

Université de Montréal

**Régionalisation et synthèse des patrons de la végétation du
Québec : Utilisation d'indices de patrons à l'échelle provinciale**

par

Kevin Partington

Département de géographie
Faculté des arts et des sciences

Mémoire présenté à la Faculté des arts et des sciences
en vue de l'obtention du grade de maîtrise
en géographie

Octobre 2012

© Kevin Partington, 2012

Université de Montréal
Faculté des études supérieures et postdoctorales

Ce mémoire intitulé :

Régionalisation et synthèse des patrons de la végétation du Québec :
Utilisation d'indices de patrons à l'échelle provinciale

Présenté par :
Kevin Partington

a été évalué par un jury composé des personnes suivantes :

Oliver Sonnentag, président-rapporteur
Jeffrey Cardille, directeur de recherche
Patrick James, membre du jury

Résumé

Le Québec est une immense province à l'intérieur de laquelle existe une grande diversité de conditions bioclimatiques et où les perturbations anthropiques et naturelles du couvert végétal sont nombreuses. À l'échelle provinciale, ces multiples facteurs interagissent pour sculpter la composition et la distribution des paysages. Les objectifs généraux de cette recherche visaient à explorer et comprendre la distribution spatiale des patrons des paysages du Québec, de même qu'à caractériser les patrons observés à partir d'images satellitaires. Pour ce faire, les patrons des paysages ont été quantifiés avec un ensemble complet d'indices calculés à partir d'une cartographie de la couverture végétale. Plusieurs approches ont été développées et appliquées pour interpréter les valeurs d'indices sur de vastes étendues et pour cartographier la distribution des patrons des paysages québécois. Les résultats ont révélé que les patrons de la végétation prédits par le Ministère des Ressources naturelles du Québec divergent des patrons de la couverture végétale observée. Ce mémoire dresse un portrait des paysages québécois et les synthétise de manière innovatrice, en plus de démontrer le potentiel d'utilisation des indices comme attributs biogéographiques à l'échelle nationale.

Mots-clés : Québec, paysage, patron, indice, groupement, exemplar, forêt, classification, représentativité, affinity propagation

Abstract

Quebec is a vast province in which bioclimatic conditions, land-uses and land-cover changes are highly diverse. At this scale, multiple drivers interact to have an impact on the composition and configuration of landscape patterns. The main objectives of this research were to explore and better understand the spatial distribution of landscape patterns across Quebec, and to characterize observed patterns as seen from satellite. To achieve these objectives, we quantified landscape patterns with an extensive set of metrics measured from a categorical land-cover map. We developed and applied several approaches to interpret metric values across large areas, and to map the distribution of Quebec landscape patterns. Results revealed that ecological subzones developed by the Ministère des Ressources naturelles et de la Faune were substantially inconsistent with observed land-cover patterns. This master thesis portrays and synthesizes Quebec landscapes in an innovative way, highlighting the considerable potential of use of landscape metrics for broad-scale biogeographic mapping.

Keywords : Quebec, landscape, pattern, metric, clustering, exemplar, forest, classification, representativeness, and affinity propagation.

Table des matières

Liste des tableaux	v
Liste des figures	vi
Liste des abréviations.....	vii
Remerciements	viii
Introduction	1
1. Contexte général de recherche du mémoire	4
1.1 Le Québec et ses paysages	4
1.2 La notion de paysage	6
1.2.1 Définition	6
1.2.2 Unités d'étude – Le paysage	7
1.3 L'étude des patrons de paysages à petites échelles géographiques	9
1.4 Les indices de patrons des paysages	11
1.5 La classification et la régionalisation des données à petite échelle géographique.....	14
1.5.1 Synthétiser le contenu d'information pour mieux comprendre.....	14
1.5.2 Affinity Propagation.....	16
1.5.3 Les travaux de Long et al. (2010)	17
1.5.4 Les arbres de classification	18
1.6 Quantification et spatialisation des divergences	19
1.7 Les données de base pour cette recherche	22
1.7.1 La couverture du sol de l'OTDD	22
1.7.2 Limites d'utilisation des données de l'OTDD	23
1.8 Objectifs généraux du mémoire.....	24
2. Premier article	27
2.1 Introduction.....	28
2.2 Methods.....	30
2.2.1 Study Area.....	30
2.2.2 EOSD land-cover data and landscape metrics	31
2.2.3 Extracting information from metric values, clustering and determining representatives	32
2.2.4 Mapping centrality of exemplars	33
2.2.5 Spatial fences	34

2.3 Results	34
2.3.1 Composite image of PCA values	34
2.3.2 Representative land-cover patterns of forested Quebec	36
2.3.3 Spatial clusters	39
2.3.4 Spatial fences	42
2.4 Discussion.....	44
2.4.1 Landscape pattern information and clustering	44
2.4.2 Representative Landscapes	45
2.4.3 Representativeness of exemplars and homogeneity of clusters	46
2.4.4 Spatial fences locate important landscape pattern differences	47
2.4.5 Limits of research and future work	47
2.5 Conclusion	47
3. Deuxième article	49
3.1 Introduction.....	50
3.2 Methods.....	53
3.2.1 Study area	53
3.2.2 Land-cover data and landscape metrics	54
3.2.3 Constrained clustering: interpreting characteristics of landscapes in vegetation subzones	55
3.2.4 Unconstrained classification of landscapes: understanding and mapping broad-scale landscape pattern.....	56
3.3 Results	57
3.3.1 Constrained clustering of landscape patterns in Southern Quebec	57
3.3.2 Unconstrained clustering of landscape patterns in Southern Quebec	65
3.3.3 Comparing constrained and unconstrained classifications of Quebec landscapes.....	70
3.4 Discussion.....	71
3.4.1 Constrained clustering.....	71
3.4.2 Unconstrained clustering.....	73
3.4.3 Spatial fences and the diversity map	74
3.5 Conclusion	75
4. Conclusion générale du mémoire.....	78
Bibliographie.....	81

Liste des tableaux

TABLE 1. MISCLASSIFICATION MATRIX FOR A CLASSIFICATION BASED ON THE THREE-LEAF TREE.	60
TABLE 2. MISCLASSIFICATION MATRIX FOR A CLASSIFICATION BASED ON THE NINE-LEAF TREE.	61

