51*(11), 933–938. doi:http://dx.doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Olstad, T. A. (2012). Understanding the science and art of ecoregionalization. *The Professional Geographer*, *64*(2), 303–308. doi:10.1080/00330124.2011.603656

Ozesmi, S. L., & Bauer, M. E. (2002). Satellite remote sensing of wetlands. *Wetlands Ecology and Management*, *10*, 381–402. doi: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., & Cumming, S. G. (2012). A remote sensing approach to biodiversity assessment and regionalization of the Canadian boreal forest. *Progress in Physical Geography*, *37*(1), 36–62. doi: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*. Ecosystems Working Group, Terrestrial Ecosystems Task Force. Province of British Columbia.

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. doi:10.1002/aqc

Roth, B., Slatton, K., & Cohen, M. (2007). On the potential for high-resolution lidar to improve rainfall interception estimates in forest ecosystems. *Frontiers in Ecology and the Environment*, *5*(8), 421–428. doi:10.1890/060119.01

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).

Running, S., Loveland, T., & Pierce, L. (1994). A vegetation classification logic based on remote sensing for use in global biogeochemical models. *Ambio*, *23*(1), 77–81.

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.

Schmidtlein, S., Tichý, L., Feilhauer, H., & Faude, U. (2010). A brute-force approach to vegetation classification. *Journal of Vegetation Science*, *21*(6), 1162–1171. doi:10.1111/j.1654-1103.2010.01221.x

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. doi: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. doi: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. doi:10.5589/m12-014

Van Leeuwen, M., & Nieuwenhuis, M. (2010). Retrieval of forest structural parameters using LiDAR remote sensing. *European Journal of Forest Research*, *129*(4), 749–770. doi:10.1007/s10342-010-0381-4

Varhola, A., & Coops, N. C. (2013). Estimation of watershed-level distributed forest structure metrics relevant to hydrologic modeling using LiDAR and Landsat. *Journal of Hydrology*, *487*, 70–86. doi:10.1016/j.jhydrol.2013.02.032

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. doi:10.1890/070001

Wang, T., Hamann, A., Spittlehouse, D. L., & Murdock, T. (2012). ClimateWNA - High-resolution spatial climate data for western North America. *Journal of Applied Meterology and Climatology*, *51*, 16–29. doi:http://dx.doi.org/10.1175/JAMC-D-11-043.1

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. doi:10.1016/j.ecolind.2015.01.032

Whittaker, R. (1967). Gradient Analysis of Vegetation. *Biological Reviews*, *49*, 207–7684.

Wulder, M. A., Bater, C. W., Coops, N. C., Hilker, T., & White, J. C. (2008). The role of LiDAR in sustainable forest management. *The Forestry Chronicle*, *84*(6), 807–826.

Wulder, M. A., Hall, R. J., Coops, N. C., & Franklin, S. E. (2004). High spatial resolution remotely sensed data for ecosystem characterization. *BioScience*, *54*(6), 511–521. doi:10.1641/0006-3568(2004)054[0511:HSRRSD]2.0.CO;2

Wulder, M. A., White, J. C., Luther, J. E., Strickland, G., Remmel, 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. doi:10.5589/m06-023

Zedler, J. B., & Kercher, S. (2005). Wetland resources: status, trends, ecosystem services, and restorability. *Annual Review of Environment and Resources*, *30*(1), 39–74. doi:10.1146/annurev.energy.30.050504.144248

5. Spatial patterns of global primary productivity regimes from 2000 to 2012

Abstract

Primary productivity and vegetation phenology are spatially and temporally variable ecosystem functions. Understanding spatial-temporal patterns can improve global change science and management. Researchers typically focus on temporal changes within or among static regions and omit dynamics of spatial configuration. Our goal was to assess global spatial-temporal variability in productivity and phenology between 2000 and 2012 using a temporally dynamic functional type classification. Fourteen functional types were defined for each year by clustering the annual sum and annual variability of the Fraction of Photosynthetically Active Radiation (fPAR) from the Moderate Resolution Imaging Spectrometer (MODIS). The 14 functional types ranged from tundra (low cumulative fPAR and highly seasonal), to tropical forests (high cumulative fPAR and low seasonality). Variability in the mean of the fPAR metrics and in two spatial pattern metrics were assessed for each functional type. Many pixels changed from one cluster to another then back again, indicating considerable short-term variability. Temporal variability in the mean of the fPAR metrics was relatively low, with changes instead primarily manifested in spatial pattern. Spatial pattern was most variable in tundra, grasslands, shrublands, and savannas. A dynamic classification demonstrated the variability in spatial patterns of primary productivity and can be used for future monitoring.

5.1 Introduction

Human well-being depends on ecosystem services such as the provision of natural resources, climate regulation, water purification, and cultural and spiritual benefits (Millennium Ecosystem Assessment 2005). Net Primary Productivity (NPP), the rate of photosynthetic production per unit area, is a basic ecosystem function that supports all other ecosystem services (Andrew et al. 2015; Carpenter et al. 2009; Kremen 2005). For instance, the extraction of timber resources and the production of oxygen through photosynthesis are both tied to primary productivity. Although the exact nature of the relationship is contested, primary productivity is also one of the most widely recognized drivers of global patterns in biodiversity (Field et al.

2008; Pianka 1966; Whittaker 2010; Wright 1983). As NPP represents the energy available to higher trophic levels (Berry et al. 2007), greater productivity is thought to directly increase the number of individuals, and species that an area can support (Currie et al. 2004; Evans et al. 2005). Vegetation phenology, the timing of developmental stages of plant life such as green-up, flowering, senescence, and length of growing season, is also a key functional trait of ecosystems that influences ecosystem processes such as pollination, carbon, water and nutrient cycling (Morisette et al. 2009; Richardson et al. 2013).

Both productivity and phenology exhibit considerable variability across space and time. Continental and global gradients in productivity and phenology reflect spatial variation in climate, and especially temperature and precipitation (Kimball 2014; Polis 1999; Richardson et al. 2013). Landscape and regional variability in these processes may also reflect finer-scale abiotic processes, and the response of species or functional groups to these patterns (Baeza et al. 2010; Diez et al. 2012; Hansen et al. 2000; Schöngart et al. 2002). The productivity and phenology of a given region or ecosystem may fluctuate as a result of inter-annual climatic variability (Mohamed et al. 2004; Piao et al. 2006; Zeng et al. 2013). As the global climate warms (IPCC 2014) persistent changes in primary productivity and vegetation phenology have also been identified. Specifically, in many areas an increase in NPP and growing season length have been observed over the long term, concomitant with increasing global mean temperature over the same time period (Boisvenue and Running 2006; Buitenwerf et al. 2015; Keenan et al. 2014; Sitch et al. 2015; Xu et al. 2013).

Monitoring spatial and temporal variability of ecosystem functions, such as productivity and phenology, with remote sensing can be used to assess changes in the delivery of ecosystem services (Andrew et al. 2014; Volante et al. 2012) and biodiversity (Coops et al. 2008). Spatial-temporal trends in productivity and phenology can also inform management and conservation (Cabello et al. 2012; Crabtree et al. 2009) and climate change science (Bonan et al. 2002; Botta et al. 2000; Holmes et al. 2015; Nelson et al. 2014). At global scales, quantitative, highly repeatable, and spatially comprehensive information regarding above-ground NPP and phenology can *only* be derived from remotely sensed data that are spatially and temporally aggregated and are representative of whole ecosystems or biomes (Alcaraz-Segura et al. 2013; Boisvenue and Running 2006; Morisette et al. 2009; Ustin and Gamon 2010; White and Nemani 2006). Such data have been used in numerous studies to assess recent global changes in productivity (e.g., de Jong et al., 2013; Jong et al., 2012; Mao et al., 2013; Zhao and Running, 2010) and phenology (e.g., Eastman et al., 2013; Jeong et al., 2011).

These global studies report changes over time in the magnitude of productivity (e.g., mean annual greenness), or the timing of phenological events (e.g., mean growing season length), and may capture spatial variation in these changes along latitudinal gradients or among static regions such as biomes, countries, or continents. The use of static regions to assess changes is limiting because ecological boundaries are dynamic and respond to intra and inter annual changes in climate. The spatial configuration of vegetation can also influence ecosystem services related to the flow of water, soil, nutrients and organisms such as pollination and carbon storage (Mitchell et al. 2013). For instance, flowering and fruiting in fragmented landscapes may be earlier and more productive at edges due to altered abiotic conditions (Herrerías-Diego et al. 2006). Such changes in phenology can affect plant reproduction and pollination (Burgess et al. 2006). As such, analyses that integrate the dynamic nature of ecosystems may highlight important spatial-temporal variation and change. For instance, as ecosystem conditions change, neighbouring units will become more or less similar but this variability is not reflected under a system of static boundaries (Handcock and Csillag 2002). A regionalization methodology that is dynamic allows the possibility of detecting spatial shifts in environmental conditions (Reygondeau et al. 2013). We are unaware of any studies that have assessed the dynamics of spatial configuration indicators of productivity or phenology at a global scale.

The goal of this research was to assess spatial-temporal variability in vegetation productivity and phenology across the globe between 2000 and 2012, at three-year intervals, using a classification based on *dynamic* ecosystem functional types. Ecosystem functional types provide a meaningful framework for understanding, monitoring, assessing, and predicting changes to ecosystems and ecosystem services (Ivits et al. 2013; Paruelo et al. 2001). To meet our goal, we used cluster analysis to classify and map remotely sensed proxies for productivity and phenology into functional types for five years from 2000 and 2012. Previous researchers have used remotely sensed proxies of primary productivity and phenology to delineate ecosystem functional types at regional (Alcaraz-Segura et al. 2013; Paruelo et al. 2001) and global scales (Ivits et al. 2013). Our analysis is unique in that we allow mapped boundaries of the functional types to change through time, rather than using a static classification, enabling the detection of changes in spatial pattern.

5.2 Methods

5.2.1 Data

Classification of ecosystem functional types in this study was based on remotely sensed estimates of the fraction of photosynthetically active radiation (fPAR) absorbed by vegetation. Whereas vegetation indices such as the Normalized Difference Vegetation Index (NDVI) are composite measures of leaf chlorophyll content, leaf area, canopy cover, and structure, fPAR is an intrinsic biophysical variable that directly influences primary productivity (Glenn et al. 2008;

Huete et al. 2011; Myneni et al. 2002). fPAR depends on vegetation type and structure and is commonly used in the estimation of ecosystem productivity and other biogeochemical cycles (Myneni et al. 2002). Values of fPAR range from 0 to 1; larger values indicate high vegetation greenness and productivity, while values approaching zero are indicative of more sparsely vegetated areas with low productivity (Myneni et al. 2002). For this study, fPAR data were obtained from the earth observing satellite Moderate Resolution Imaging Spectrometer (MODIS) at a spatial resolution of 0.05 degrees (approximately 5.5 km at the equator). As MODIS data are collected daily, cloud screening algorithms can be applied to produce cloud-free composites by selecting the clearest pixels within a range of dates. In our case, we used MODIS eight-day composites, totalling 46 images per year. The fPAR data were compiled for the years 2000, 2003, 2006, 2009 and 2012 to capture the range of variability within this time period.

To classify ecosystem functional types, two fPAR metrics were calculated for each of the five years of interest: total annual sum ($fPAR_{sum}$), and the annual coefficient of variation ($fPAR_{cv}$). The total annual sum is representative of annual net primary productivity (Goward et al. 1985). Higher values are indicative of areas that are productive throughout the year and/or that are highly productive for the growing season. As the maximum fPAR value for any pixel is 1, and the number of images utilized per year was 46, $fPAR_{sum}$ in theory ranges from 0 to 46. Excluding non-vegetated areas for which MODIS fPAR data are absent, our calculated values of $fPAR_{sum}$ ranged from 0.1 to 42 for the years assessed. The annual coefficient of variation of fPAR ($fPAR_{cv}$) quantifies the seasonality of primary productivity, and represents phenological variation. High variability in fPAR (or NDVI) occurs where the climate exhibits extremes within a year and where vegetation is highly temperature or precipitation limited, such as in annual and desert grasslands, arctic and sub-arctic environments, or where there are cyclical agricultural practices in place (Coops et al. 2008; Coops et al. 2009; Reed et al. 1994). Values of $fPAR_{cv}$ in our dataset ranged from 0 to 6.7.

5.2.2 Classification and mapping of functional types

A quantitative regionalization approach was used to classify and map functional types. The two fPAR metrics, $fPAR_{sum}$ and $fPAR_{cv}$ from each of the five target years, were used in a cluster analysis to generate a map of global ecosystem functional types. Cluster analysis is an unsupervised classification technique that partitions unlabelled data into natural groups based on (dis)similarity (Jain 2010; Jain et al. 1999).We applied a non-spatial clustering that did not enforce spatial contiguity in the classification. We used two-step clustering in the statistical software SPSS (v22) because it can accommodate very large datasets. In two-step clustering, raw data points are first partitioned into a set of pre-clusters. Then, a hierarchical agglomerative clustering method consecutively joins the pre-clusters into a smaller number of clusters based on their distance in feature space (Mooi and Sarstedt 2011). All five years of the fPAR measures were incorporated into one data frame, standardized to z-scores, then clustered.

Previous studies have mapped 10 to 15 global ecosystem functional types (Bonan et al. 2002; Friedl et al. 2010; Ivits et al. 2013; Poulter et al. 2011). To maximize thematic resolution yet enable general comparisons with these and other global mapping products (e.g., Olson et al. 2001), we specified that 14 clusters should be delineated. Non-vegetated pixels (fPAR = 0) were excluded from the classification. After running the cluster analysis, cluster membership values were mapped for each of the five years (2000, 2003, 2006, 2009, and 2012). An overall map representing all years was also produced by calculating the majority (modal) cluster value of each pixel. The maps were smoothed using a 3x3 majority filter. Clusters were interpreted by assessing the statistical distribution (e.g., mean, median, range) of each cluster with respect to

$fPAR_{sum}$ and $fPAR_{cv}$. To provide additional context, we also determined the dominant land cover type (Friedl et al. 2010) and biome type (Olson et al. 2001) within each of the resultant clusters.

5.2.3 Spatial–temporal variability and change

To assess temporal variation in the clustering, change matrices were built for each of four change periods: 2000 to 2003, 2003 to 2006, 2006 to 2009, and 2009 to 2012. Each matrix contained the number of pixels classified as each cluster type for each date, thus providing information on the type and magnitude of change. We mapped the results for each year, and summed the results of the four individual matrices to generate one table representing total change over all years. In order to determine whether changes were sustained and directional or short-term and variable, we also assessed the frequency of change for each individual pixel changed (zero to four times). For those pixels that changed more than two times, we then assessed the number of unique values. Pixels with four unique cluster classifications represent a very different type of change than pixels that changed four times, but had only two unique values during that time.