Liste des figures

FIGURE 1. EXEMPLE DE PAYSAGE DÉCOUPÉ SELON LES BORDURES DU FEUILLET TOPOGRAPHIQUE 31P05 DU SRNC À L'ÉCHELLE 1:50000.....	8
FIGURE 2. EXEMPLE DE PAYSAGE DÉCOUPÉ SELON LE DISTRICT ÉCOLOGIQUE 31L009 DU MRNF.....	9
FIGURE 3. THE PROVINCE OF QUÉBEC BOUNDARIES AND, IN BACKGROUND, THE EOSD LAND COVER CHARACTERIZES FORESTED LANDSCAPES.....	31
FIGURE 4. COMPOSITE IMAGE CONSTITUTED OF THE THREE FIRST PRINCIPAL COMPONENTS AS OPTICAL BANDS (PC1 IN RED, PC2 IN GREEN, PC3 IN BLUE).....	36
FIGURE 5. CLUSTERING SET OF FIVE WITH REPRESENTATIVE LANDSCAPES (EXEMPLARS) LOCATED WITHIN THEIR CLUSTERS IN QUEBEC.....	38
FIGURE 6. CLUSTERING OUTPUTS (TOP ROW) COMPARED WITH HIERARCHICAL DIVISIONS OF THE MRNF ECOLOGICAL CLASSIFICATIONS (BOTTOM ROW).....	41
FIGURE 7. FROM THE 54 CLUSTERING LEVELS, SEGMENTS OF REOCCURRING SPATIAL CLUSTER BORDERS WERE DRAWN TO LOCATE THE MOST PRONOUNCED VARIATION AMONG ADJACENT LANDSCAPES IN FORESTED QUEBEC.....	43
FIGURE 8. SOUTHERN QUEBEC CATEGORIZED BY THREE VEGETATION SUBZONES (Z11-ORANGE: DECIDUOUS FOREST; Z12-GREEN: MIXED-WOOD FOREST; Z21-BLUE: CONTINUOUS BOREAL FOREST), AND BY ECOLOGICAL DISTRICTS (DARK GRAY CONTOUR LINES POLYGONS).....	53
FIGURE 9. A THREE-LEAF TREE FOR 2489 LANDSCAPES CONSTRAINED TO MRNF VEGETATION SUBZONES (NORTHERN DECIDUOUS FOREST, MIXED-WOOD FOREST, CONTINUOUS BOREAL FOREST).....	58
FIGURE 10. RESULTING CLASSIFICATION OF THE MAXIMAL NINE-LEAF TREE FOR 2489 LANDSCAPES CONSTRAINED TO MRNF VEGETATION SUBZONES (NORTHERN DECIDUOUS FOREST, MIXED-WOOD FOREST, CONTINUOUS BOREAL FOREST).....	59
FIGURE 11. SPATIAL CHARACTERISTICS OF THE CONSTRAINED CLASSIFICATIONS.....	63
FIGURE 12. THE MISCLASSIFICATION MAP OF CONSTRAINED CLASSIFICATIONS IDENTIFIES WHERE THE REFERENCE MRNF CLASSIFICATION IS INCOMPATIBLE WITH OBSERVED LAND-COVER PATTERNS.....	65
FIGURE 13. UNCONSTRAINED CLUSTERING OF THREE CLUSTERS CALCULATED WITH AFFINITY PROPAGATION USING LANDSCAPE PATTERN METRICS.....	66
FIGURE 14. SPATIAL FENCES GENERATED FROM REPEATED CLUSTERING ITERATIONS.....	68
FIGURE 15. THE SPATIAL CLUSTER DIVERSITY MAP CHARACTERIZES THE PURENESS OF THE LANDSCAPES WITHIN 30 KM BY 30 KM BLOCKS.....	69
FIGURE 16. DIVERSITY OF UNCONSTRAINED SPATIAL PATTERN REGIONS WITHIN CONSTRAINED CLASSIFICATION INCONSISTENCIES.....	71

Liste des abréviations

AP	<i>Affinity Propagation</i> : Algorithme de groupage qui permet simultanément de grouper les éléments similaires ensemble et d'en déterminer les éléments représentatifs de chacun (<i>exemplars</i>). Développé par Frey et Dueck de l'Université de Toronto (2007).
AVHRR	<i>Advanced Very High Resolution Radiometer</i> .
CLARA	<i>Clustering for LARge Applications</i> : Algorithme de groupage (Long et al. 2010).
EOSD	<i>Earth Observation for Sustainable Developments of Forests</i> : Cartographie matricielle de la couverture boisée du Canada produite à partir d'imageries Landsat vers l'an 2000 (Wulder et al. 2008a)
Landsat	Programme d'observation de la Terre par imagerie satellitaire développé par la NASA depuis le début des années 1970 (NASA 2012b).
LDCM	<i>Landsat Data Continuity Mission</i> : Projet de plate-forme qui remplacera l'actuelle et défectueuse Landsat-7 ETM+. Le projet est censé démarrer ses activités à la fin de 2012 (NASA 2012a).
MNR	<i>Ministère des Ressources naturelles du Québec</i> .
MNRF	<i>Ministère des Ressources naturelles et de la Faune du Québec</i> .
MODIS	<i>Moderate Resolution Imaging Spectroradiometer</i> .
OTDD	<i>Observation de la Terre pour le Développement Durable des forêts</i> : Traduction de EOSD.
SAFORAH	<i>System of Agents for Forest Observation Research with Advanced Hierarchies</i> : Réseau de distribution de données géospatiales (www.saforah.org)

Remerciements

Ce mémoire de maîtrise est le résultat d'un plus de deux ans de travail. Il ne m'aurait été possible d'accomplir ces travaux de recherche sans la patience, la confiance et surtout l'appui de plusieurs personnes. Parmi eux, je souhaite tout d'abord remercier profondément mes parents Louise et Michel de m'avoir soutenu dans mes décisions depuis leur pas si lointain Cap-de-la-Madeleine. À Pamela, qui a su me tenir bonne compagnie malgré l'arythmie qu'engendre la recherche universitaire et qui a toujours su me conseiller dans mes innombrables présentations. À mon directeur, Jeffrey, sans qui je n'aurais jamais écrit ces lignes s'il n'avait été de m'avoir accepté dans son laboratoire dès le cheminement *Honours*. À Marc-Élie et Mylène, pour leur grande contribution inattendue dans cette recherche. À mes collègues du laboratoire et mes amis, Alexis, Mathieu, Jean-Baptiste, Kate, et David pour avoir égayé les longs mois passés dans le déprimant Strathcona. Aux «complexés» qui ont su stimuler l'intellect du 4^e étage avec les nombreuses pauses café et les croque-complexes. À mes amis Marc-André et Mike, qui ont su me tenir compagnie virtuellement pendant toutes ces innombrables heures passées devant le moniteur de mon ordinateur. Enfin, à Benjamin et Jacob, à qui je souhaite une longue vie remplie de bonheur.

Introduction

Les classifications écologiques et les cartographies de couverture et d'utilisation du sol se sont établies comme des produits de référence pour les études environnementales aux échelles régionales, nationales, continentales et maintenant globales. En effet, les scientifiques autant que les gestionnaires les utilisent pour étudier les phénomènes géographiques, par exemple la distribution des espèces végétales ou encore pour gérer les ressources à l'intérieur d'unités spatiales connues, comme les unités d'aménagement forestier. L'ajout d'information complémentaire par le biais de données auxiliaires (indices, données de recensements, données historiques, etc.) permet d'étendre les possibilités d'application d'études d'envergure provinciale et nationale, et d'adapter les cartographies existantes aux nouveaux besoins scientifiques. Les indices de patrons de paysage sont un bel exemple de données auxiliaires qui peuvent être associées aux classifications. À partir de classifications de la couverture terrestre, ils permettent de quantifier les patrons des parcelles, autant en composition qu'en configuration, dans les paysages. Ils ont ainsi la capacité à distinguer quantitativement les patrons. Toutefois, il existe des limites notoires à leur utilisation, notamment lorsqu'ils sont utilisés à grande échelle (choix, interprétation, production énorme de données, exigences de calculs, etc.). Conséquemment, ces indices n'ont que très peu été explorés à des fins de cartographie et de classification écologique, mais il n'en demeure pas moins qu'ils offrent un fort potentiel pour la caractérisation des patrons à petite échelle géographique puisque des recherches récentes ont démontré leur utilité pour les études environnementales à grande échelle.

La Province de Québec est un immense territoire dont le façonnement des patrons des paysages est influencé par une gamme de facteurs biophysiques et de perturbations du couvert végétal de diverses origines. Les classifications écologiques de référence, comme le système hiérarchique de classification écologique du Ministère des Ressources naturelles et de la Faune du Québec (MRNF), expriment bien la complexité et la diversité des patrons des paysages québécois en découpant, selon des critères établis en fonction de l'échelle d'étude, le territoire en différentes régions. En général, ce type de classification se base – et avec raisons – sur des facteurs biophysiques stables dans l'environnement. Naturellement,

certaines des facteurs environnementaux, par exemple le climat, les précipitations et les perturbations, varient au gré du temps et ils peuvent avoir potentiellement des répercussions sur les paysages. Considérant la variabilité des facteurs, nous pouvons poser la question à savoir si les patrons prédits dans les ouvrages de référence concordent bien avec les actuels patrons observables. Or, certains scientifiques pourraient souhaiter connaître l'état réel ou observable, plutôt que théorique, des patrons des paysages du Québec, car cette information pourrait avoir des impacts sur l'interprétation des classifications écologiques et sur les connaissances de la représentativité des conditions exprimées dans celle-ci.

À partir de ce constat, nous nous sommes posé quelques questions générales : est-il possible de régionaliser les paysages observés uniquement à partir d'indices de patrons? Quels sont les paysages les plus représentatifs du Québec? Quels types de patrons sont-ils caractérisés? Y a-t-il des différences majeures entre les paysages observés et les paysages prédits par la classification de référence? Si oui, où ces différences sont-elles situées? Les réponses à ces questions générales contribueraient, non seulement, à nos connaissances sur les patrons des paysages québécois, mais elles contribueraient aussi à étendre nos connaissances sur leur applicabilité à l'échelle nationale et à l'évaluation de la représentativité des classifications déjà produites. Bref, les objectifs généraux de ce mémoire sont : d'explorer les capacités réelles des indices à discriminer les différents types de patrons des paysages à l'échelle du Québec, de caractériser et mieux comprendre la distribution spatiale des patrons des paysages à l'échelle du Québec et d'évaluer les différences des résultats avec la classification du MRNF. Pour ce faire, nous avons développé quelques idées et adapté quelques méthodes innovatrices (classification, régionalisation ou groupage) afin de dresser un portrait des paysages du Québec.

Ce mémoire est divisé en quatre chapitres. Le premier expose le contexte général duquel a émergé l'idée de cet ouvrage et il établit les notions théoriques qui ont servi à motiver celui-ci. Le deuxième chapitre présente une étude sous la forme d'un article scientifique qui traite de l'utilisation de centaines d'indices de patrons à des fins de régionalisation des paysages du Québec. Nous y avons identifié les paysages les plus

représentatifs en plus de générer des régions de paysages en fonction de leurs patrons et d'évaluer l'homogénéité spatiale de ces régions. Nous y dressons un portrait large des paysages québécois tel que vu du ciel. Le troisième chapitre contient une deuxième étude sous forme d'article scientifique qui démontre la différence des paysages observés, c'est-à-dire, qui sont caractérisés par les indices, par rapport à une classification écologique existante à l'échelle du Québec méridional. Nous y reprenons le même exercice général de régionalisation développé au chapitre deux, mais cette fois avec des paysages de formes et tailles variables. La variabilité et la complexité spatiales des patrons de paysage y sont aussi démontrées. Finalement, le quatrième et dernier chapitre présente les grandes lignes conclusives de ce manuscrit.

1. Contexte général de recherche du mémoire

Ce mémoire de maîtrise de recherche intègre plusieurs notions et touche à quelques sous-disciplines de la géographie : l'écologie du paysage, la biogéographie, les systèmes d'information géographique et la télédétection. Plus spécifiquement, le sujet général porte sur l'étude des patrons du couvert végétal du Québec tel qu'observé à partir d'images satellitaires et d'indices mesurant la composition et la configuration des parcelles (*patches*) des paysages. Dans les pages qui suivent, le contexte spécifique de la recherche, les concepts clés et les bases nécessaires à la juste compréhension des deux chapitres suivants seront élaborés.

1.1 Le Québec et ses paysages

La région d'intérêt principal de ce projet se définit par les frontières de la province de Québec, Canada. Plus spécifiquement, le premier article (chapitre 2) se concentre sur la quasi-totalité du Québec. Il s'agit en fait de la partie boisée de la province qui s'étend de sa frontière sud 45°N jusqu'à environ la latitude 58°N. En effet, la partie la plus septentrionale de la province qui correspond aux domaines bioclimatiques des toundras arctiques arbustives et herbacées a dû être exclue faute de données disponibles (explications à la section 1.7 et au chapitre 2). Dans le cas du deuxième article (chapitre 3), l'étude s'est effectuée sur le Québec méridional tel qu'entendu dans l'ouvrage de Robitaille et Saucier (1998), c'est-à-dire la section au sud du 53°N approximativement. En date de rédaction de ce mémoire, cette zone d'étude correspond à la limite du découpage des districts écologiques du MRNF (MRN 2000).

Pour ces deux chapitres, nous nous sommes référés et comparés aux « *sous-zones de végétation* » qui sont décrites par le MRNF comme des zones caractérisant « la physionomie de la végétation à des fins de successions » (MRNF 2010). Même si elles sont grossières, ces zones traduisent bien la variabilité des facteurs environnementaux qui influencent le couvert végétal. Elles transmettent l'étendue bioclimatique présente au Québec qui a un impact, autant sur la composition que sur la structure de la végétation

(Robitaille & Saucier 1998). Il s'agit donc d'un choix raisonnable vu l'échelle d'étude utilisée dans ce mémoire.

Que soit au Québec ou ailleurs, les patrons des paysages sont façonnés par l'interaction d'une multitude de facteurs biophysiques et de diverses sources de perturbations. Les conditions biotiques et abiotiques du terrain régulent la végétation dominante à l'échelle régionale, comme le gradient climatique, le régime hydrologique et l'exposition à la lumière (Robitaille & Saucier 1998, Bailey 2004). L'historique d'utilisations du sol (agriculture, coupes forestières contemporaines et anciennes) a aussi un impact sur le couvert actuel de la végétation puisqu'elle implique une reconfiguration des formes dans le paysage ainsi que sa composition en espèce (de Blois et al. 2001, Boucher et al. 2009). Ajoutons qu'en raison des interactions avec les processus biotiques (compétition, herbivorisme, etc.) qui surviennent dans le paysage, les patrons sont, de plus, rétroactivement modifiés (Turner et al. 2001, Turner 2005). Qu'ils soient naturels, comme avec les feux de forêt, ou anthropiques, par le biais des coupes forestières, les régimes de perturbations constituent aussi des facteurs majeurs dans la dynamique des paysages du Québec (Bergeron et al. 2001). Dans le contexte des paysages québécois, l'ensemble de ces facteurs peut se manifester singulièrement ou conjointement, à diverses intensités et dans une temporalité variable. En d'autres termes, ces facteurs autant naturels qu'anthropiques assurent une dynamique spatiotemporelle des patrons des paysages à l'échelle de la province (Urban et al. 1987).