Several indicators of spatial and temporal variability in the ecosystem functional types were assessed. First, for each of the five years of data, mean annual greenness ($fPAR_{sum}$) and mean annual seasonality ($fPAR_{cv}$) were calculated for each cluster. Each cluster was then described in terms of its long-term variability (coefficient of variation and standard deviation, respectively) in each of these two metrics. The novelty of our dynamic regionalization based on global fPAR measures is that it allows for an assessment of the change in spatial pattern over time. Two components of spatial pattern in the clusters were assessed: *composition* and *configuration*. Composition refers to the area of a cluster, indicating what and how much is present (Fahrig 2005; Gustafson 1998). We calculated the area of each cluster at each time step by summing the total number of pixels in each. Note that the sum of pixels in each cluster is also contained in the change matrices previously described; the key difference is that to measure variability in spatial pattern, we then computed the coefficient of variation of cluster area over time.

Configuration refers to explicit spatial arrangements such aggregation, fragmentation, and adjacency (Fahrig 2005; Gustafson 1998). To assess spatial configuration, we calculated a measure of aggregation using FRAGSTATS (McGarigal and Ene 2013), a tool designed to calculate landscape metrics from categorical maps. We chose a single measure of spatial configuration, *Clumpiness*, because it is bounded (ranges from -1 to +1), thus facilitating interpretability (Crews-Meyer 2004), and responds linearly to changes in aggregation regardless of class area (Neel et al. 2004). Values of clumpiness approaching -1 occur in maximally disaggregated situations, while values approaching +1 occur in maximally clumped situations. We emphasize that the spatial resolution and total landscape extent remained constant at each time step, thus avoiding challenges with metric comparison (e.g., Wu, 2004; Wu et al., 2002). Clumpiness was calculated for each cluster for each of the five years (2000, 2003, 2006, 2009, 2012), and the coefficient of variation computed.

5.3 Results

5.3.1 Classification and mapping of functional types

The 14 derived clusters (functional types) correspond logically to existing global biome and land cover classifications (Table 5.1), and each functional type had a unique (and quantitative) distribution with regards to cumulative annual greenness and annual seasonality of greenness. Clusters 1 through 4 were generally found at high elevations and/or high latitudes, with relatively low annual greenness and relatively high seasonality (Figures 5.1 and 5.2).

Clusters 1 and 2 were tundra functional types and were the least green (mean $fPAR_{sum}$ values of 6.3 and 7.9, respectively) and most seasonal (mean $fPAR_{cv}$ values of 1.4 and 1.2, respectively) of all clusters (Figure 5.1, Table 5.1). Cluster 3 was similar to Clusters 1 and 2 in terms of annual greenness (mean $fPAR_{sum} = 7.8$), but was less seasonal (mean $fPAR_{cv} = 0.75$). Cluster 3 was grass-dominated, and lay primarily within the Eurasian Steppe and the Great Plains region of North America. Cluster 4 was an open, shrubby functional type with a slightly higher annual greenness than the first three (mean $fPAR_{sum} = 11.4$) and its degree of seasonality was intermediate between Clusters 2 and 3 (mean $fPAR_{cv} = 0.98$). It was found in Scandinavia, Siberia, Alaska and northern Canada, as well as the Aspen Parkland region in central Canada.

Cluster 5 was grass or shrub-dominated, and had low annual greenness (mean $fPAR_{sum} = 6.2$) similar to Clusters 1 through 3, but was much less seasonal than those clusters (mean $fPAR_{cv} = 0.45$). It occurred at lower latitudes, primarily in arid regions of Eurasia, central Australia, northern Africa, southwestern USA, and southern Chile. Clusters 6 and 7 were also grass and shrub dominated, but with a slightly higher annual greenness than Cluster 5 (mean $fPAR_{sum} = 9.4$ and 12.1, respectively). They occurred in semi-arid regions such as Southern Africa, Australia, India, and the western United States. Clusters 8 and 9 were dominated by forests, with moderately-high annual greenness (mean $fPAR_{sum} = 18.5$ and 15.1, respectively), and moderately-high annual seasonality (mean $fPAR_{cv} = 0.58$ and 0.80, respectively). They occurred primarily at moderately-high latitudes across large areas of Europe, Eurasia, and North America within the *boreal forests / taiga* biome.

Cluster 10 through 14 were characterized by high annual greenness and low seasonality. Cluster 10 was dominated by savannas and scattered across temperate, subtropical and especially tropical regions of the globe (Table 5.1, Figure 5.2). It had a mean annual greenness of 25.7 and a mean annual seasonality of 0.22. Cluster 11 was slightly less productive (mean $fPAR_{sum}$ = 22.6) and slightly more seasonal than Cluster 10 (mean $fPAR_{cv}$ = 0.42). It occurred primarily within temperate broadleaf and mixed forests in central Europe and North America. Cluster 12 was dominated by savannas with a mean annual greenness value of 18.8, and a mean $fPAR_{cv}$ of 0.24. It occurred in regions such as northeastern Australia, parts of southern and central Africa, and the Indian subcontinent. Clusters 13 and 14 occurred in the tropical Amazon and Congo River basins, and had very low seasonality (mean $fPAR_{cv}$ = 0.05 and 0.12, respectively), and very high annual greenness (mean $fPAR_{sum}$ = 38.0 and 31.1, respectively). Cluster 13 was predominately tropical forest, while Cluster 14 (substantiated by a visual assessment of higher spatial resolution satellite imagery on GoogleEarth) was more open, representing pastures, croplands, or regenerating forests.

Table 5.1. Relationship of clusters with existing land cover and biome types and our functional interpretation.

Cluster	Dominant biome[1] (% of Cluster)	2nd Dominant biome[1] (% of Cluster)	Dominant land cover type[2] (% of Cluster)	2nd Dominant land cover type[2] (% of Cluster)	Interpreted functional type
1	Tundra (67)	Boreal forests / taiga (32)	Open shrublands (80)	Grasslands (15)	Tundra (Low Productivity)
2	Boreal forests / taiga (51)	Tundra (42)	Open shrublands (72)	Grasslands (14)	Tundra (Low Productivity)
3	Temperate grasslands, savannas, & shrublands (44)	Deserts & xeric shrublands (21)	Grasslands (77)	Croplands (10)	Grasslands (Low Productivity)
4	Boreal forests / taiga (56)	Temperate grasslands, savannas, & shrublands (17), Tundra (17)	Open shrublands (31)	Woody savannas (17)	Forests & shrublands (Low Productivity), and Grasslands (Moderate Productivity)
5	Deserts & Xeric shrublands (59)	Tropical & subtropical grasslands, savannas, & shrublands (13)	Open shrublands (65)	Grasslands (19)	Grasslands (Low Productivity)
6	Deserts & xeric shrublands (26)	Tropical & subtropical grasslands, savannas, & shrublands (23)	Grasslands (59)	Open shrublands (17)	Grasslands & shrublands (Moderate Productivity)
7	Deserts & xeric shrublands (34)	Tropical & subtropical grasslands, savannas, & shrublands (22)	Open shrublands (55)	Grasslands (23)	Grasslands & shrublands (Moderate Productivity)
8	Boreal forests / taiga (42)	Temperate broadleaf & mixed forests (29)	Mixed forest (34)	Croplands (20)	Forests (Moderate Productivity)
9	Boreal forests / taiga (58)	Temperate broadleaf & mixed forests (18)	Woody savannas (19)	Croplands (18)	Forests (Moderate Productivity)
10	Tropical & subtropical grasslands, savannas, & shrublands (37)	Temperate broadleaf & mixed forests (23)	Woody savannas (30)	Savannas (17)	Forests & savannas (High Productivity)
11	Temperate broadleaf & mixed forests (36)	Boreal forests / taiga (24)	Mixed forest (40)	Savannas (15)	Forests (High Productivity)

12	Tropical & subtropical grasslands, savannas, & shrublands (38)	Mediterranean forests, woodlands & scrub (15)	Savannas (29)	Croplands (22)	Grasslands & savannas (High productivity)
13	Tropical & subtropical moist broadleaf forests (85)	Temperate broadleaf & mixed forests (4)	Evergreen broadleaf forest (86)	Cropland / Natural vegetation mosaic (5)	Forests (High Productivity)
14	Tropical & subtropical moist broadleaf forests (18)	Temperate broadleaf & mixed forests (19), Tropical & subtropical grasslands, savannas, & shrublands (19)	Woody savannas (22)	Evergreen broadleaf forest (18) & Mixed forest (18)	Forests & shrublands (High productivity)

---

[1] (Olson et al. 2001)

[2] (Friedl et al. 2010)

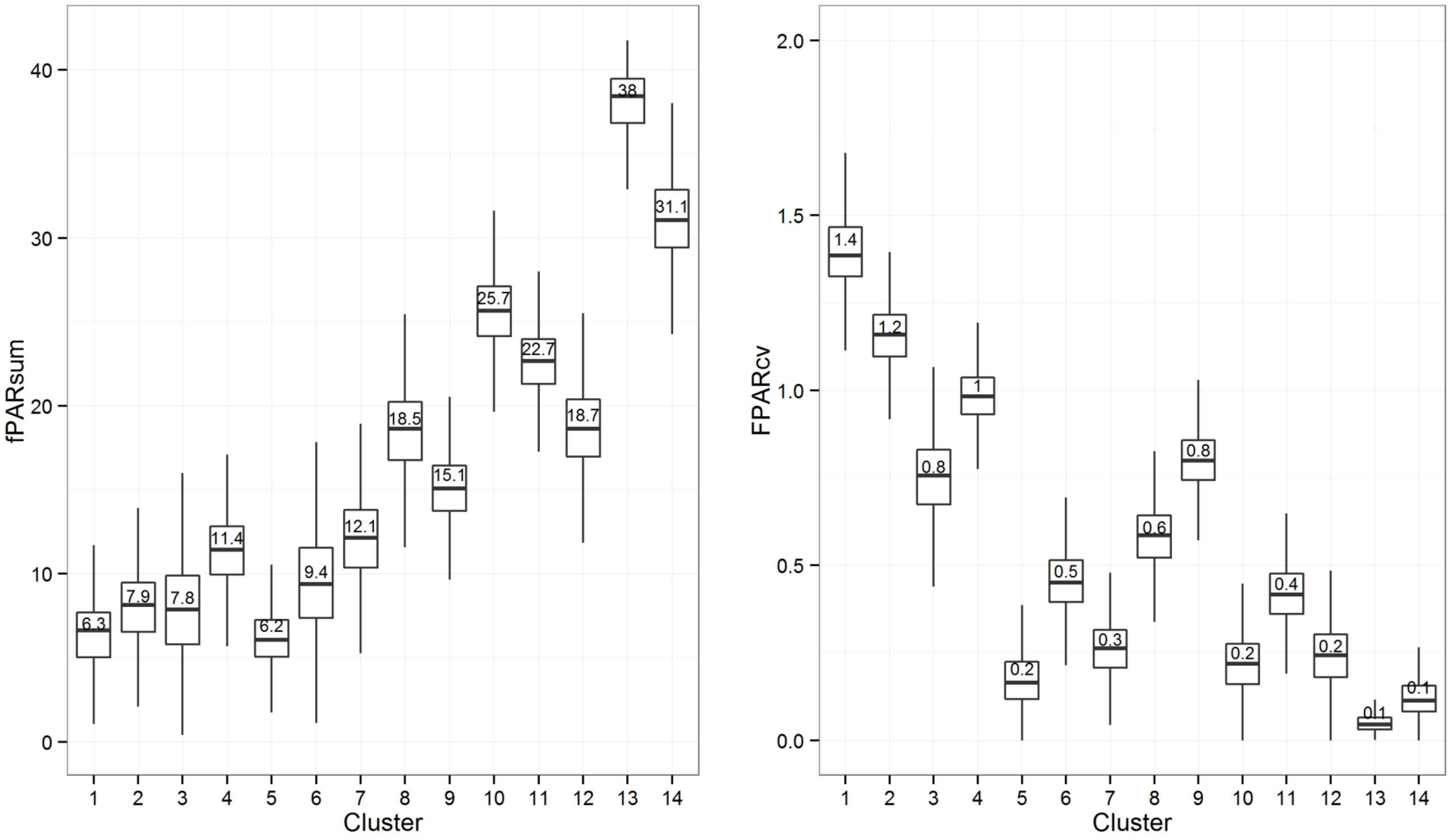


Figure 5.1. Distribution of total annual productivity ($fPAR_{sum}$) and seasonality in productivity ($fPAR_{cv}$) for the 14 clusters, with mean values labelled. Boxplots were produced with values for all years combined.