Dépendamment de notre définition (ou conception) du mot paysage et de nos objectifs d'utilisation de ce mot, la grande diversité des paysages québécois peut s'apprécier de multiples façons. D'une part, une personne pourrait le faire en voyageant et en observant les différents patrons que l'on retrouve, par exemple, à travers les paysages agricoles des Basses-Terres du Saint-Laurent et ceux de la forêt boréale d'Abitibi. Ensuite, des outils tels que les globes virtuels nous offrent, depuis quelques années, l'opportunité de visualiser la surface terrestre et d'y interpréter les patrons (ex. *Google Earth* et *World*

Wind). Comme géographes, nous pouvons aussi nous référer à des ouvrages comme les cartes « *Végétation du Québec-Labrador : Formations et grands domaines* » (Richard 1995), et « *Zones de végétation et domaines bioclimatiques du Québec* » (MRNF 2010) ou encore l'ouvrage de Robitaille et Saucier (1998): « *Les paysages régionaux du Québec méridional* » qui nous présentent une synthèse cartographique et descriptive des paysages québécois. Ces références ne sont que quelques exemples d'outils qui témoignent de manière frappante de la grande diversité de la composition des paysages et de la nécessité d'en comprendre la distribution spatiale. Puis, nous pouvons nous tourner vers les images satellitaires, et les cartographies de la couverture du sol et d'utilisation du sol dans l'optique d'étudier le territoire tel que vu du ciel. Ces produits satellitaires permettent d'ailleurs un suivi précis, pixel par pixel, des conditions réelles du sol ainsi qu'un suivi temporel grandissant à mesure que les archives d'images sont stockées.

1.2 La notion de paysage

1.2.1 Définition

La notion de paysage est assez intuitive. Toutefois, sa définition demeure subjective, voire personnelle, et peut varier d'une personne à une autre, d'une discipline scientifique à une autre. Peu importe le point de vue, il devient nécessaire de définir un paysage et d'expliquer comment il sera utilisé dans ce mémoire afin d'éviter toutes fausses interprétations. Pour ce faire, nous avons utilisé des définitions évoquées par des auteurs qui ont contribué à l'émergence de l'écologie du paysage. Tout d'abord, Urban et al. (1987) a proposé de définir un paysage comme « [...] une mosaïque de formes hétérogènes, de types de végétation et d'utilisations du sol. ». Puis, Turner et al. (2001) a défini un paysage comme « [...] une aire spatialement hétérogène dans au moins un facteur d'intérêt ». En d'autres termes, nous pouvons interpréter et adapter ces définitions afin d'établir qu'un paysage se délimite par un cadre géographique connu, qu'il se caractérise par au moins deux composantes différentes (hétérogénéité) au sol et par une variation de configurations

spatiales des patrons (formes spatiales), et que cela résulte d'une dynamique spatiotemporelle des facteurs biotiques et abiotiques.

1.2.2 Unités d'étude – Le paysage

Depuis le début des années 1980 particulièrement, le paysage est utilisé comme unité géographique d'étude en écologie du paysage (Turner 2005) et cela a ouvert la possibilité d'étudier les processus écologiques de l'échelle du paysage (Trzcinski et al. 1999), régionale (Lawler & Edwards 2002), nationale (Wickham et al. 2008) jusqu'à l'échelle de la Terre (Riitters et al. 2000). Dans ce type d'études, un paysage est typiquement défini comme une image satellitaire ou une photographie aérienne qui ont été préalablement classifiées selon des critères pertinents à l'étude (par ex., une classification du couvert végétal) et qui correspondent à l'échelle des phénomènes ou processus étudiés (Turner et al. 2001). Les analyses environnementales ont ainsi passé de l'échelle du site, au paysage, jusqu'à l'échelle nationale. Conséquemment, les processus écologiques demeuraient à être étudiés. Dans le cas du paysage, l'intérêt s'est développé autour de l'hétérogénéité des patrons et la prémisse qu'il y a interaction avec les processus écologiques (O'Neill et al. 1988).

La figure 1 présente un exemple typique de paysage utilisé en écologie du paysage et il constitue l'unité d'étude du premier article (chapitre 2). Nous pouvons y distinguer des parcelles de différentes compositions du couvert végétal représentées par différentes couleurs et circonscrites à l'intérieur d'un cadre défini et régulier. Ce découpage est identique avec celui d'études similaires sur les paysages forestiers du Canada (Wulder et al. 2008a, Wulder et al. 2008c, Cardille et al. 2012). Il s'agit d'un cadre idéal pour l'étude des paysages à l'échelle nationale puisqu'il correspond aux délimitations des cartes topographiques gouvernementales bien connues des écologistes et géographes québécois.

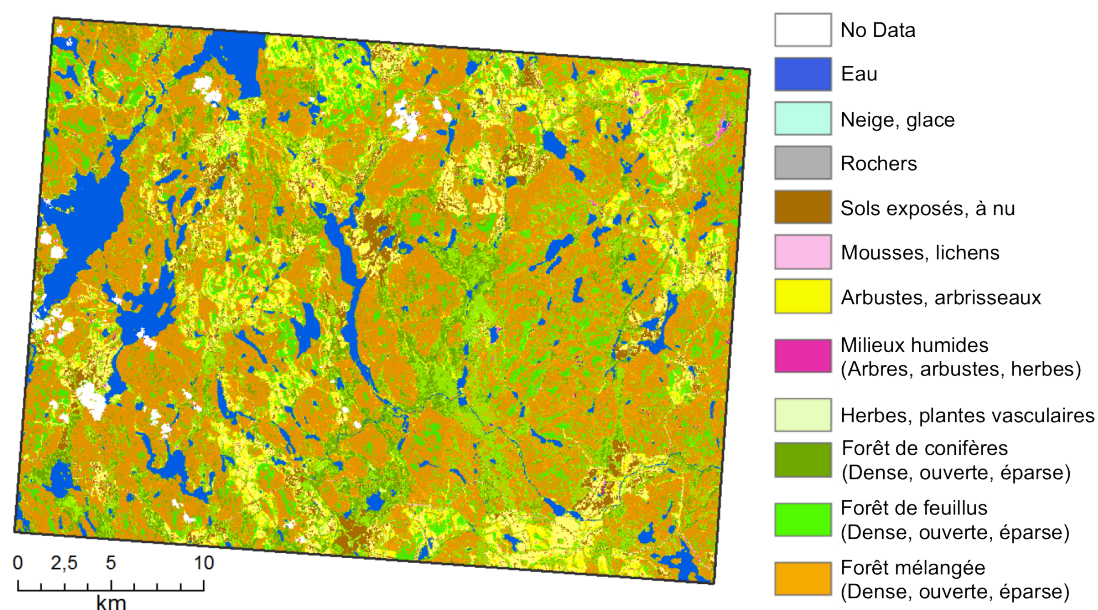


Figure 1. Exemple de paysage découpé selon les bordures du feuillet topographique 31P05 du SRNC à l'échelle 1:50000. La mosaïque est constituée de parcelles de différentes couleurs représentant différentes classes de couverture du sol. Sources des données : (RNC 2009, SAFORAH 2012).