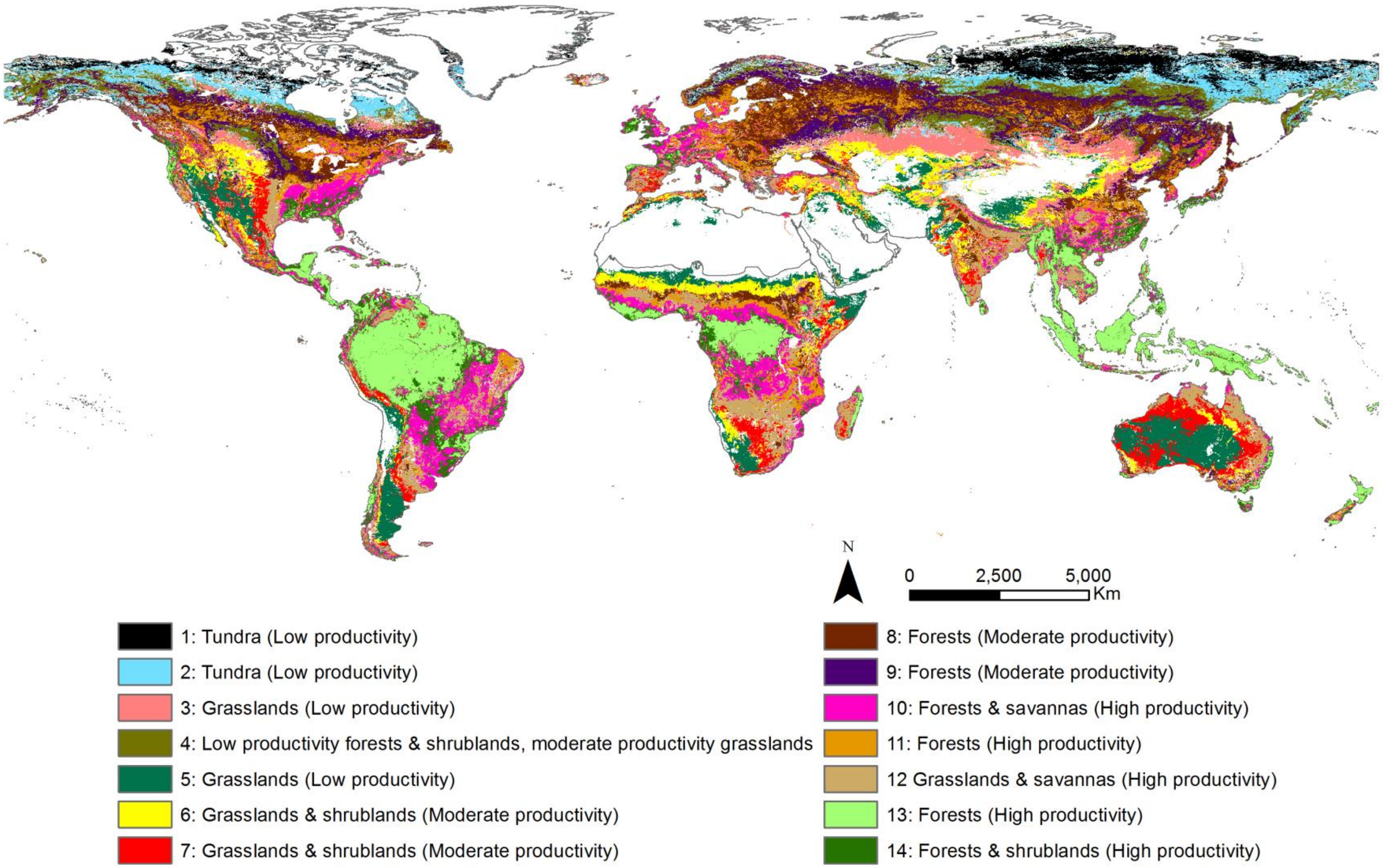


Figure 5.2. Fourteen global ecosystem functional types were delineated that represent unique combinations of annual greenness (fPAR_{sum}) and annual seasonality (fPAR_{cv}). Depicted is the majority (modal) value between 2000 and 2012.

5.3.2 Spatial–temporal variability and change

An assessment of change frequency and uniqueness indicated that the majority (50%) of pixels with a high frequency of change had only two unique values through time, and 44% had three unique values over time. The cumulative change matrix (Table 5.2) indicates that most changes represented pixels changing from one functional type to another functional type that was relatively similar (e.g., Figure 5.1). For instance, pixels that changed from Cluster 1 (low productivity tundra) were primarily reclassified as Cluster 2 (also tundra but with slightly higher annual greenness and slightly lower seasonality), while change in Cluster 2 was primarily towards Cluster 1 (Table 5.3). Pixels from Clusters 3 and 5 (both representing low productivity grasslands) primarily changed to Cluster 6 (moderately productive grasslands and shrublands). Change in Cluster 6 was dominantly to Cluster 3, and pixels that changed over time in Cluster 7 (also moderately productive grasslands and shrublands) mainly became part of Cluster 5. Pixels that changed from Cluster 4 (forests, shrublands and grasslands with low to moderate productivity) were for the most-part reclassified as Cluster 2. Pixels that changed in Cluster 8 (moderately productive forests) were mostly reclassified as Cluster 11 (highly productive forests), while the dominant change for pixels in Cluster 9 (moderately productive forests) was to Cluster 8. Cluster 10 (highly productive forests and savannas) changes were primarily to Cluster 11, and Cluster 11 changes were primarily to Cluster 8. Changes in Cluster 12 (highly productive grasslands and shrublands) were predominately to Cluster 7. Any pixels from Cluster 13 (highly productive forests) that changed over time were primarily reclassified as Cluster 14 (highly productive forests and shrublands), while pixels changing from Cluster 14 were mostly reclassified as Cluster 9.

Table 5.2. Change in cluster area over time, calculated from change matrices (e.g., Table 5.3) for each year (not shown) by differencing row sums and column sums for each cluster. Change greater than |10%| is bolded.

Cluster	**Percent change (2000-2003)[1]**	**Percent change (2003-2006)[2]**	**Percent change (2006-2009)[3]**	**Percent change (2009-2012)[4]**
1	**-18.76**	**23.12**	**-11.38**	0.76
2	-5.78	2.24	**13.95**	**-10.77**
3	1.52	1.85	-3.45	**15.84**
4	**12.05**	**-12.56**	4.53	-7.67
5	5.59	**-11.95**	**20.70**	**-13.85**
6	1.18	**12.81**	-6.81	-6.31
7	**-14.97**	8.34	**-14.36**	**20.86**
8	**10.92**	-6.97	2.52	1.42
9	5.47	-4.42	1.74	5.04
10	0.50	2.09	-5.09	-0.84
11	5.99	-6.25	2.88	-2.48
12	**-13.31**	**12.40**	-4.85	**12.86**
13	0.98	-1.23	-0.34	-1.47
14	1.80	0.06	-5.71	2.96

[1] Relative to cluster area in year 2000

[2] Relative to cluster area in year 2003

[3] Relative to cluster area in year 2006

[4] Relative to cluster area in year 2009

Table 5.3. Change matrix representing pixel counts changing from one cluster to another, summed over all four time periods (2000-2003, 2003-2006, 2006-2009, and 2009-2012). Values along the diagonal represent pixels that did *not* change. The largest change count in each row is bolded (e.g., greatest change in Cluster 1 was to Cluster 2).

								TO								
		Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11	Cluster 12	Cluster 13	Cluster 14	sum
	Cluster 1	576,128	**384,910**	36,237	110,572	203	384	3	85	7,347	2		7			1,115,878
	Cluster 2	**344,929**	675,032	84,772	310,306	304	1,442	12	406	24,984	0	1	7			1,442,195
	Cluster 3	37,700	76,980	574,140	98,448	16,128	**145,867**	2,227	7,697	40,365	4	48	268		1	999,873
	Cluster 4	116,963	**273,033**	117,627	837,668	26	4,861	32	8,259	250,060	3	101	78		0	1,608,711
F	Cluster 5	295	252	16,961	139	1,134,045	**177,082**	129,552	447	142	4	30	916		1	1,459,866
R	Cluster 6	407	995	**161,926**	5,373	148,699	702,726	118,482	49,948	20,816	54	693	16,692		2	1,226,813
O	Cluster 7	4	5	3,497	93	**146,383**	118,970	598,824	23,952	903	799	3,212	130,539	1	43	1,027,225
M	Cluster 8	118	470	8,012	9,396	746	42,062	22,223	1,179,932	223,332	4,340	**245,845**	67,927		40	1,804,443
	Cluster 9	7,126	23,123	34,579	215,288	253	16,369	801	**244,040**	885,368	24	6,854	943			1,434,768
	Cluster 10		2	3	0	11	112	1,676	5,853	69	1,235,891	**187,485**	124,781	1,511	147,159	1,704,553
	Cluster 11		7	215	165	23	1,262	6,207	**249,155**	7,104	173,033	876,252	88,649	27	4,706	1,406,805
	Cluster 12	4	8	547	52	1,203	14,678	**133,718**	65,969	1,386	114,863	80,367	1,044,009	76	3,263	1,460,143
	Cluster 13							8			2,011	34	149	1,739,863	**93,233**	1,835,298
	Cluster 14			1		4	7	60	66	3	**158,934**	4,818	3,118	84,368	802,738	1,054,117
	sum	1,083,674	1,434,817	1,038,517	1,587,500	1,448,028	1,225,822	1,013,825	1,835,809	1,461,879	1,689,962	1,405,740	1,478,083	1,825,846	1,051,186	19,580,688

Change matrices were generated for each of the four change periods and used to calculate the net difference in cluster extent (i.e., total magnitude of change) for each cluster at each time step. From 2000 to 2003, Cluster 1 changed the most of all clusters, decreasing ~19% in area (Figures 5.3 and 5.4, Table 5.3). Cluster 4 (forests, shrublands and grasslands with low to moderate productivity), Cluster 7, Cluster 8 (moderately productive forests) and Cluster 12 also changed by at least ±10%. Conversely, the change in area from 2000 to 2003 in Cluster 10 (highly productive forest and savanna) and Cluster 13 represented less than 1% of their area in 2000. In the next time step (from 2003 to 2006), Cluster 1 again changed the most, this time gaining area equivalent to ~23% of its area in 2003. Clusters 4, 5, and 6 (all low to moderately productive forests, grasslands, or shrublands), and Cluster 12 changed by ~12% -13%. Cluster 14 (highly productive forests and shrublands) changed the least, increasing in area by only 0.06%. From 2006 to 2009, Cluster 5 changed the most, increasing in area by ~21%; Cluster 13 changed the least, losing approx. 0.3% of its area. Finally, in the last time step (from 2009 to 2012), Cluster 7 changed the most, increasing in area by ~21%, while the area of Cluster 1 changed by only ~0.8%.

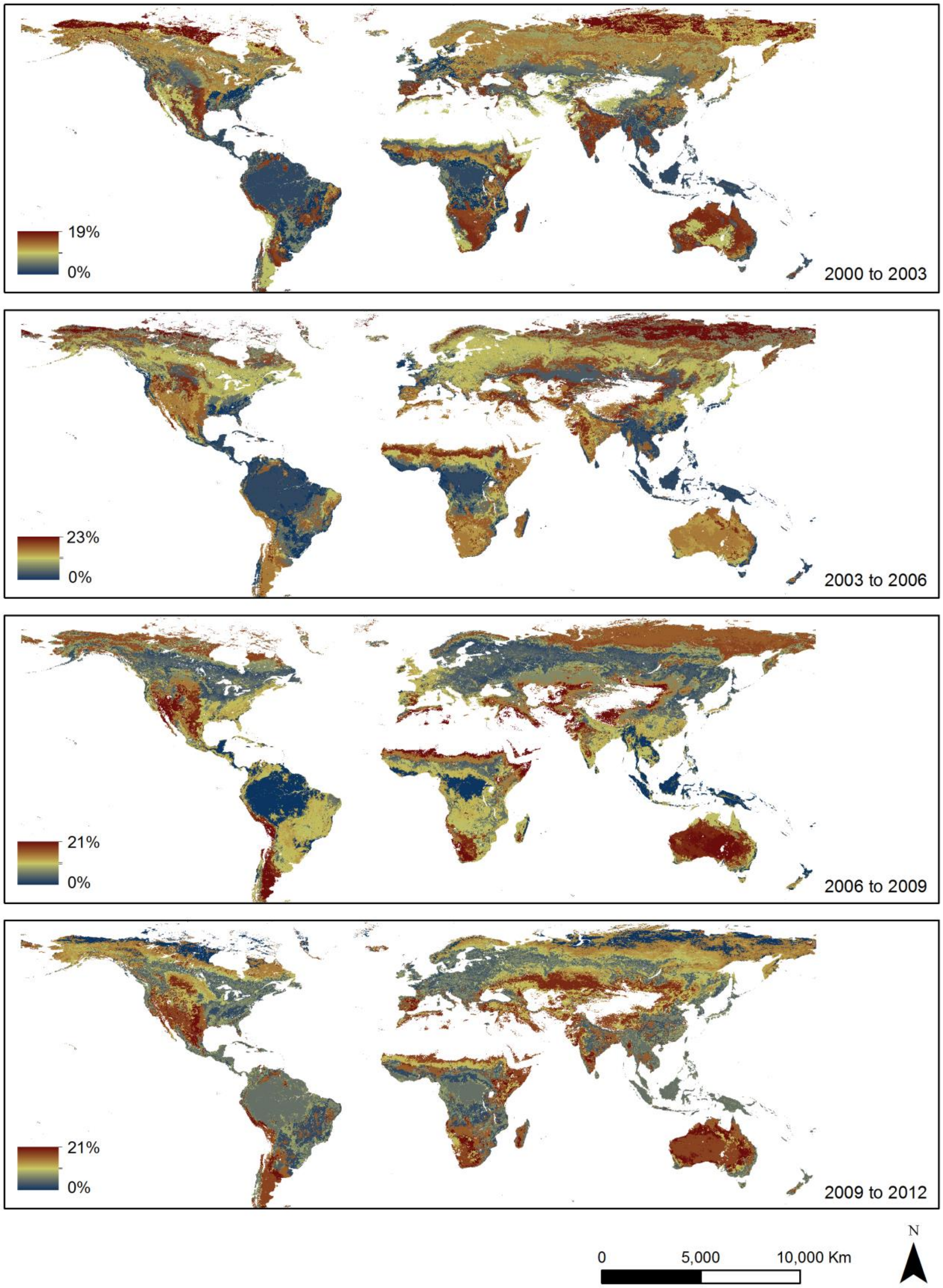


Figure 5.3. Change per ecosystem functional type (cluster) at each step in time, corresponding to Table 5.3.