Pour le deuxième article, nous avons opté pour un autre type de découpage des paysages, celui correspondant aux limites géographiques des districts écologiques (figure 2) délimités par le MRNF dans le cadre de la classification hiérarchique de classification écologique du Québec (Robitaille & Saucier 1998). D'un district écologique à un autre, le découpage varie en taille (superficie et périmètre) et en forme. Nous avons utilisé ce découpage pour le deuxième article (chapitre 3) car il s'agit d'une unité d'aménagement des ressources naturelles pour le Québec méridional, c'est-à-dire au sud du 53°N, et parce qu'il découpe les paysages de manière plus « naturelle » et cela, avec un maillage plus fin comparativement au découpage proposé à la figure 1. Ce découpage est conséquemment bien connu des scientifiques et des gestionnaires qui travaillent sur le Québec méridional. En date actuelle de rédaction de ce mémoire, le découpage des districts écologiques situés au nord de la sous-zone de la forêt boréale continue n'était pas encore disponible (MRN

2000). Pour cette raison, l'aire d'étude du chapitre 3 ne s'étend qu'aux limites actuelles couvertes par les districts écologiques dans le Québec méridional (Robitaille & Saucier 1996).

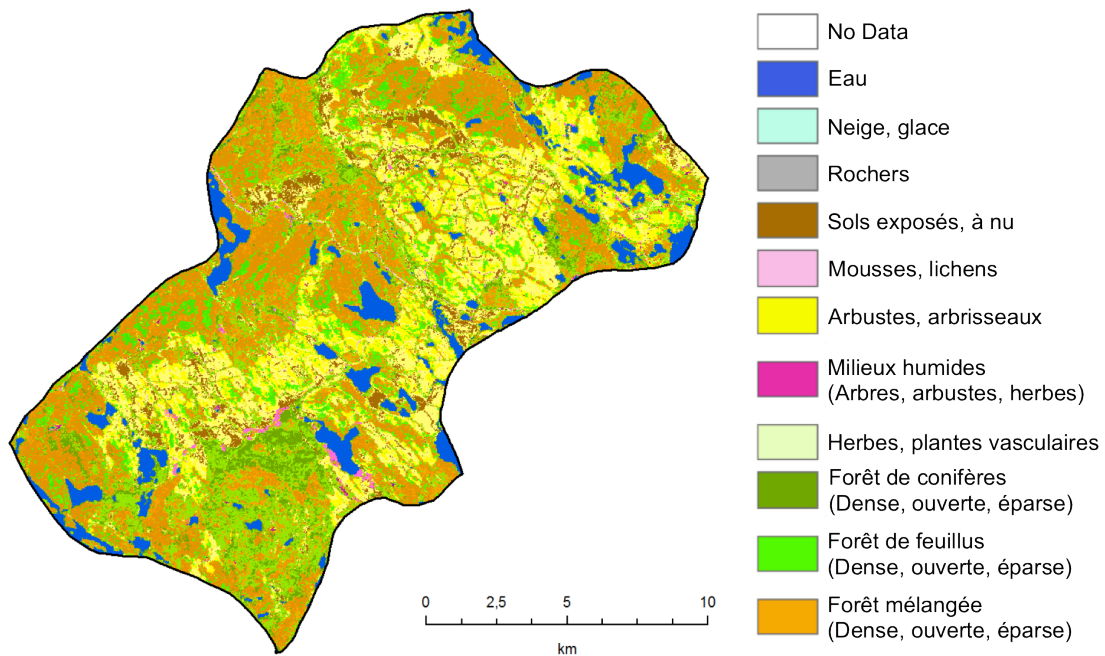


Figure 2. Exemple de paysage découpé selon le district écologique 31L009 du MRNF. 2489 de ces paysages ont été conservés pour le chapitre 3, chacun variant en taille et en forme. On y reconnaît la même classification (couleurs) que sur la figure 1. Sources des données : (MRNF 2005, SAFORAH 2012)

1.3 L'étude des patrons de paysages à petites échelles géographiques

Récemment, un intérêt marqué s'est développé pour l'évaluation des paysages à l'échelle nationale qui, le plus souvent, est réalisée à partir d'imageries satellitaires et de plus en plus complétées avec l'intégration de données auxiliaires (Riitters et al. 2002, Kupfer 2006, Wulder et al. 2008c, Cardille & Lambois 2010, Cardille et al. 2012). Les produits dérivés d'images satellitaires permettent de compresser les variations spectrales

(valeurs numériques de pixels) et temporelles pour représenter l'état actuel ou passé de la surface terrestre d'une région donnée (Lillesand et al. 2008, Cardille & Lambois 2010). Par exemple, ces produits nous informent sur les différents types de couvertures et d'utilisation du sol (Homer et al. 2004). Récemment, l'ajout de données auxiliaires à ces mégas produits a permis une plus grande flexibilité quant à la finalité d'utilisation de ceux-ci. Pour en donner quelques exemples, les cartographies de la fragmentation des forêts nationales ont été développées en utilisant des indices de patrons de paysages (Kupfer 2006, Wulder et al. 2008c), alors que d'autres auteurs ont poussé l'analyse afin de caractériser la distribution et la productivité des terres agricoles à l'échelle mondiale (Monfreda et al. 2008, Ramankutty et al. 2008).

L'intérêt pour les évaluations des paysages aux échelles nationales et continentales s'explique, en partie, par le fait que certaines recherches ont démontré l'appropriation sans cesse constante des ressources naturelles par l'Homme, particulièrement pour les ressources issues du produit de la photosynthèse (Vitousek et al. 1986, Cardille & Foley 2003, Foley et al. 2005, Ramankutty et al. 2008). Cette appropriation grandissante du territoire terrestre par l'Homme a pour conséquence de modifier les paysages autant en composition qu'en configuration, et ce, à toutes les échelles géographiques (Wulder et al. 2008c, Cardille & Lambois 2010, Wulder et al. 2011a, Cardille et al. 2012). En effet, l'ampleur marquante de la signature humaine sur les patrons de paysages a d'ailleurs été conclue dans plusieurs études scientifiques et elle se traduit non seulement par les coupes forestières, mais aussi via l'étalement urbain, la conversion de forêts en terres agricoles, etc. (Vitousek et al. 1997, Grimm et al. 2008, Cardille & Lambois 2010).

D'autre part, il a été démontré que la modification des paysages s'effectue actuellement à des échelles temporelles et spatiales jamais atteintes auparavant (Peters et al. 2008), ce qui a pour effet de modifier les processus écologiques. Les relations entre certains processus écologiques et les patrons des paysages à multiples échelles ont notamment été soulevées dans plusieurs publications (Holling 1992, King et al. 2004, Peters et al. 2004,

Peters et al. 2007, Peters et al. 2008). Une certaine part des activités humaines a donc un impact sur les processus écologiques à bien des échelles. Dès lors, il devient pertinent de générer des connaissances sur l'état des paysages à l'échelle nationale qui tiennent compte des patrons réels, et qui incluent les modifications d'origine anthropique et non uniquement de la caractérisation des conditions théoriques ou prédites par les facteurs naturels dominants (Ellis & Ramankutty 2008).

Bref, l'information géographique qui caractérise les patrons des paysages est d'une nécessité et d'une richesse indéniable autant pour les gestionnaires que pour les écologistes qui s'intéressent et qui souhaitent connaître l'état des ressources naturelles et du développement du territoire. Toutefois, il n'en demeure pas moins que ce genre d'exercice produit une masse incommensurable de données et qu'il peut devenir complexe de gérer ces produits (Cardille et al. 2005, Cardille & Lambois 2010, Cardille et al. 2012). Devant cette masse sans cesse croissante d'information, il devient pressant de développer des outils scientifiques efficaces afin de faciliter la compréhension et l'interprétation de ces données (Cardille et al. 2005, Mezard 2007).