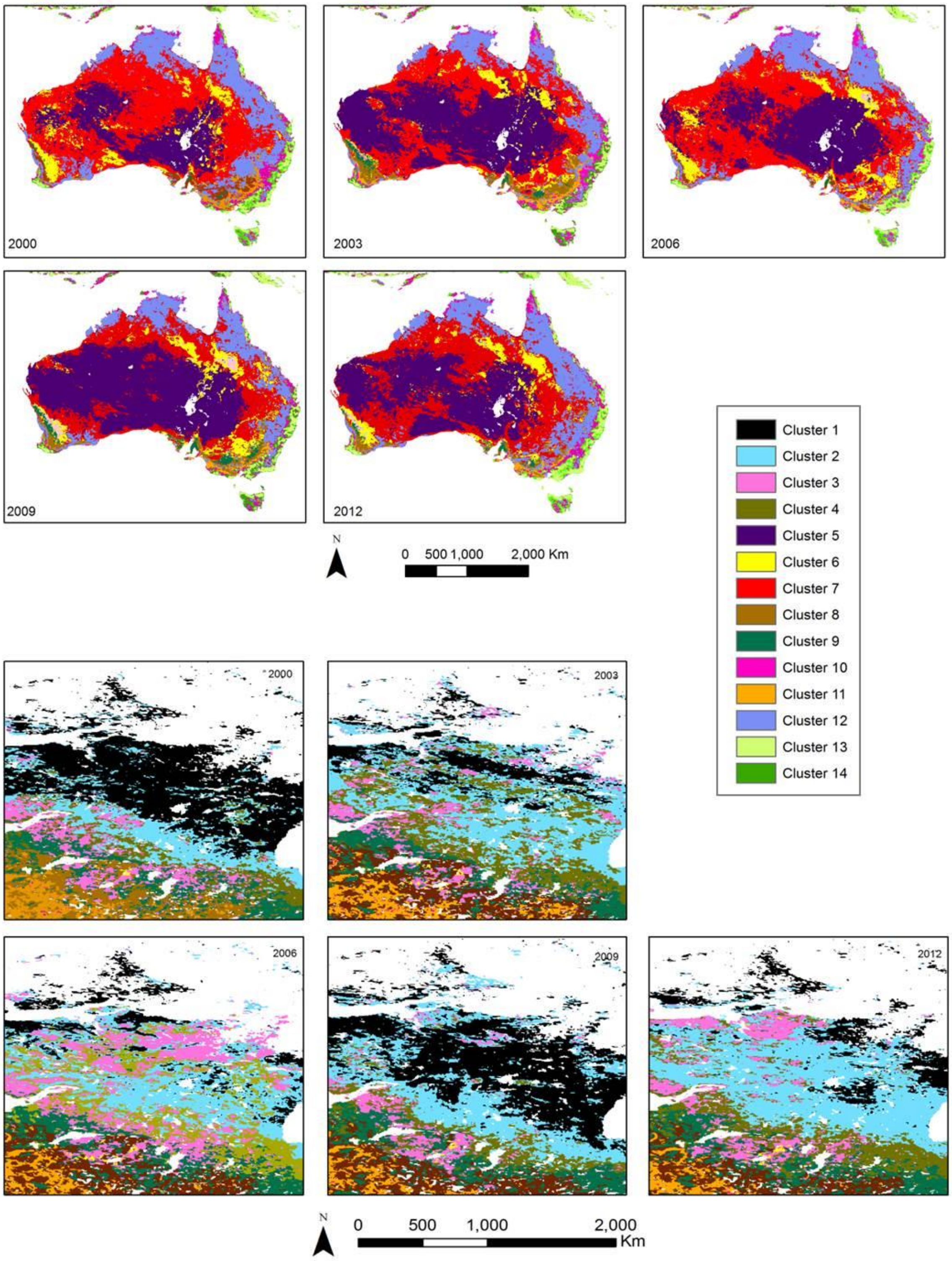


Figure 5.4. Close up of changes occuring in particularly dynamic clusters and regions from 2000 to 2012. The top panel shows the back and forth nature of Clusters 5 and 7 in Australia, while the bottom panel shows similar dynamics between Clusters 1 and 2 in Arctic Canada.

The mean values of $\mathrm{fPAR}_{\mathrm{sum}}$ and $\mathrm{fPAR}_{\mathrm{cv}}$ for each cluster were somewhat variable over time (Figure 5.5). Annual greenness ($\mathrm{fPAR}_{\mathrm{sum)}}$ was highly variable in Cluster 1 (CV = 0.043), moderately high in Cluster 5 (CV = 0.018), and less variable in the other clusters (CV <0.010). Annual seasonality ($\mathrm{fPAR}_{\mathrm{cv}}$) was most variable in Cluster 1 but was quite stable in all other clusters. The spatial pattern, and in particular composition (area) of the 14 clusters was more varible through time (Figure 5.5). For all clusters, variation in composition (area) was greater than variation in configuration (clumpiness). Variability in area was particularly high in Cluster 1 (tundra), Cluster 5 (low productivity grasslands), Cluster 7 (moderately productivity grasslands and shrublands), and Cluster 12 (highly productive grasslands, savannas, and shrublands) (CV > 0.06). Variability in cluster area was much lower in the highly productive forests represented by Cluster 13 (CV = 0.01). The most variable (CV ≥ 0.014) clusters in terms of spatial configuration (clumpiness) were Clusters 1 and 2 (tundra), Cluster 3 (low productivity grasslands), and Cluster 7, while Cluster 13 had the lowest variation in this indicator (CV = 0.001).

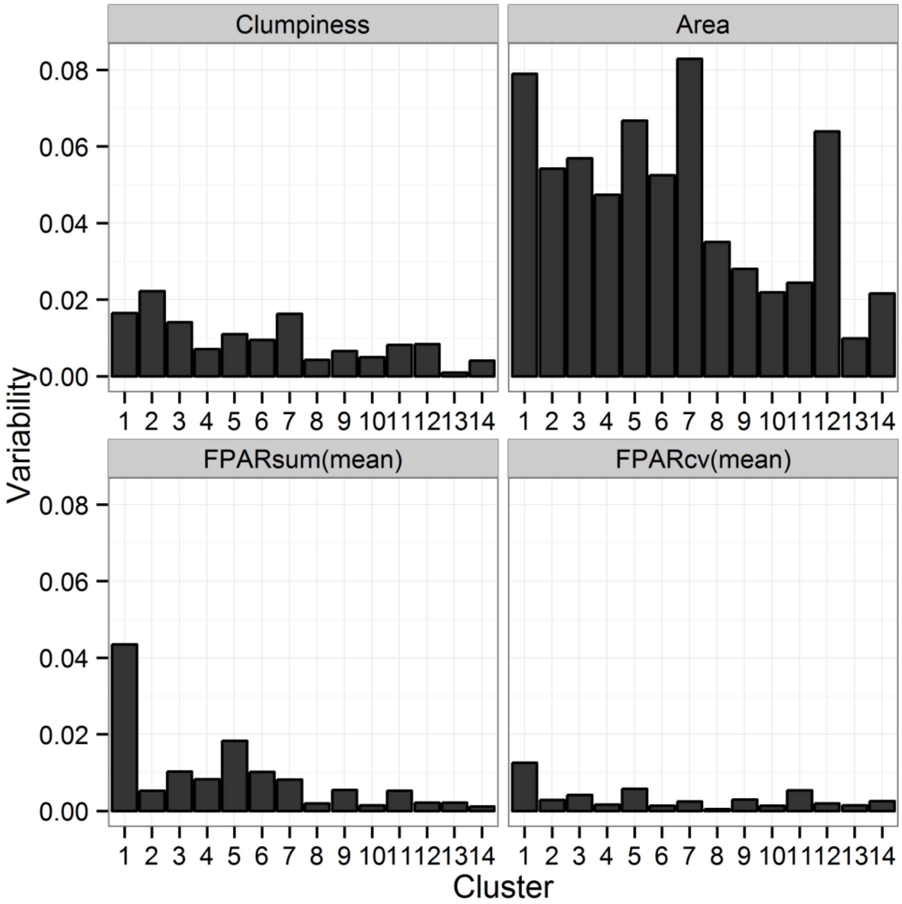


Figure 5.5. Spatial-temporal variability in global ecosystem functional types

5.4 Discussion

5.4.1 Classification and mapping of functional types

Our functional type classification captures an informative proportion of existing global biome and land cover maps, but conveys information regarding functional variability that is unique and complementary to biome and land use mapping. Regionalizations aggregate information and allow for environmental assessment, management and planning (Loveland and Merchant 2004). Depending on scale, perceived importance, and objectives, regions can be constructed using any number or type of environmental variables, including productivity, land cover, topography, geology, climate, and phytographic and zoographic information. Functional types represent collections of species that respond to, or affect one or more ecosystem processes or conditions in a similar manner (Hooper et al. 2005) and regionalizations representing functional types may also be used to assess, monitor, manage and predict changes to ecosystems. Remote sensing can capture many functional properties of vegetation including structure and productivity (Ustin and Gamon 2010). Yet, in the global change literature, regions used to summarize temporal variability in remotely sensed proxies of functions like primary productivity (e.g., fPAR, NDVI) are normally defined by other datasets such as land cover (e.g., Jong et al. 2012; Eastman et al. 2013). We used two remotely sensed indicators of primary productivity and phenology in a data-driven methodology to classify and map ecosystem functional types (Figure 5.2). Primary productivity can vary independently of changes in land cover or vegetation structure because the latter changes more slowly in response to climatic or anthropogenic perturbations (Paruelo et al. 2001). Our clusters capture a combined measure representative of annual greenness and seasonality and we note the two measures ($fPAR_{sum}$ and $fPAR_{cv}$) are inversely related. Pixels that shift from one cluster to the next may be interpreted as areas where climatic change or disturbance, over the time period assessed, led to changes in the cumulative sum and intra-annual variability of fPAR. Such a transition represents a change in ecosystem function, and not necessarily a structural change.

5.4.2 Spatial–temporal variability and change

Many of the changes observed at one time step in this study were not sustained through time, suggesting our analysis is largely capturing short-term vegetation responses to climatic variability or land use and disturbance dynamics. Pixels generally changed from one cluster to another with relatively similar distributions of $fPAR_{sum}$ and $fPAR_{cv}$ values, then back again at a later date (Figure 5.4, Table 5.3). For instance, in Australia in the year 2000, Cluster 7 was more abundant than Cluster 5 (Figure 5.4). Three years later, the opposite was true. In 2006, the amount of Cluster 7 increased again, corresponding to a decrease in the area of Cluster 5. This back and forth nature continued through 2009 and 2012. In Arctic Canada, Clusters 1 and 2 in Arctic Canada had a similar dynamic over time (Figure 5.4). Changes in primary productivity are influenced by long-term global climate change (Boisvenue and Running 2006; Buitenwerf et al. 2015; Keenan et al. 2014; Sitch et al. 2015; Xu et al. 2013). Primary productivity can also vary as a result of climatic patterns and anomalies such as *El Niño Southern Oscillation* (*ENSO*) (Nemani et al. 2003; Peng et al. 2012) and droughts (Zhao and Running 2010), as well as natural disturbances such as volcanic eruptions (Lucht et al. 2002), fire (Goetz et al. 2006; Hicke et al. 2003), and land use/land cover change (Mao et al. 2013). A stable regionalization where boundaries remained fixed over time would not capture this spatial variability in ecosystem function.

Since the boundaries between our regions varied through time, processes leading to variability in annual greenness or annual seasonality were overwhelmingly manifested as changes in spatial composition (area) of the functional types (Figures 5.4 and 5.5). Clusters exhibited variability over time with respect to area (Figure 5.4), often increasing in area at one time step and decreasing at the next (Table 5.2). Clusters at high latitudes representing tundra and taiga ecosystems (e.g., Clusters 1, 2, 4) and those in temperate or subtropical regions representing arid or semi-arid ecosystems (e.g., Clusters 5, 7, 12) had the highest magnitude of change in spatial area, while tropical forests (Clusters 13 and 14) changed the least.

High spatial-temporal variability in the area of arid and semi-arid areas is most likely due to the effect of climatic variability on primary productivity. Other processes such as shrub encroachment and agriculture interact with climate to influence vegetation dynamics and spatial configuration in arid and semi-arid ecosystems (Asner et al. 2004; Knauer et al. 2014) but, at the spatial extent and resolution of this analysis (5 km pixels, global extent), are likely less obvious than broader climatic processes. In arid and semi-arid regions, water is the dominant limiting factor to vegetation productivity, although temperature and solar radiation also interact with water availability to impose constraints on growth (Fensholt et al. 2012; Nemani et al. 2003). Vegetation productivity and phenology in these ecosystems can be particularly sensitive to variability in rainfall; a rain event can produce a sudden flush, or pulse, of growth in vegetation (Noy-Meir 1973). In a study in west-central North America, annual grasses such as cheatgrass (*Bromus tectorum*) were found to be exceptionally reactive to variation in rainfall, and extensive areas of these green areas were easily detectable from space (Bradley and Mustard 2005). Across multiple biomes in North America, Knapp and Smith (2001) found inter-annual variability in above-ground NPP to be considerably greater in grasslands than in forests in North America, and highlighted that the magnitude of production pulses following precipitation pulses were likewise much greater for grasslands. Globally, Jong et al.'s (2012) study identified sequential periods of abrupt greening followed by gradual browning for semi-arid ecosystems such as shrublands and grasslands. Savannas, which consist of both woody and herbaceous vegetation and are found in tropical and subtropical regions, are also known to be a very highly dynamic ecosystem strongly limited by rainfall (Sankaran et al. 2005). Variability is especially high at the drier end of the savanna spectrum, where Ma et al. (2013) found savannas dominated by annual grasses in northern Australia to be the most responsive to variation in precipitation. Woody species, and to a lesser extent, perennial grasses, with their deeper roots that can access water stored below ground, are slower to grow and show less variable responses to precipitation patterns (Bradley and Mustard 2005; Ma et al. 2013; Rich et al. 2008). Some tropical forests exhibit high variability in leaf area and photosynthesis due to seasonality of rainfall and cloud cover (Bi et al. 2015; Huete et al. 2006; Myneni et al. 2007; Xu et al. 2015). In other tropical systems, the ability of trees' roots to access and redistribute soil water helps buffer these systems against seasonal droughts, and trees may remain green and actively photosynthesize for most of the year (Davidson et al. 2012; Ma et al. 2013). Variability in productivity in the Amazon may also be due to large scale forest clear-cutting, grazing, and subsequent agricultural abandonment (Foley et al. 2007; Vieira et al. 2003). Our results demonstrate that on a global scale, vegetation greenness in tropical forests is relatively stable throughout the year here relative to drier, non-forested ecosystems but corroborate the greater variability of tropical pastures (e.g., Cluster 14) compared to tropical forests (e.g., Cluster 13) (Huete et al. 2006).

The results of our research align with studies that have identified increasing trends in remotely sensed indices of productivity in some regions in recent decades (e.g., Beck and Goetz 2012; Mao et al. 2013; Buitenwerf et al. 2015) At high latitudes and high elevations, primary productivity is strongly limited by temperature, although moisture and solar radiation are also limiting (Bliss and Matveyeva 1992; Nemani et al. 2003). Warmer temperatures in Arctic and alpine regions can lead to earlier flowering and greening, and a longer growing season (Oberbauer et al. 2013), and there is substantial evidence of shrub proliferation and increased above-ground biomass (e.g., Lantz et al. 2012; Myers-Smith et al. 2011; Hill and Henry 2011; Fraser et al. 2014; Elmendorf et al. 2012). Primary productivity is also variable over space and time due to variations in snow and ice cover and melt (Choler 2015; Grippa et al. 2005; Stow et al. 2004), which in turn is affected by temperature variability (Rumpf et al. 2014). Thus cumulative annual greenness and variability of greenness will vary considerably if snow cover lasts a month longer or a month less from one year to the next. We also note that remotely sensed estimates of primary productivity in snow-covered areas will increase following snowmelt even if unrelated to increased vegetation growth (Beck et al. 2006). In the sub-arctic and northern Boreal regions (e.g., as represented by portions of Cluster 4), both in-situ and remote-sensing studies have found periods of greening and browning, the latter which has been attributed to drought and temperature stress (Goetz et al. 2011; Jong et al. 2012; Lloyd and Bunn 2007). Large forest fires are also a natural disturbance that can occur over very large, contiguous areas with millions of hectares typically burning annually in the northern Boreal (Kasischke and Turetsky 2006; Stocks et al. 2002) and may explain some of the variability in Cluster 4 (as well as Clusters 8 and 9).

Clusters with high temporal variability in spatial pattern tend to be spatially heterogeneous on the landscape (i.e., occurring within a mosaic of multiple classes for any given time period). For instance, Clusters 1 and 2 occur in Arctic regions where the extreme environment can lead to landscapes with highly patchy (i.e., spatially discontinuous) vegetation (Billings and Mooney 1968; Bliss and Matveyeva 1992) and the spatial configuration of trees at their northern limit is highly variable and dynamic (Harper et al. 2011). Clusters 3, 5, 6 and 7, 10, and 12 which represent grasslands, savannas, and shrublands, are often heterogeneous in that woody vegetation is interspersed in various degrees with herbaceous vegetation, or patches of bare-ground may be present. Cluster 7 for example, includes linear swaths of quite green vegetation within the valleys of arid canyon lands of the western United States (e.g., Nevada and Utah). We expect greenness within some of these ecosystems to be more spatially expansive during the wet-season, whereas greater spatial heterogeneity in greenness would arise in dry periods because of the aforementioned differences in the response of woody and herbaceous species to drought.

Scientists have called for research to address variability in ecosystem functional types caused by climatic and anthropogenic impacts (Ivits et al. 2013) and spatial information has been suggested as a useful addition to a time-based collection of ecosystem change indicators (Kéfi et al. 2014). At local and regional scales, studies have shown that the spatial configuration of productivity and phenological responses affects and is affected by other ecological processes such as seed dispersal, hydrology, pollination, and carbon storage (e.g., Asner et al. 2004). In this study, we have demonstrated that a dynamic regionalization approach is an effective framework to capture temporal variability in the spatial heterogeneity of ecosystem function at global scales.

5.5 Conclusion

We used remotely sensed fPAR data to delineate 14 global ecosystem functional types, each representing a unique productivity and phenology regime. Such classifications, or regionalizations, can be a vital component of global biodiversity and ecosystem service assessment and conservation efforts, as demonstrated by the popularity and widespread use of WWF’s global ecoregion map (Olson et al. 2001). As well, a discrete functional type classification is complementary to monitoring continuous, pixel-based values of productivity or phenology because the classification captures and presents a simplified representation of spatial heterogeneity in ecosystem function. Specifically, a regionalization provides context to changing fPAR values by providing a multivariate or joint-measure assessment of changes in ecosystem function. Our dynamic framework is novel in that it allows an assessment of spatial-temporal variability in ecosystem functional units that would not be possible using a static regionalization.

We have demonstrated that the spatial patterns of primarily productivity in arid and semi-arid ecosystems, including polar regions, are considerably more variable than in tropical forests. Climatic change and land use are major pressures in these ecosystems. Future work should examine the explicit relationship between spatial pattern of global ecosystem functions and hypothesized drivers of such change, such as climatic variability and anthropogenic disturbance. Understanding these linkages will facilitate management and projections of future change. As more data become available and computational power increases, our methodology may be repeated using longer temporal sequences at a higher temporal resolution, enabling the creation of a long-term baseline. Deviations from this baseline spatial-temporal variability may then be detected, signalling areas where more fine-scale monitoring and management efforts may need to be prioritized. Earth observation data provides systematic, repeatable, and globally comprehensive data at a variety of spatial and temporal resolutions that can be used in a quantitative, data-driven and dynamic framework for monitoring changes to essential ecosystem processes in these and other ecosystems.

5.6 References

Alcaraz-Segura, D., Paruelo, J., Epstein, H., & Cabello, J. (2013). Environmental and human controls of ecosystem functional diversity in temperate South America. *Remote Sensing, 5*, 127–154. doi:10.3390/rs5010127

Andrew, M. E., Wulder, M. A., & Nelson, T. A. (2014). Potential contributions of remote sensing to ecosystem service assessments. *Progress in Physical Geography*, *38*(3), 328–353. doi:10.1177/0309133314528942

Andrew, M. E., Wulder, M. A., Nelson, T. A., & Coops, N. C. (2015). Spatial data, analysis approaches, and information needs for spatial ecosystem service assessments: a review. *GIScience & Remote Sensing*, *52*(3), 344–373. doi:10.1080/15481603.2015.1033809

Asner, G. P., Elmore, A. J., Olander, L. P., Martin, R. E., & Harris, A. T. (2004). Grazing systems, ecosystem responses, and global change. *Annual Review of Environment and Resources, 29*(1), 261–299. doi:10.1146/annurev.energy.29.062403.102142

Baeza, S., Lezama, F., Piñeiro, G., Altesor, A., & Paruelo, J. M. (2010). Spatial variability of above-ground net primary production in Uruguayan grasslands: a remote sensing approach. *Applied Vegetation Science*, *13*(1), 72–85. doi:10.1111/j.1654-109X.2009.01051.x

Beck, P. S., Atzberger, C., Høgda, K. A., Johansen, B., & Skidmore, A. K. (2006). Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sensing of Environment*, *100*(3), 321–334. doi:10.1016/j.rse.2005.10.021

Beck, P. S., & Goetz, S. J. (2012). Corrigendum: Satellite observations of high northern latitude vegetation productivity changes between 1982 and 2008: ecological variability and regional differences. *Environmental Research Letters, 7*(2), 029501. doi:10.1088/1748-9326/7/2/029501

Berry, S., Mackey, B., & Brown, T. (2007). Potential applications of remotely sensed vegetation greeness to habitat analysis and the conservation of dispersive fauna. *Pacific Conservation Biology*, *13*, 120–127.

Bi, J., Knyazikhin, Y., Choi, S., Park, T., Barichivich, J., Ciais, P., et al. (2015). Sunlight mediated seasonality in canopy structure and photosynthetic activity of Amazonian rainforests. *Environmental Research Letters*, *10*(6), 064014. doi:10.1088/1748-9326/10/6/064014

Billings, W. D., & Mooney, H. A. (1968). The ecology of arctic and alpine plants. *Biological Review*, *43*, 481–529. doi:10.1111/j.1469-185X.1968.tb00968.x

Bliss, L., & Matveyeva, N. (1992). Circumpolar arctic vegetation. In F. S. I. Chapin, R. Jeffries, J. Reynolds, G. Shaver, & J. Svoboda (Eds.), *Arctic ecosystems in a changing climate: an ecophsiological perspective* (pp. 59–89). San Diego, California: Academic Press Inc.

Boisvenue, C., & Running, S. W. (2006). Impacts of climate change on natural forest productivity - evidence since the middle of the 20th century. *Global Change Biology*, *12*(5), 862–882. doi:10.1111/j.1365-2486.2006.01134.x

Bonan, G. B., Levis, S., Kergoat, L., & Oleson, K. W. (2002). Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Global Biogeochemical Cycles*, *16*(2), 5–1–5–23. doi:10.1029/2000GB001360

Botta, A., Viovy, N., Ciais, P., Friedlingstein, P., & Monfray, P. (2000). A global prognostic scheme of leaf onset using satellite data. *Global Change Biology*, *6*, 709–725.

Bradley, B. A., & Mustard, J. F. (2005). Identifying land cover variability distinct from land cover change: Cheatgrass in the Great Basin. *Remote Sensing of Environment*, *94*(2), 204–213. doi:10.1016/j.rse.2004.08.016

Buitenwerf, R., Rose, L., & Higgins, S. I. (2015). Three decades of multi-dimensional change in global leaf phenology. *Nature Climate Change*, *5*(4), 364–368. doi:10.1038/nclimate2533

Burgess, V. J., Kelly, D., Robertson, A. W., & Ladley, J. J. (2006). Positive effects of forest edges on plant reproduction: literature review and a case study of bee visitation to flowers of Peraxilla tetrapetala (Loranthaceae). *New Zealand Journal of Ecology*, *30*(2), 179–190.

Cabello, J., Fernández, N., Alcaraz-Segura, D., Oyonarte, C., Piñeiro, G., Altesor, A., et al. (2012). The ecosystem functioning dimension in conservation: insights from remote sensing. *Biodiversity and Conservation*, *21*(13), 3287–3305. doi:10.1007/s10531-012-0370-7

Carpenter, S., Mooney, H., Agard, J., Capistrano, D., DeFries, R., Diaz, S., et al. (2009). Science for managing ecosystem services: Beyond the Millennium Ecosystem Assessment. *Proceedings of the National Academy of Sciences*, *106*(5), 1305–1312.

Choler, P. (2015). Growth response of temperate mountain grasslands to inter-annual variations in snow cover duration. *Biogeosciences*, *12*(12), 3885–3897. doi:10.5194/bg-12-3885-2015

Coops, N. C., Wulder, M. A., Duro, D. C., Han, T., & Berry, S. (2008). The development of a Canadian dynamic habitat index using multi-temporal satellite estimates of canopy light absorbance. *Ecological Indicators*, *8*(5), 754–766. doi:10.1016/j.ecolind.2008.01.007

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. doi:10.1016/j.ecoinf.2008.09.005

Crabtree, R., Potter, C., Mullen, R., & Sheldon, J. (2009). A modeling and spatio-temporal analysis framework for monitoring environmental change using NPP as an ecosystem indicator. *Remote Sensing of Environment*, *113*, 1486–1496.

Crews-Meyer, K. A. (2004). Agricultural landscape change and stability in northeast Thailand: historical patch-level analysis. *Agriculture, Ecosystems & Environment*, *101*(2-3), 155–169. doi:10.1016/j.agee.2003.09.024

Currie, D. J., Mittelbach, G. G., Cornell, H. V., Field, R., Guegan, J.-F., Hawkins, B. A., et al. (2004). Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters*, *7*(12), 1121–1134. doi:10.1111/j.1461-0248.2004.00671.x

Davidson, E. a., de Araújo, A. C., Artaxo, P., Balch, J. K., Brown, I. F., C. Bustamante, M. M., et al. (2012). The Amazon basin in transition. *Nature*, *481*(7381), 321–328. doi:10.1038/nature10717

De Jong, R., Verbesselt, J., Zeileis, A., & Schaepman, M. (2013). Shifts in global vegetation activity trends. *Remote Sensing*, *5*(3), 1117–1133. doi:10.3390/rs5031117

Diez, J. M., Ibáñez, I., Miller-Rushing, A. J., Mazer, S. J., Crimmins, T. M., Crimmins, M. A., et al. (2012). Forecasting phenology: from species variability to community patterns. *Ecology letters*, *15*(6), 545–53. doi:10.1111/j.1461-0248.2012.01765.x

Eastman, J., Sangermano, F., Machado, E., Rogan, J., & Anyamba, A. (2013). Global trends in seasonality of Normalized Difference Vegetation Index (NDVI), 1982–2011. *Remote Sensing*, *5*(10), 4799–4818. doi:10.3390/rs5104799

Elmendorf, S. C., Henry, G. H. R., Hollister, R. D., Björk, R. G., Boulanger-Lapointe, N., Cooper, E. J., et al. (2012). Plot-scale evidence of tundra vegetation change and links to recent summer warming. *Nature Climate Change*, *2*(6), 453–457. doi:10.1038/nclimate1465

Evans, K. L., Warren, P. H., & Gaston, K. J. (2005). Species-energy relationships at the macroecological scale: a review of the mechanisms. *Biological reviews of the Cambridge Philosophical Society*, *80*(1), 1–25. doi:10.1017/S1464793104006517

Fahrig, L. (2005). When is a landscape perspective important? In J. Wiens & M. Moss (Eds.), *Issues and Perspectives in Landscape Ecology* (pp. 3–10). Cambridge University Press, Cambridge UK.

Fensholt, R., Langanke, T., Rasmussen, K., Reenberg, A., Prince, S. D., Tucker, C., et al. (2012). Greenness in semi-arid areas across the globe 1981-2007 - an Earth Observing Satellite based analysis of trends and drivers. *Remote Sensing of Environment*, *121*, 144–158. doi:10.1016/j.rse.2012.01.017

Field, R., Hawkins, B. A., Cornell, H. V., Currie, D. J., Diniz-Filho, J. A. F., Guégan, J.-F., et al. (2008). Spatial species-richness gradients across scales: a meta-analysis. *Journal of Biogeography*, *36*(1), 132–147. doi:10.1111/j.1365-2699.2008.01963.x

Foley, J. a, Asner, G. P., Costa, M. H., Coe, M. T., Gibbs, H. K., Howard, E. a, et al. (2007). Amazonia revealed: forest degradation and loss of ecosystem goods and services in the Amazon Basin. *Frontiers in Ecology and the Environment*, *5*, 25–32.

Fraser, R. H., Lantz, T. C., Olthof, I., Kokelj, S. V., & Sims, R. a. (2014). Warming-induced shrub expansion and lichen decline in the western Canadian Arctic. *Ecosystems*, 1151–1168. doi:10.1007/s10021-014-9783-3

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*(1), 168–182. doi:10.1016/j.rse.2009.08.016

Glenn, E. P., Huete, A. R., Nagler, P. L., & Nelson, S. G. (2008). Relationship between remotely-sensed vegetation indices, canopy attributes, and plant physiological processes: what vegetation indices can and cannot tell us about the landscape. *Sensors*, *8*, 2136–2160.

Goetz, S. J., Epstein, H. E., Bhatt, U. S., Jia, G. J., Kaplan, J. O., Lischke, H., et al. (2011). Ch 2: Recent changes in Arctic vegetation: satellite observations and simulation model predictions. In G. Gutman & A. Reissell (Eds.), *Eurasian Arctic Land Cover and Land Use in a Changing Climate* (pp. 9–36). Dordrecht: Springer Netherlands. doi:10.1007/978-90-481-9118-5

Goetz, S. J., Fiske, G. J., & Bunn, A. G. (2006). Using satellite time-series data sets to analyze fire disturbance and forest recovery across Canada. *Remote Sensing of Environment*, *101*(3), 352–365. doi:10.1016/j.rse.2006.01.011

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. doi:10.1007/BF00033449

Grippa, M., Kergoat, L., Le Toan, T., Mognard, N. M., Delbart, N., L'Hermitte, J., & Vicente-Serrano, S. M. (2005). The impact of snow depth and snowmelt on the vegetation variability over central Siberia. *Geophysical Research Letters*, *32*(21), 1–4. doi:10.1029/2005GL024286

Gustafson, E. J. (1998). Quantifying landscape spatial pattern: What is the state of the art ? *Ecosystems*, *1*(2), 143–156.