1.4 Les indices de patrons des paysages

Depuis bientôt 30 ans, des indices de patrons de paysages (*landscape metrics*) ont été développés par les écologistes du paysage et sont utilisés pour quantifier les patrons (Cushman et al. 2008). Plus précisément, ces indices permettent de mesurer la composition (ou caractéristiques non spatiales) et la configuration spatiale des paysages à partir de cartes classifiées (O'Neill et al. 1988, Turner et al. 2001). Ces indices, lorsqu'utilisés judicieusement, servent à quantifier l'état des composantes et des fonctions écologiques des paysages observés, à étudier les changements spatiotemporels des patrons, puis à permettre la comparaison des paysages à travers un cadre géographique et statistique défini (Lausch & Herzog 2002, Cardille et al. 2005). À titre d'exemple, ces indices sont typiquement utilisés pour établir les relations possibles entre l'hétérogénéité des paysages et les

processus écologiques, et ce, à des échelles allant d'une simple parcelle de terrain jusqu'à un continent entier (Kie et al. 2002, Fahrig 2003, Turner 2005).

Les indices de patrons calculent les proportions, la diversité, la complexité des parcelles et la connectivité des classes d'occupation du paysage. Les indices spatiaux calculent des valeurs traduisant la contagion, le périmètre et l'aire des parcelles, le poids des dimensions de parcelles, etc. (Turner et al. 2001, Wulder et al. 2008c). Puis, il y a des indices spatiaux qui caractérisent la forme des parcelles dans les paysages. Ils servent principalement à quantifier la façon dont sont configurées les parcelles (bordures totales, indice des formes de paysage, etc.) (De Clercq et al. 2006). Ensuite, parmi les indices non spatiaux, il existe également ceux qui calculent la distribution de la superficie ou la composition des paysages (nombre de parcelles, la superficie totale des parcelles d'une catégorie, la taille moyenne des parcelles, etc.).

Il existe une panoplie d'indices de patrons, des centaines en fait, mais il subsiste des difficultés notoires quant à leur utilisation. Les difficultés les plus communes ont rapport à la corrélation entre les indices (Fortin et al. 2003), leur choix en fonction du type de question, l'interprétation des valeurs et la gestion du nombre de données générées (Riitters et al. 1995, Li & Wu 2004, Cushman et al. 2008). Une partie des raisons qui expliquent les difficultés d'utilisation vient du fait que ces indicateurs sont dépendants de l'échelle et de la résolution spatiales, des dimensions des paysages et de la classification, que certains sont des indexes, alors que d'autres sont des proportions ou même des nombres sans unités, et j'en passe, les exemples étant très nombreux (Wu et al. 2000, Turner et al. 2001, Li & Wu 2004). Il devient alors difficile de saisir tout le sens des résultats, de comparer plusieurs indices et donc, de comparer les paysages. En réponse à cela, plusieurs chercheurs se sont penchés sur la question d'identification d'un ensemble restreint et facilement compréhensible d'indices de patrons sans toutefois en arriver à des réponses universelles pour l'étude de paysages diversifiés à l'échelle nationale (Riitters et al. 1995, Cain et al. 1997, De Clercq et al. 2006, Cushman et al. 2008).

Malgré la capacité qu'ont ces indices à distinguer les différents patrons (Herzog & Lausch 2001), peu de gens les ont utilisés afin de régionaliser les différents types de paysages à grande échelle (O'Neill et al. 1988, Riitters et al. 1995, Long et al. 2010). Effectivement, les difficultés associées à ces indices peuvent freiner leur utilisation à grande échelle. Néanmoins, des méthodes d'analyse multifactorielle, particulièrement l'analyse en composantes principales, se sont révélées capables d'éliminer la redondance et la corrélation statistique des valeurs (Riitters et al. 1995, Cain et al. 1997, Cardille & Lambois 2010, Cardille et al. 2012). Dès lors, il devient possible d'effectuer ces transformations avec confiance afin de caractériser les patrons des paysages à grande échelle.

Typiquement, les études qui ont utilisé les indices à des fins de régionalisation cartographique n'ont utilisé que quelques indices et cela, à des fins très spécifiques. L'outil « *Metaland* » développé par Cardille et al. (2005) en est un bon exemple, car il s'agit d'un outil internet qui permet l'analyse spatiale et statistique des paysages et où l'on peut sélectionner et visualiser la répartition spatiale de valeurs d'indices pour les États-Unis d'Amérique. De plus, l'évaluation des forêts aux échelles nationales s'effectue aussi en utilisant une combinaison de quelques indices de patrons, particulièrement la fragmentation des paysages forestiers (Kupfer 2006, Wulder et al. 2008c). Pour leur part, Long et al. (2010) ont régionalisé les différents types de fragmentation des paysages d'une région forestière de 5,5 millions d'hectares en Colombie-Britannique au Canada. Ils ont utilisé un total de cinq indices dans le but de grouper les patrons de paysages similaires. Le comportement à grande échelle des indices dans une optique d'étude de la distribution spatiale des patrons demeure donc un sujet de recherche à explorer.

Dans ce mémoire, nous avons élargi le spectre d'application des indices de patrons tel qu'approché dans les recherches présentées ci-dessus. Nous avons calculé et traité un jeu constitué de centaines d'indices sur un territoire vaste représenté par la Province de Québec afin d'y caractériser les différents types de patrons (chapitres 2 et 3). Par le biais de

différentes approches, nous avons aussi démontré la capacité des indices à distinguer les patrons dans le but de prédire des classifications écologiques existantes (chapitre 3).

1.5 La classification et la régionalisation des données à petite échelle géographique

1.5.1 Synthétiser le contenu d'information pour mieux comprendre

Compte tenu des quantités imposantes de données qui sont générées lors des études d'envergures nationales et continentales, il devient rapidement nécessaire de développer des outils d'analyse et d'organisation des données afin de pouvoir gérer efficacement ce flux astronomique de données, et de mieux les comprendre et les interpréter (Mezard 2007). Pour ce faire, les chercheurs ont communément recours à des méthodes d'analyse quantitative multifactorielle et à des algorithmes de groupage (*clustering*). Les algorithmes de groupage réunissent les éléments similaires pour former des ensembles homogènes selon, généralement, un nombre de groupes préalablement défini par l'utilisateur (Hargrove & Hoffman 2004). Les groupes formés permettent d'avoir une vue d'ensemble des différentes caractéristiques à l'intérieur d'un jeu de données. Plus simples à interpréter, car ils sont en nombre réduit, les groupes nous permettent, par exemple, d'extraire les statistiques des attributs typiques moyens propres à chaque groupe. En complément, certaines méthodes de groupage poussent les capacités d'analyse plus loin alors qu'elles permettent d'identifier le ou les éléments représentatifs de chaque groupe. Les éléments représentatifs des groupes peuvent être calculés de différentes façons et, conséquemment, ceux-ci sont nommés différemment, notamment les « *centroids* » (Hargrove & Hoffman 1999), les « *medoids* » (Van der Laan et al. 2003, Long et al. 2010), ou même les « *exemplars* » (Frey & Dueck 2007). Même s'ils sont les éléments représentatifs, il faut bien les différencier et ne pas les confondre comme étant similaires. Ces éléments représentatifs sont calculés et obtenus de manières différentes.