Handcock, R. N., & Csillag, F. (2002). Ecoregionalization assessment: Spatio-temporal analysis of net primary production across Ontario. *Ecoscience*, *9*(2), 219–230.

Hansen, A., Rotella, J., Kraska, M., & Brown, D. (2000). Spatial patterns of primary productivity in the Greater Yellowstone Ecosystem. *Landscape Ecology*, *15*, 505–522.

Harper, K. A., Danby, R. K., De Fields, D. L., Lewis, K. P., Trant, A. J., Starzomski, B. M., et al. (2011). Tree spatial pattern within the forest–tundra ecotone: a comparison of sites across Canada. *Canadian Journal of Forest Research*, *41*(3), 479–489. doi:10.1139/X10-221

Herrerías-Diego, Y., Quesada, M., Stoner, K. E., & Lobo, J. A. (2006). Effects of forest fragmentation on phenological patterns and reproductive success of the tropical dry forest tree Ceiba aesculifolia. *Conservation Biology*, *20*(4), 1111–1120. doi:10.1111/j.1523-1739.2006.00370.x

Hicke, J. A., Asner, G. P., Kasischke, E. S., French, N. H. F., Randerson, J. T., Collatz, G. J., et al. (2003). Postfire response of North American boreal forest net primary productivity analyzed with satellite observations. *Global Change Biology*, *9*(8), 1145–1157. doi:10.1046/j.1365-2486.2003.00658.x

Hill, G. B., & Henry, G. H. R. (2011). Responses of High Arctic wet sedge tundra to climate warming since 1980. *Global Change Biology*, *17*(1), 276–287. doi:10.1111/j.1365-2486.2010.02244.x

Holmes, K. R., Coops, N. C., Nelson, T. A., Fontana, F. M. A., & Wulder, M. A. (2015). Indicators of vegetation productivity under a changing climate in British Columbia, Canada. *Applied Geography*, *56*, 135–144. doi:10.1016/j.apgeog.2014.11.020

Hooper, D., Chapin, F. S., Ewel, J., Hector, A., Inchausti, P., Lavorel, S., et al. (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological Monographs*, *75*(1), 3–35.

Huete, A. R., Didan, K., Shimabukuro, Y. E., Ratana, P., Saleska, S. R., Hutyra, L. R., et al. (2006). Amazon rainforests green-up with sunlight in dry season. *Geophysical Research Letters*, *33*(6), 2–5. doi:10.1029/2005GL025583

Huete, A. R., Didan, K., van Leeuwen, W., Miura, T., & Glenn, E. (2011). MODIS Vegetation Indices. In B. Ramachandran (Ed.), *Land Remote Sensing and Global Environmental Change, Remote Sensing and Digital Image Processing* (pp. 579–602). Springer. doi:10.1007/978-1-4419-6749-7

IPCC. (2014). *Climate Change 2014 Synthesis Report. Contribution of Working Groups I, II, and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. (R. Pachauri & L. Meyer, Eds.). IPCC, Geneva, Switzerland.

Ivits, E., Cherlet, M., Horion, S., & Fensholt, R. (2013). Global biogeographical pattern of ecosystem functional types derived From earth observation data. *Remote Sensing*, *5*(7), 3305–3330. doi:10.3390/rs5073305

Jain, A. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters, 31*(8), 651–666. doi:10.1016/j.patrec.2009.09.011

Jain, A., Murty, M., & Flynn, P. (1999). Data clustering: a review. *ACM computing surveys (CSUR)*.

Jeong, S.-J., Ho, C.-H., Gim, H.-J., & Brown, M. E. (2011). Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982-2008. *Global Change Biology, 17*(7), 2385–2399. doi:10.1111/j.1365-2486.2011.02397.x

Jong, R., Verbesselt, J., Schaepman, M. E., & Bruin, S. (2012). Trend changes in global greening and browning: contribution of short-term trends to longer-term change. *Global Change Biology, 18*(2), 642–655. doi:10.1111/j.1365-2486.2011.02578.x

Kasischke, E. S., & Turetsky, M. R. (2006). Recent changes in the fire regime across the North American boreal region - Spatial and temporal patterns of burning across Canada and Alaska. *Geophysical Research Letters, 33*(9). doi:10.1029/2006GL025677

Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., et al. (2014). Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Climate Change, 4*(7), 598–604. doi:10.1038/nclimate2253

Kéfi, S., Guttal, V., Brock, W. A., Carpenter, S. R., Ellison, A. M., Livina, V. N., et al. (2014). Early warning signals of ecological transitions: methods for spatial patterns. *PloS one, 9*(3), e92097. doi:10.1371/journal.pone.0092097

Kimball, J. S. (2014). Vegetation phenology. In E. Njoku (Ed.), *Encyclopedia of Earth Sciences Series* (pp. 886–890). New York: Springer New York. doi:10.1007/978-0-387-36699-9_188

Knapp, A., & Smith, M. (2001). Variation among biomes in temporal dynamics of aboveground primary production. *Science, 291*, 481–485.

Knauer, K., Gessner, U., Kuenzer, C., & Dech, S. (2014). Remote sensing of vegetation dynamics in West Africa – a review. *International Journal of Remote Sensing, 35*(17), 6357–6396. doi:10.1080/01431161.2014.954062

Kremen, C. (2005). Managing ecosystem services: what do we need to know about their ecology? *Ecology letters, 8*(5), 468–79. doi:10.1111/j.1461-0248.2005.00751.x

Lantz, T. C., Marsh, P., & Kokelj, S. V. (2012). Recent shrub proliferation in the Mackenzie Delta Uplands and microclimatic implications. *Ecosystems, 16*(1), 47–59. doi:10.1007/s10021-012-9595-2

Lloyd, A. H., & Bunn, A. G. (2007). Responses of the circumpolar boreal forest to 20th century climate variability. *Environmental Research Letters, 2*(4), 045013. doi:10.1088/1748-9326/2/4/045013

Loveland, T. R., & Merchant, J. M. (2004). Ecoregions and ecoregionalization: geographical and ecological perspectives. *Environmental Management*, *34*(S1), S1–S13. doi:10.1007/s00267-003-5181-x

Lucht, W., Prentice, I. C., Myneni, R., Sitch, S., Friedlingstein, P., Cramer, W., et al. (2002). Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. *Science*, *296*(5573), 1687–1689. doi:10.1126/science.1071828

Ma, X., Huete, A. R., Yu, Q., Coupe, N. R., Davies, K., Broich, M., et al. (2013). Spatial patterns and temporal dynamics in savanna vegetation phenology across the North Australian Tropical Transect. *Remote Sensing of Environment*, *139*, 97–115. doi:10.1016/j.rse.2013.07.030

Mao, J., Shi, X., Thornton, P., Hoffman, F., Zhu, Z., & Myneni, R. (2013). Global latitudinal-asymmetric vegetation growth trends and their driving mechanisms: 1982–2009. *Remote Sensing*, *5*(3), 1484–1497. doi:10.3390/rs5031484

McGarigal, K., & Ene, E. (2013). *Fragstats v4.2: Spatial Pattern Analysis Program for Categorial and Continuous Maps.* Producted by the authors at the University of Massachusetts, Amherst. http://www.umass.edu/landeco/research/fragstats/fragstats.html.

Millennium Ecosystem Assessment. (2005). *Ecosystems and human well-being: synthesis*. Island Press, Washington, DC.

Mitchell, M. G. E., Bennett, E. M., & Gonzalez, A. (2013). Linking landscape connectivity and ecosystem service provision: Current knowledge and research gaps. *Ecosystems*, *16*(5), 894–908. doi:10.1007/s10021-013-9647-2Mohamed, M. A. A., Babiker, I. S., Chen, Z. M., Ikeda, K., Ohta, K., & Kato, K. (2004). The role of climate variability in the inter-annual variation of terrestrial net primary production (NPP). *The Science of the Total Environment*, *332*(1-3), 123–37. doi:10.1016/j.scitotenv.2004.03.009

Mooi, E., & Sarstedt, M. (2011). Chapter 9: Cluster analysis. In *A Concise Guide to Market Research*. Berlin Heidelberg: Springer-Verlag.

Morisette, J. T., Richardson, A. D., Knapp, A. K., Fisher, J. I., Graham, E. A., Abatzoglou, J., et al. (2009). Tracking the rhythm of the seasons in the face of global change: phenological research in the 21st century. *Frontiers in Ecology and the Environment*, *7*(5), 253–260. doi:10.1890/070217

Myers-Smith, I. H., Hik, D. S., Kennedy, C., Cooley, D., Johnstone, J. F., Kenney, A. J., & Krebs, C. J. (2011). Expansion of canopy-forming willows over the ywentieth ventury on Herschel Island, Yukon Territory, Canada. *Ambio*, *40*(6), 610–623. doi:10.1007/s13280-011-0168-y

Myneni, R., Hoffman, S., Knyazikhin, Y., Privette, J., Glassy, J., Tian, Y., et al. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment*, *83*, 214–231. doi:10.1016/S0034-4257(02)00074-3

Myneni, R., Yang, W., Nemani, R. R., Huete, A. R., Dickinson, R. E., Knyazikhin, Y., et al. (2007). Large seasonal swings in leaf area of Amazon rainforests. *Proceedings of the National Academy of Sciences*, *104*(12), 4820–4823. doi:10.1073/pnas.0611338104

Neel, M. C., McGarigal, K., & Cushman, S. A. (2004). Behavior of class-level landscape metrics across gradients of class aggregation and area. *Landscape Ecology*, *19*(4), 435–455.

Nelson, T. A., Coops, N. C., Wulder, M. A., Perez, L., Fitterer, J. L., Powers, R. P., & Fontana, F. M. A. (2014). Predicting climate change impacts to the Canadian boreal forest. *Diversity*, *6*(1), 133–157. doi:10.3390/d6010133

Nemani, R. R., Keeling, C. D., Hashimoto, H., Jolly, W. M., Piper, S. C., Tucker, C. J., et al. (2003). Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, *300*(5625), 1560–3. doi:10.1126/science.1082750

Noy-Meir, I. (1973). Desert Ecosystems: Environment and Producers. *Annual Review of Ecoloy and Systematics*, *4*, 25–51. doi:10.1146/annurev.es.04.110173.000325

Oberbauer, S. F., Elmendorf, S. C., Troxler, T. G., Hollister, R. D., Rocha, A. V, Bret-Harte, M. S., et al. (2013). Phenological response of tundra plants to background climate variation tested using the International Tundra Experiment. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *368*(1624), 20120481. doi:10.1098/rstb.2012.0481

Olson, D. M., Dinerstein, E., Wikramanaya, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., et al. (2001). Terrestrial ecoregions of the world: a new map of life on Earth. *BioScience*, *51*(11), 933–938. doi:http://dx.doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Paruelo, J. M., Jobbágy, E. G., & Sala, O. E. (2001). Current distribution of ecosystem functional types in temperate South America. *Ecosystems*, *4*(7), 683–698. doi:10.1007/s10021-001-0037-9

Peng, D., Zhang, B., Liu, L., Fang, H., Chen, D., Hu, Y., & Liu, L. (2012). Characteristics and drivers of global NDVI-based FPAR from 1982 to 2006. *Global Biogeochemical Cycles*, *26*(3), 1–15. doi:10.1029/2011GB004060

Pianka, E. (1966). Latitudinal gradients in species diversity: a review of concepts. *American Naturalist*, *100*(910), 33–46.

Piao, S., Fang, J., Zhou, L., Ciais, P., & Zhu, B. (2006). Variations in satellite-derived phenology in China's temperate vegetation. *Global Change Biology*, *12*(4), 672–685. doi:10.1111/j.1365-2486.2006.01123.x

Polis, G. (1999). Why are parts of the world green? Multiple factors control productivity and the distribution of biomass. *Oikos*, *86*(1), 3–15.

Poulter, B., Ciais, P., Hodson, E., Lischke, H., Maignan, F., Plummer, S., & Zimmermann, N. E. (2011). Plant functional type mapping for earth system models. *Geoscientific Model Development*, *4*(4), 993–1010. doi:10.5194/gmd-4-993-2011

Reed, B. C., Brown, J. F., VanderZee, D., Loveland, T. R., Merchant, J. W., & Ohlen, D. O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, *5*(5), 703–714. doi:10.2307/3235884

Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., & Maury, O. (2013). Dynamic biogeochemical provinces in the global ocean. *Global Biogeochemical Cycles*, *27*(4), 1046–1058. doi:10.1002/gbc.20089

Rich, P. M., Breshears, D. D., & White, A. B. (2008). Pnelology of mixed woody-herbaceous ecosyetms following extreme events: net and differential responses. *Ecology*, *89*(2), 342–352. doi:10.1890/06-2137.1

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., & Toomey, M. (2013). Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agricultural and Forest Meteorology*, *169*, 156–173. doi:10.1016/j.agrformet.2012.09.012

Rumpf, S. B., Semenchuk, P. R., Dullinger, S., & Cooper, E. J. (2014). Idiosyncratic responses of high arctic plants to changing snow regimes. *PLoS ONE*, *9*(2), 1–10. doi:10.1371/journal.pone.0086281

Sankaran, M., Hanan, N. P., Scholes, R. J., Ratnam, J., Augustine, D. J., Cade, B. S., et al. (2005). Determinants of woody cover in African savannas. *Nature*, *438*(7069), 846–849. doi:10.1038/nature04070

Schöngart, J., Piedade, M. T. F., Ludwigshausen, S., Horna, V., & Worbes, M. (2002). Phenology and stem-growth periodicity of tree species in Amazonian floodplain forests. *Journal of Tropical Ecology*, *18*(04), 581–597. doi:10.1017/S0266467402002389

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., et al. (2015). Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, *12*(3), 653–679. doi:10.5194/bg-12-653-2015

Stocks, B. J., Mason, J. a., Todd, J. B., Bosch, E. M., Wotton, B. M., Amiro, B. D., et al. (2002). Large forest fires in Canada, 1959–1997. *Journal of Geophysical Research*, *108*(D1). doi:10.1029/2001JD000484

Stow, D. A., Hope, A., McGuire, D., Verbyla, D., Gamon, J., Huemmrich, F., et al. (2004). Remote sensing of vegetation and land-cover change in Arctic Tundra Ecosystems. *Remote Sensing of Environment*, *89*(3), 281–308. doi:10.1016/j.rse.2003.10.018

Ustin, S. L., & Gamon, J. A. (2010). Remote sensing of plant functional types. *New Phytologist*, *186*, 795–816. doi:10.1111/j.1469-8137.2010.03284.x

Vieira, I. C. G., De Almeida, A. S., Davidson, E. a., Stone, T. a., Reis De Carvalho, C. J., & Guerrero, J. B. (2003). Classifying successional forests using Landsat spectral properties and ecological characteristics in eastern Amazônia. *Remote Sensing of Environment*, *87*(4), 470–481. doi:10.1016/j.rse.2002.09.002

Volante, J. N., Alcaraz-Segura, D., Mosciaro, M. J., Viglizzo, E. F., & Paruelo, J. M. (2012). Ecosystem functional changes associated with land clearing in NW Argentina. *Agriculture, Ecosystems & Environment*, *154*, 12–22. doi:10.1016/j.agee.2011.08.012

White, M. A., & Nemani, R. R. (2006). Real-time monitoring and short-term forecasting of land surface phenology. *Remote Sensing of Environment*, *104*(1), 43–49. doi:10.1016/j.rse.2006.04.014

Whittaker, R. J. (2010). Meta-analyses and mega-mistakes: calling time on meta-analysis of the species richness-productivity relationship. *Ecology*, *91*(9), 2522–2533.