Les éléments représentatifs ont l'avantage de synthétiser la variabilité intrinsèque des variables des groupes qu'ils représentent et sont plus faciles à interpréter que les tableaux statistiques traditionnels (Cardille & Lambois 2010, Long et al. 2010, Cardille et al. 2012). D'ailleurs, l'identification de paysages représentatifs dans le but de peindre les patrons de paysages typiques des vastes étendues de territoire est un objectif souhaitable selon plusieurs auteurs (Hargrove & Hoffman 2004, Keller et al. 2008, Peters et al. 2008) puisque cela ferait un pas de plus dans l'établissement d'un réseau de paysages représentatifs à l'échelle nationale. Ces sites représentatifs pourraient permettre, par exemple, d'évaluer la représentativité des sites actuels de protection environnementale par opposition aux conditions réelles (ex. les parcs nationaux) (Hargrove & Hoffman 2004, Cardille et al. 2012), et d'établir un réseau de sites représentant les conditions environnementales des paysages (Keller et al. 2008).

Les développements informatiques récents ont ouvert la porte aux scientifiques pour l'utilisation de ces algorithmes à des fins de régionalisation des paysages. Dans notre domaine d'intérêt, cela se traduit par la classification de domaines environnementaux (Coops et al. 2009), la délimitation d'écorégions (Hargrove & Hoffman 1999) et de régions de patrons spatiaux (Long et al. 2010), ou encore par la régionalisation des paysages selon de multiples caractéristiques biophysiques (Mackey et al. 2008). En effet, plusieurs facteurs sont en cause dans l'élaboration de ces mégas produits qui intègrent des données auxiliaires, par exemple : la capacité de traitement croissante des postes informatiques personnels, le développement des systèmes d'information géographique (SIG) et la disponibilité d'archives d'images des satellites d'observation de la Terre comme celle du projet Landsat disponible depuis les années 1970 (Wulder et al. 2008b).

L'élaboration de ce type de classifications environnementales devient intéressante, car elles sont quantitatives, reproductibles et flexibles par opposition aux méthodes classiques de classifications écologiques plus subjectives et effectuées à l'aide d'experts (Hargrove & Hoffman 1999). Les méthodes plus communes exigent en effet beaucoup de

connaissances du terrain ainsi que d'excellentes habiletés à pouvoir tracer, par exemple, des zones contiguës les plus homogènes possible. Dans ces cas, les ressources nécessaires sont énormes et il est souvent impossible de régénérer automatiquement des cartes advenant le cas d'un changement dans les variables de base ou d'une mise à jour. À l'inverse, les approches de régionalisation quantitative sont adaptables et peuvent être reproduites au besoin, et ce, avec les variables souhaitées. Cette flexibilité est possible, car les groupes (ou régions) sont formés dans l'espace « donnée » (c.-à-d. dans les dimensions statistiques) plutôt que forcé à l'intérieur de polygones contigus dans l'espace géographique. Bien que ces approches soient beaucoup plus objectives, il n'en demeure pas moins que le choix des données à traiter doit être subjectivement effectué (Hargrove & Hoffman 2004).

Dans les chapitres 2 et 3, nous approfondirons l'utilisation à petite échelle géographique des indices de patrons à des fins de régionalisation des patrons des paysages. Comme mentionnés auparavant dans ce chapitre, les problèmes associés à l'utilisation des indices de patrons ont limité leur usage comme attributs environnementaux à ces échelles d'études. Néanmoins, O'Neill et al. (1988) ainsi que Wickham et Norton (1994) ont été des précurseurs en cherchant à démontrer la distribution des différents types de patrons de paysages. Nous avons donc entrepris d'élargir le spectre de ces recherches en poussant l'analyse à l'utilisation de centaines d'indices et de régionaliser les différents types de paysages du Québec en fonction de leurs patrons.

1.5.2 Affinity Propagation

L'algorithme « Affinity Propagation » (AP) développé par Frey et Dueck (2007) est conçu pour former efficacement des groupes parmi d'immenses quantités de données et, simultanément, identifier l'élément le plus représentatif de chaque groupe. Selon les créateurs, l'algorithme a la capacité d'effectuer cette opération 100 fois plus rapidement en une seule exécution que la meilleure des 10 000 exécutions de l'algorithme *k-centres*, et ce, avec plus d'exactitude. Les auteurs ont d'ailleurs démontré la puissance de cet algorithme

en résumant le texte de leur article en quatre phrases grâce à celui-ci. Puis, l'algorithme a été adapté pour traiter des données génétiques (Frey & Dueck 2007, Leone et al. 2007), pour résumer le contenu de vidéos (Xie et al. 2008) et, plus près de la géographie, AP a servi à classifier des images (Dueck et al. 2007), et effectuer la sélection de bandes d'images hyperspectrales (Qian et al. 2009). Bref, l'utilisation de cet algorithme est justifiée pour grouper de manière efficace un jeu de données complexes comme le sont les indices de patrons.

Par ailleurs, Cardille et al. (2010, 2012) ont adapté l'algorithme de Frey et Dueck (2007) afin d'identifier les paysages les plus représentatifs des États-Unis et du Canada forestier. Par contre, ils se sont limités à utiliser AP pour sa capacité à déterminer les exemplars (paysages représentatifs). Conséquemment, ils ont laissé de la place pour la recherche dans ce domaine, car ils n'ont pas exploité les résultats de groupage des paysages avec l'utilisation d'indices de patrons. Néanmoins, leurs résultats sont conclusifs et permettent d'affirmer que les indices de patrons utilisés dans un contexte de groupage peuvent se révéler utiles et remplis de potentiel. Ainsi, les auteurs ont pu synthétiser la grande étendue et la variabilité des patrons des paysages. Ces recherches ont de plus montré les applications potentielles d'une telle approche, par exemple, pour la gestion des ressources naturelles, en proposant une méthode d'évaluation de la représentativité du réseau de parcs nationaux vis-à-vis les patrons des paysages forestiers.

1.5.3 Les travaux de Long et al. (2010)

Long et al. (2010) ont, pour leur part, utilisé les capacités de groupage de l'algorithme CLARA (*Clustering for LArge Applications*) afin de régionaliser les patrons des paysages à partir, eux aussi, d'indices de patrons. Ils ont également identifié les paysages représentatifs de chacun des groupes formés, appelés *medoids*. Ainsi, leurs résultats ont démontré les différents types de patrons de fragmentation des paysages forestiers d'une région couvrant 5,5 millions d'hectares en Colombie-Britannique. Bien que convaincants, en vertu de leurs objectifs, les indices utilisés et les résultats ne se

concentrent que sur la proportion de la forêt par une classification dichotomique forêt/non-forêt. Bien qu'ils répondent aux objectifs de l'étude, les résultats « cachent », selon nous, une part de la réalité avec les indices calculés à partir de la classification dichotomique forêt/non-forêt plutôt que d'inclure les patrons d'autres types de végétation ou de classes.

Il y a donc une partie des deux chapitres suivants (2 et 3) qui récupère l'idée de régionalisation des patrons présentée par Long et al. (2010). Nos contributions viennent du fait que nous avons exploité une somme beaucoup plus grande d'indices de patrons sur un territoire de superficie beaucoup plus grande, que nous avons caractérisé un plus grand nombre de paysages définis par une résolution et une classification différentes, en plus d'utiliser AP pour y arriver.