Wright, D. (1983). Species-energy theory: an extension of species-area theory. *Oikos*, *41*(3), 496–506.

Wu, J. (2004). Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecology*, *19*(2), 125–138. doi:10.1023/B:LAND.0000021711.40074.ae

Wu, J., Shen, W., Sun, W., & Tueller, P. T. (2002). Empirical patterns of the effects of changing scale on landscape metrics, 761–782.

Xu, L., Myneni, R., Chapin III, F. S., Callaghan, T. V., Pinzon, J. E., Tucker, C. J., et al. (2013). Temperature and vegetation seasonality diminishment over northern lands. *Nature Climate Change*, *3*(3), 581–586. doi:10.1038/nclimate1836

Xu, L., Saatchi, S. S., Yang, Y., Myneni, R. B., Frankenberg, C., Chowdhury, D., & Bi, J. (2015). Satellite observation of tropical forest seasonality: spatial patterns of carbon exchange in Amazonia. *Environmental Research Letters*, *10*(8), 084005. doi:10.1088/1748-9326/10/8/084005

Zeng, F.-W., Collatz, G., Pinzon, J., & Ivanoff, A. (2013). Evaluating and Quantifying the Climate-Driven Interannual Variability in Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) at Global Scales. *Remote Sensing*, *5*(8), 3918–3950. doi:10.3390/rs5083918

Zhao, M., & Running, S. W. (2010). Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science*, *329*(5994), 940–3. doi:10.1126/science.1192666

6. Discussion and Conclusion

6.1 Summary and Contributions

The overall goal of this dissertation is to advance the mapping and monitoring of biodiversity indicators, nationally and globally, through modelling and frameworks that leverage the information content of remote sensing. In Chapter 2 recommendations are provided for free, open-source, and analysis-ready remotely sensed data products to capture the environmental variables needed for modelling species richness at continental scales. Chapter 3 generated the first spatially continuous, 30 m spatial resolution maps of tree species distributions across the entire forested region of Saskatchewan, Canada. Chapter 4 delineated and described detailed ecosystem structural classes within a remote region of coastal British Columbia. In Chapter 5 global spatial-temporal variability in ecosystem function was quantified through the use of a novel dynamic regionalization and spatial pattern analyses.

The limited availability and geographically, temporally, and taxonomically biased nature of biodiversity data within Canada and across the globe has prompted development of biodiversity targets focused on knowledge development and information access (Table 1.1). One challenge to monitoring is that global biodiversity and remotely sensed datasets, such as global-scale expert range maps of species distributions, and global land cover maps, are often not at a high enough spatial resolution for regional or even national biodiversity monitoring and conservation planning (Secades et al. 2014). Spatial resolutions need to be appropriate (e.g., at ≤ 30 m) for practical conservation at local and regional scales (Buchanan et al. 2009). A first and important contribution of this research is the baseline biodiversity data provided. As opposed to discrete point observations of species observations and sampled ecosystem inventory data, the maps of tree species and ecosystem distributions produced in this research are spatially continuous. As well, vector-based ecosystem maps that are delineated through manual interpretation of remotely sensed data lack the spatial precision of the quantitatively derived ecosystem classes produced in Chapter 4, where each pixel within the spatially continuous map is associated with a quantitative value. Additionally, in Chapters 4 and 5, remotely sensed data from Landsat, RapidEye, and LiDAR systems were used to generate medium-resolution (20 and 30 m) biodiversity data for Canada, which more closely aligns with the needs of users at local and regional scales. The maps of species and ecosystem distributions produced in this research cover compositional, structural and/or functional attributes of biodiversity at both the species and ecosystem levels, helping to fill gaps in biodiversity data from multiple perspectives, and laying the foundation for future monitoring studies.

A second contribution of this research is a description of how different components of biodiversity, at diverse spatial scales, may be mapped and monitored with remote sensing. The capability of remote sensing to directly and explicitly map and monitor biodiversity features useful for monitoring has not always been well demonstrated (Skidmore et al. 2015). In Chapter 2, links are drawn between key environmental predictors of broad scale species richness and remotely sensed products, with recommendations made for each of six global biogeographic realms. In Chapter 5, global primary productivity regimes are delineated in space, with boundaries corresponding to the primary productivity data itself, rather than data less relevant to monitoring ecosystem processes such as land cover. All of the research chapters focus on well understood metrics of biodiversity such as species presence, species richness, ecosystem structure, and ecosystem productivity. The use of terminology familiar to ecologists builds a common reference frame that can facilitate interdisciplinary collaboration, and ultimately, improved development and adoption of remote sensing for biodiversity monitoring (Pettorelli et al. 2014).

A third contribution is the demonstration of how progress towards individual conservation goals and targets, both globally and within Canada, may be monitored with remote sensing. While there is a plethora of remotely sensed data available, linking data and proposed indicators to clear targets or objectives is a fundamental analytical step that requires careful consideration (e.g., Failing and Gregory 2003). All of the research chapters in this dissertation link to conservation targets (Table 6.1) and provide guidance for linking data to indicators. The framework described in Chapter 2 for estimating geographic gradients in species richness at broad scales may be used to assess progress towards the effectiveness of conservation targets such as reducing rates of habitat loss, preventing extinction of threatened species, and ensuring a minimum level of protected area coverage. Moderate spatial resolution (30 m) maps of tree species distributions over large areas (Chapter 3) may be useful for assessing progress towards conservation targets such as sustainable forestry and reducing habitat loss, and for assessing the effectiveness of protected areas. Moderate spatial resolution (20 m) maps that characterize ecosystem structure (as in Chapter 4) may be used to assess progress towards conservation targets such as sustainable forest management, the safeguarding of ecosystem services, reducing the rate of habitat loss, and also may be used to assess the effectiveness of protected areas (Table 6.1). Finally, the dynamic ecosystem functional type classification described in Chapter 5 may be used to assess conservation targets such as sustainable forest management, decreased habitat loss, and maintenance of ecosystem services and ecosystem resilience.

Table 6.1 Summary of indicators assessed in each research chapters with reference to desirable attributes of indicators

Research Result	**Relevant Aichi & Canadian Targets**	**Ease of implementation / repeatability**	**Level of interpretability / ecological meaning**	**Cost-effectiveness**	**Applicable spatial extent and resolution**	**Key contribution and recommendations**
Ch. 2 – Species richness modelling recommendations	Global: 5, 11, 12, 15 Canada: 1	Moderately-high	High	High (many free, open, and analysis-ready datasets available)	Regional to continental extents. Choice of inputs will be region-specific. Constrained by spatial resolution of species data (Hurlbert and White 2005; Rahbek 2004).	Models must be parametrized separately for each of six global biogeographic realms: free and open source remotely sensed data products are recommended for each region
Ch. 3 – Tree species distribution mapping	Global: 5, 7, 11 Canada: 6	Moderately-high	High.	Moderately-high	Local to regional. Requires extensive image processing and also adequate species occurrence data. Unlikely to be successful w image spatial resolutions >30 m	Landsat BAP is useful for mapping detailed species distributions over large areas
Ch. 4 – Ecosystem structural classification	Global: 5, 7, 11, 14, 15 Canada: 1, 3, 6	Moderately-high	Moderate	Low (LiDAR is costly)	Local to regional. Choice of input variables may need to vary by location. Over larger areas, a sampling approach with LiDAR and lower resolution / more affordable RS data may be necessary and possible	-LiDAR can capture 3D vegetation structure, detailed terrain indices, and spatially explicit representation of the landscape to complement traditional expert-based mapping. -Quantitative regionalization is transparent and repeatable.
Ch. 5 – Dynamic ecosystem functional type classification	Global: 5, 14, 15 Canada: 6, 17	Moderately-high	Moderate	High (free, open, and analysis-ready datasets)	Local to global extents.	-Dynamic regionalization useful for capturing short-term spatial-temporal variation in ecosystem function. -Spatial configuration and spatial composition are useful metrics for tracking change

The fourth contribution of this research was a demonstration of the availability and utility of free, open-source, and analysis-ready data at moderate to low spatial resolutions for biodiversity research at both global and sub-global scales. Cost and availability are another considerable challenge to the use of remote sensing for biodiversity mapping and monitoring (W. Turner et al. 2003). With limited funds, organizations are reluctant to devote funds to data or product development if the rewards are uncertain (Vanden Borre et al. 2011). Ecologists and biologists tend to believe that higher spatial resolution data is always better (Nagendra et al. 2013) and that lower resolution data cannot match their needs (Kerr and Ostrovsky 2003; Turner et al. 2003). However, the cost of very high resolution data may limit ecologists from using any remotely sensed data. Free,open-source, and analysis-ready data can help in the global advancement and standardization of biodiversity monitoring by encouraging and facilitating use of remotely sensed data (Wulder and Coops 2014). Furthermore, improved knowledge integration and accessibility is a global and Canadian target for the conservation of biodiversity (Table 1.1). With the exception of the LiDAR and RapidEye data used in Chapter 5, all research chapters used free and open-source data from Landsat or MODIS.

6.2 Limitations and Recommendations for Biodiversity Monitoring in Canada

Canada is a large country, with many remote areas that are unmanaged and poorly monitored, and can surely benefit from a spatially extensive and systematic biodiversity monitoring system driven by remote sensing (Duro et al. 2007; Hermosilla et al. 2015). In the following sections, the applicability of the various research chapters to Canada-wide monitoring is discussed, and specific recommendations regarding data types and methodologies are made.

Species richness monitoring in Canada

Chapter 2 indicated that remotely sensed data can provide many of the important environmental covariates in broad scale models of species richness across the globe. Previous research in Canada suggests that remotely sensed estimates of primary productivity (e.g., fPAR) in particular can be good indicators of the spatial gradient of biodiversity in this country when combined with other environmental datasets. Specifically, environmental regionalizations based on remotely sensed indices of primary productivity have been conducted for various regions in Canada and related to species richness for various taxa including birds (Coops, et al. 2009; Fitterer et al. 2013) and butterflies (Andrew et al. 2011, 2012). Chapter 2 indicated that temperature (e.g., remotely sensed Land Surface Temperature from MODIS) is likely to be an important indicator of biodiversity in Canada in addition to fPAR. The value of Land Surface Temperature data for species richness modelling in Canada should be further investigated in future work. Future work should also test the application of these datasets for all additional regions and species within Canada, in order to provide more confidence in the results and use as an indicator.

Broad scale models of species richness may utilize a variety of remotely sensed predictor data. These remotely sensed data may be of any spatial resolution, but high resolution data will have to be coarsened to match the resolution of species range maps if those are used to parametrize the model (e.g., Hurlbert and White 2005; Rahbek 2004). A possible limitation is that remotely sensed datasets such as land surface temperature and fPAR data are only available for the past ~15 years, whereas climate norms are typically developed from at least 30 years of data. Nonetheless, given the current pace of global change, new methods and shorter-term averages for developing climate normals may be appropriate (Arguez and Vose 2011; Livezey et al. 2007). A second limitation for Canada in particular, is that even though these regions are vegetated and snow-free for several months each year, no fPAR values are made available for the high arctic regions of Canada because they are classified as non-vegetated in the corresponding land cover product. With primary productivity and growing seasons expected to continue to increase for the region (Kaplan 2003), seasonal biophysical data from remote sensors should be provided. Given the high variability of vegetation productivity in northern regions discussed in Chapter 3, the lack of MODIS fPAR data in the high Arctic is unfortunate. Models of species richness built on data from recent decades could be used to forecast potential consequences of global change in Canada, similarly to recent forecasting studies of other biodiversity indicators (e.g., Holmes et al. 2015; Nelson et al. 2014). For regional studies within the Canadian Arctic, Landsat-derived NDVI may be beneficial for biodiversity monitoring (e.g., Gould 2000). Both Landsat and MODIS are free and open-source data-sets with broad spatial coverage.

Canada-wide tree species distribution mapping

In Chapter 3, the first ever 30 m spatial resolution, spatially comprehensive map of tree species distributions was developed for the forested region of Saskatchewan. Future work could apply the approach described in Chapter 4 to all of the forested regions of Canada. The National Forest Inventory (NFI) photo plot dataset used to develop the RandomForest models provides systematic and standardized species data from a network of sampling points across the entire country (Gillis et al. 2005). Annual, Canada-wide Best Available Pixel (BAP) Landsat composites are also now available for 1998 to 2012 (Hermosilla et al. 2015). Future work may seek to assess the performance of the modelling methodology in areas of the country with different types and different numbers of tree species.

Additionally, Hermosilla et al (2015)demonstrate how metrics of forest disturbance and recovery may be calculated using the BAP approach . Such metrics could perhaps be used to add structural attributes to the species distribution maps. Although Landsat data are free and open-source, the extensive nature of the image processing and data storage issues may limit repeatability of the method. In theory, however, the distribution models may be repeated with regular frequency (e.g., annually) as new Landsat BAP composites are compiled. In future years as additional composite products are created, the species maps could be used to contextualize forest change. For instance, fire and forest harvest impacts on the relative distribution of particular species could be determined. Lower spatial resolution data with larger footprints such as MODIS may (Beaudoin et al. 2014) or may not (Nagendra et al. 2013; Wulder, Hall, et al. 2004) provide the spatial resolution needed to accurately map species composition in Canada.

Canada-wide ecosystem structural mapping

Chapter 4 highlighted the similarities and differences between a quantitative, unsupervised approach to ecosystem mapping and an expert-based method of ecosystem mapping in coastal British Columbia. The method is easily repeated and in theory could be used at a range of spatial scales and extents. Other provinces in Canada also use expert-based ecosystem mapping similar to British Columbia (e.g., Ontario’s Ecological Land Classification system), and thus it may be of interest to generate quantitative remotely sensed driven maps of these regions as well. If applied Canada wide, our data-driven approach could be a way of providing nationally standardized (i.e., data-driven) maps, with consistent data sources and methods, rather than a mixture of data at different spatial resolutions, spatial coverage, and following different provincial standards along jurisdictional boundaries.

The main challenge of this research for biodiversity monitoring (i.e., for repeat coverage) is the prohibitive cost of LiDAR, and the fact that it generally has limited spatial and temporal coverage (Skidmore et al. 2015). One solution to the costly nature of LiDAR may be to use a sampling approach over large areas (Wulder, White, Bater, et al. 2012; Wulder, White, Nelson, et al. 2012). Specifically, LiDAR data could be collected over a set of small, spatially discrete areas, related statistically to more affordable and spatially extensive data such as Landsat, then extrapolated to areas that are covered only by the latter. Alternatively, satellite-borne LiDAR data with spatially comprehensive coverage across much of the globe may be a possibility. Forest structural data was captured globally between 2003 and 2009 by a LiDAR system known as the Geoscience Laser Altimeter System (GLAS), aboard the Ice, Cloud and land Elevation SATellite (ICESAT), yet a recent study found large discrepancies in forest structural measurements between these data and regional airborne LiDAR data in Canada (Bolton et al. 2013). A new global LiDAR system purpose-built for vegetation structural mapping is planned to begin capturing data beginning in 2019, but with coverage planned for 50°S to 50°N, applicability in Canada of the new Global Ecosystem Dynamics Investigation (GEDI) will be somewhat limited.

Dynamic ecosystem functional type mapping in Canada

The dynamic ecosystem functional type classification described in Chapter 5 could be applied at local to global extents, with different remotely sensed estimates of primary productivity (e.g., fPAR, NDVI, EVI, or NPP) as appropriate or as available. Dynamic ecosystem functional type mapping could easily be completed within the boundaries of Canada (with the exception of the high Arctic if using fPAR data from MODIS). However, an advantage of the global approach is that it provides context to the Canadian results. For example, change in the Arctic regions of Canada may be compared to that in Arctic regions of Eurasia. With free, open-source, and analysis-ready data such as MODIS fPAR, the method is very cost effective. However, the same two limitations discussed previously (temporal and geographic coverage of MODIS fPAR data) are also applicable here. Unfortunately, few, if any, other sensors match the high temporal resolution of MODIS data, however, fusion of data from Landsat 7 ETM+, Landsat 8, and Sentinel-2 may be an option to generate the temporal coverage needed to capture intra-annual spatial-temporal variability in ecosystem function.

Higher spatial resolution sensors such as Landsat may be more ideal for spatial pattern characterization. Although spatial configuration and composition were both shown to be dynamic at the global scale in Chapter 2, linking pattern to process is a challenge at this scale. Landscape pattern is shaped by, and affects ecological processes, and thus measurements of pattern, pattern variability, and changes in pattern and pattern variability, may be used to better understand those processes (Fahrig 2005; Fraterrigo and Rusak 2008; M. Turner 2010; Underwood et al. 2000). While forest fragmentation is generally agreed upon as a useful global biodiversity indicator (Skidmore et al. 2015; Geijzendorffer et al. 2015; www.bipindicators.net), the use of dynamic data-driven boundaries and the assessment of spatial composition and configuration in non-forested biomes is a novel idea that could be very beneficial to understand temporal patterns and process in biodiversity. However, many processes such as shrub encroachment in grasslands, agricultural dynamics (e.g., crop alternating), windthrow, forest harvesting, and insect disturbance may occur at spatial scales < 5 km; at coarser scales these effects are spatially aggregated. The optimal balance between spatial and temporal resolution will be project-specific.

6.3. Conclusion

Biodiversity is a complex, multifaceted and hierarchical concept that includes genes, species and ecosystems. This research has demonstrated some of the potential of remotely sensed data for mapping compositional, structural and functional attributes of biodiversity at the species and ecosystem level and at local, regional, and global scales. First, I highlighted that remote sensing can provide the necessary environmental correlates needed to model very broad scale species richness gradients, including underutilized climatic variables. Opting to use remotely sensed climate variables over interpolated climate datasets may lead to improved models in remote areas where the availability and quality of such data may be limited. Further, the frequent acquisition of the remote sensing data means these types of models can be used to generate newer, shorter-term climate normals reflecting our changing climate, and can to monitor potential changes to global species distributions with ease. Second, spatially comprehensive maps of tree species distributions were predicted for an unprecedented spatial resolution (30 m) across a large area of Saskatchewan, Canada. In this work, a pixel-based image compositing technique was shown to have potential to generate similar species distribution maps, and thus to monitor changes in forest composition, at a national scale and at frequent intervals. Third, ecosystem units were delineated in coastal British Columbia, characterizing three-dimensional vegetation structure and detailed topographic features. These baseline conditions in vegetation structure can be monitored overtime as new LiDAR data are acquired. Finally, in Chapter 5 global spatial-temporal variability in primary productivity was quantified through the use of a novel dynamic regionalization and spatial pattern analyses. The suggested framework for monitoring a key ecosystem function may be extended in time as new data are acquired and results indicate that more detailed monitoring should be targeted to arid, semi-arid, and polar environments.

Together these research chapters show how remote sensing can provide low-cost, repeatable, data on biodiversity indicators at a variety of spatial scales to facilitate large-area monitoring. The use of remote sensing data to monitor indicators of biodiversity will become more and more feasible as remaining issues related to cost, availability, and ease of use are addressed. These concessions will reduce the need to make trade-offs between large area coverage, temporal resolution and spatial resolution. Importantly, the broad spatial coverage and regular frequency over which remotely sensed data are collected will need to be increasingly combined with ecological data collected *in-situ*, as well as other spatial environmental datasets (e.g., climate, land use) in quantitative models to improve our understanding of biodiversity over large areas. As the availability and integration of remotely sensed data and data collected *in-situ* increases, we will improve our understanding of drivers and responses of biodiversity. With greater knowledge, management and conservation planning activities can be advanced to help meet global biodiversity conservation goals.

6.4 References

Andrew, M. E., Wulder, M. A., & Coops, N. C. (2011). How do butterflies define ecosystems? A comparison of ecological regionalization schemes. *Biological Conservation, 144*(5), 1409–1418. doi:10.1016/j.biocon.2011.01.010

Andrew, M. E., Wulder, M. A., Coops, N. C., & Baillargeon, G. (2012). Beta-diversity gradients of butterflies along productivity axes. *Global Ecology and Biogeography, 21*(3), 352–364. doi:10.1111/j.1466-8238.2011.00676.x

Arguez, A., & Vose, R. S. (2011). The definition of the standard WMO climate normal: The key to deriving alternative climate normals. *Bulletin of the American Meteorological Society*, *92*(6), 699–704. doi:10.1175/2010BAMS2955.1

Beaudoin, A., Bernier, P. Y., Guindon, L., Villemaire, P., Guo, X. J., Stinson, G., et al. (2014). Mapping attributes of Canada's forests at moderate resolution through k NN and MODIS imagery. *Canadian Journal of Forest Research*, *44*(5), 521–532. doi:10.1139/cjfr-2013-0401

Bolton, D. K., Coops, N. C., & Wulder, M. A. (2013). Investigating the agreement between global canopy height maps and airborne Lidar derived height estimates over Canada. *Canadian Journal of Remote Sensing*, *39*(S1), 139–151. doi:10.5589/m13-036

Buchanan, G. M., Nelson, A., Mayaux, P., Hartley, A., & Donald, P. F. (2009). Delivering a global, terrestrial, biodiversity observation system through remote sensing. *Conservation Biology*, *23*(2), 499–502. doi:10.1111/j.1523-1739.2008.01083.x

Coops, N. C., Wulder, M. A., & Iwanicka, D. (2009). Exploring the relative importance of satellite-derived descriptors of production, topography and land cover for predicting breeding bird species richness over Ontario, Canada. *Remote Sensing of Environment*, *113*(3), 668–679. doi:10.1016/j.rse.2008.11.012

Duro, D. C., Coops, N. C., Wulder, M. A., & Han, T. (2007). Development of a large area biodiversity monitoring system driven by remote sensing. *Progress in Physical Geography*, *31*(3), 235–260. doi:10.1177/0309133307079054

Fahrig, L. (2005). When is a landscape perspective important? In J. Wiens & M. Moss (Eds.), *Issues and Perspectives in Landscape Ecology* (pp. 3–10). Cambridge University Press, Cambridge UK.

Failing, L., & Gregory, R. (2003). Ten common mistakes in designing biodiversity indicators for forest policy. *Journal of Environmental Management*, *68*(2), 121–132. doi:10.1016/S0301-4797(03)00014-8

Fitterer, J., Nelson, T., Coops, N. C., Wulder, M. A., & Mahony, N. A. (2013). Exploring the ecological processes driving geographical patterns of breeding bird richness in British Columbia, Canada. *Ecological Applications*, *23*(4), 888–903.

Fraterrigo, J. M., & Rusak, J. a. (2008). Disturbance-driven changes in the variability of ecological patterns and processes. *Ecology Letters*, *11*(7), 756–70. doi:10.1111/j.1461-0248.2008.01191.x

Geijzendorffer, I. R., Regan, E. C., Pereira, H. M., Brotons, L., Brummitt, N., Gavish, Y., et al. (2015). Bridging the gap between biodiversity data and policy reporting needs: An Essential Biodiversity Variables perspective. *Journal of Applied Ecology*, n/a–n/a. doi:10.1111/1365-2664.12417

Gillis, M. D., Omule, A. Y., & Brierley, T. (2005). Monitoring Canada's forests: The National Forest Inventory. *The Forestry Chronicle*, *81*(2), 214–221. doi:10.5558/tfc81214-2

Gould, W. (2000). Remote sensing of vegetation, plant species richness, and regional biodiversity hotspots. *Ecological Applications*, *10*(6), 1861–1870.

Hermosilla, T., Wulder, M. a., White, J. C., Coops, N. C., & Hobart, G. W. (2015). An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. *Remote Sensing of Environment*, *158*, 220–234. doi:10.1016/j.rse.2014.11.005

Holmes, K. R., Coops, N. C., Nelson, T. A., Fontana, F. M. A., & Wulder, M. A. (2015). Indicators of vegetation productivity under a changing climate in British Columbia, Canada. *Applied Geography*, *56*, 135–144. doi:10.1016/j.apgeog.2014.11.020

Hurlbert, A. H., & White, E. P. (2005). Disparity between range map- and survey-based analyses of species richness: patterns, processes and implications. *Ecology Letters*, *8*(3), 319–327. doi:10.1111/j.1461-0248.2005.00726.x

Kaplan, J. O. (2003). Climate change and Arctic ecosystems: 2. Modeling, paleodata-model comparisons, and future projections. *Journal of Geophysical Research*, *108*(D19), 8171. doi:10.1029/2002JD002559

Kerr, J. T., & Ostrovsky, M. (2003). From space to species: ecological applications for remote sensing. *Trends in Ecology & Evolution*, *18*(6), 299–305. doi:10.1016/S0169-5347(03)00071-5

Livezey, R. E., Vinnikov, K. Y., Timofeyeva, M. M., Tinker, R., & van den Dool, H. M. (2007). Estimation and extrapolation of climate normals and climatic trends. *Journal of Applied Meteorology and Climatology*, *46*(11), 1759–1776. doi:10.1175/2007JAMC1666.1

Nagendra, H., Lucas, R., Honrado, J. P., Jongman, R. H. G., Tarantino, C., Adamo, M., & Mairota, P. (2013). Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecological Indicators*, *33*, 45–59. doi:10.1016/j.ecolind.2012.09.014

Nelson, T. A., Coops, N. C., Wulder, M. A., Perez, L., Fitterer, J. L., Powers, R. P., & Fontana, F. M. A. (2014). Predicting climate change impacts to the Canadian boreal forest. *Diversity*, *6*(1), 133–157. doi:10.3390/d6010133

Pettorelli, N., Safi, K., Turner, W., & B, P. T. R. S. (2014). Satellite remote sensing, biodiversity research and conservation of the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1643), 20130190. doi:10.1098/rstb.2013.0190

Rahbek, C. (2004). The role of spatial scale and the perception of large-scale species-richness patterns. *Ecology Letters*, *8*(2), 224–239. doi:10.1111/j.1461-0248.2004.00701.x

Secades, C., O'Connor, B., Brown, C., & Walpole, M. (2014). *Earth Observation for Biodiversity Monitoring: A review of current approaches and future opportunities for tracking progress towards the Aichi Biodiversity Targets.* CBD Technical Series No 72. Montreal, Canada: Secretariat of the Convention on Biological Diversity.

Skidmore, A. K., Pettorelli, N., Coops, N. C., Geller, G., Hansen, M. C., Lucas, R., et al. (2015). Environmental science: Agree on biodiversity metrics to track from space. *Nature*, *523*(7561), 403–405. doi:10.1038/523403a

Turner, M. (2010). Disturbance and landscape dynamics in a changing world. *Ecology*, *91*(10), 2833–2849.

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., & Steininger, M. (2003). Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution*, *18*(6), 306–314. doi:10.1016/S0169-5347(03)00070-3

Underwood, A. J., Chapman, M. G., & Connell, S. D. (2000). Observations in ecology: You can't make progress on processes without understanding the patterns. *Journal of Experimental Marine Biology and Ecology*, *250*(1-2), 97–115. doi:10.1016/S0022-0981(00)00181-7

Vanden Borre, J., Paelinckx, D., Mücher, C. a., Kooistra, L., Haest, B., De Blust, G., & Schmidt, A. M. (2011). Integrating remote sensing in Natura 2000 habitat monitoring: Prospects on the way forward. *Journal for Nature Conservation*, *19*(2), 116–125. doi:10.1016/j.jnc.2010.07.003

Wulder, M. A., & Coops, N. C. (2014). Make Earth observations open access. *Nature*, *513*, 30–31.

Wulder, M. A., Hall, R. J., Coops, N. C., & Franklin, S. E. (2004). High spatial resolution remotely sensed data for ecosystem characterization. *BioScience*, *54*(6), 511–521. doi:10.1641/0006-3568(2004)054[0511:HSRRSD]2.0.CO;2

Wulder, M. A., White, J. C., Bater, C. W., Coops, N. C., Hopkinson, C., & Chen, G. (2012). Lidar plots — a new large-area data collection option: context, concepts, and case study. *Canadian Journal of Remote Sensing*, *38*(05), 600–618. doi:10.5589/m12-049

Wulder, M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., et al. (2012). Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment*, *121*, 196–209. doi:10.1016/j.rse.2012.02.001