1.5.4 Les arbres de classification

Les arbres de classification sont des outils statistiques simples à comprendre qui permettent de produire des classificateurs précis ou d'extraire la structure prédictive d'une classification existante (Breiman et al. 1984). Puisqu'ils sont non-paramétriques, les arbres de classification ont la capacité de traiter des jeux de données déséquilibrées et complexes (De'ath & Fabricius 2000). Avec l'augmentation des bases de données à échelles nationale et continentale, et des multiples sources possibles, il devient de plus en plus intéressant d'explorer ce type d'outils statistiques capables de les combiner efficacement afin de retenir les informations pertinentes (Cardille & Clayton 2007).

Les arbres de classification fonctionnent à partir de données classifiées (ou catégoriques) alors que les arbres de régression se basent sur des données numériques continues (De'ath & Fabricius 2000). Dans ce mémoire, nous avons utilisé les arbres de classification. Hansen et al. (2000) proposent cette description des arbres de classification : « Les arbres sont des classificateurs non paramétriques et hiérarchiques qui prédisent l'appartenance des classes en divisant de manière récursive un jeu de données en groupes homogènes ». De plus, ils permettent d'identifier de manière supervisée les variables

explicatives pour des réponses catégoriques ou des classes (Questier et al. 2005). Cette méthode statistique a été utilisée par plusieurs auteurs et pour une multitude d'objectifs comme d'identifier les facteurs biophysiques de présence-absence de certains taxons de coraux (De'ath & Fabricius 2000); dans le domaine de la santé, ils ont servi comme outil décisionnel de diagnostics (Breiman et al. 1984); en télédétection, à classifier des couvertures du sol (Hansen et al. 2000) à partir de données auxiliaires (Cardille et al. 2002, Cardille & Foley 2003); en biologie marine, à modéliser la distribution spatiale de mammifères marins en fonction de variables environnementales et de données de proies (Friedlaender et al. 2006); et en écologie du paysage, pour identifier des indices de patrons pertinents à être utilisés (Lausch & Herzog 2002). En écologie du paysage, les arbres de classification ont déjà servi à sélectionner un jeu d'indices de patrons pertinents (Lausch & Herzog 2002). Ces études viennent appuyer l'utilisation des arbres de classification à traiter efficacement des données aussi hétérogènes et complexes que les indices de patrons.

Dans le chapitre 3, nous nous sommes basés sur une catégorisation existante des paysages, les sous-zones de végétation du Québec méridional du MRNF, et, par conséquent, avons utilisé les arbres de classification afin d'extraire la partie prédictive des indices de patrons à des fins de classifications des paysages. De plus, ces arbres de classification ont servi de classificateurs pour évaluer la capacité de cette méthode à catégoriser des paysages en se basant sur des indices de patrons. Comme cette méthode classifie en groupes homogènes et cela implique nécessairement une simplification de la réalité. Cette généralisation engendre nécessairement des erreurs de classification et des divergences. Nous avons souhaité connaître l'emplacement géographique des divergences et évaluer la capacité des indices de patrons à petite échelle géographique.

1.6 Quantification et spatialisation des divergences

Dans l'exercice de grouper des paysages en régions, il existe certains avantages à utiliser des méthodes quantitatives par rapport aux méthodes dites « basées-experts »

(Hargrove & Hoffman 1999). Parmi ceux-ci, il est possible de calculer l'homogénéité ou l'hétérogénéité des groupes (Kelling et al. 2009). Conséquemment, il est aussi possible de représenter la variabilité de la précision – ou de la qualité – d'un phénomène dans l'espace géographique (Serra et al. 2011). Bien qu'il soit inutile de prouver la pertinence et l'utilité que jouent les systèmes d'information géographique dans la science d'aujourd'hui, rares sont les cas où ceux-ci sont mis à profit afin de quantifier l'incertitude ou les divergences des résultats et d'en faire une représentation spatiale (Zhang & Goodchild 2002). Or dans un contexte de recherche où des centaines de variables sont exploitées afin d'y voir émerger des résultats (Kelling et al. 2009), il devient important de connaître la qualité des modèles cartographiques résultants, à savoir, non seulement quelle est la précision totale, par exemple d'une classification écologique, mais aussi quelles sont les erreurs et où sont-elles situées (Mackey et al. 1988). Le cas contraire pourrait éventuellement mener à une mésinterprétation des résultats, par exemple.

Les frontières aux zones de transition (ex. entre deux types de patrons ou d'écosystèmes) peuvent être graduelles ou aigües, mais sont le plus souvent délimitées par des entités discrètes spatialement définies (lignes ou polygones) et qui ne traduisent pas ce phénomène (Fortin & Edwards 2001). La variabilité spatiale du phénomène transitoire des patrons n'est pas représentée et donc, les résultats peuvent induire en erreur. Hargrove et Hoffman (2004) ont proposé de produire des « clôtures spatiales » (*spatial fences*) qui représentent les changements environnementaux instantanés, c'est-à-dire de délimiter les changements qui apparaissent brusquement dans l'espace géographique. Ces clôtures spatiales n'apparaissent qu'aux endroits où il y a un minimum de transition aigüe de part et d'autre laissant parfois ouverts les endroits où ces changements s'effectuent graduellement dans l'espace géographique (McIntire 2004). Conséquemment, ces clôtures apportent une autre dimension d'information géographique aux délimitations écologiques communément formées par des polygones fermés faisant abstraction de la variabilité naturelle des conditions de terrain. Dans chacun des chapitres 2 et 3, nous avons récupéré cette idée de clôture spatiale afin de tracer les « frontières persistantes » des différents types de patrons

de paysages, c'est-à-dire d'identifier les endroits où, systématiquement, il y a une délimitation importante entre différents groupes de paysages différents.

Les erreurs de modélisation cartographique ou de classification – aussi inévitables soient-elles – sont généralement considérées comme gênantes ou pires, à éviter. Mais la complexité des phénomènes environnementaux, d'une part, et l'utilisation de données géographiques, d'autre part, impliquent nécessairement des incertitudes (Zhang & Goodchild 2002). Ces incertitudes sont pour le plus souvent exprimées par des données statistiques (probabilités de mauvaise classification, racine carrée de l'erreur moyenne, etc.) n'informent que très peu sur le phénomène spatial réel (Lillesand et al. 2008). Au contraire, selon Zhang et Goodchild (2002), les erreurs spatiales pourraient plutôt être vues comme des informations déterminantes – voir, une information géographique à part entière. Par exemple, certaines régions comme les zones de transitions de végétation (c.-à-d. des écotones) sont des zones sensibles aux changements environnementaux et par conséquent, sont propices à démontrer des divergences par rapport aux conditions prédites (Fortin & Edwards 2001). Il y a là, matière à investigation.

Dans les deux chapitres suivants (2 et 3), nous avons évalué la variabilité et l'homogénéité spatiale de nos résultats, et nous avons tiré avantage des multiples modèles obtenus afin de mettre à la lumière le maximum d'information contenue dans nos résultats. D'une part, nous avons apporté une autre dimension aux résultats de régionalisation des patrons du Québec en y intégrant la centralité, ou la mesure de représentativité, des paysages représentatifs par rapport aux paysages constituant leur groupe. La centralité des paysages représentatifs peut se traduire par une représentation spatiale de l'homogénéité intragroupe dans l'espace géographique. Cette approche de représentation des résultats a été adaptée des recherches de Hargrove et Hoffman (1999, 2004) ainsi que, dans une moindre mesure, de l'étude de Serra et al. (2011). Comme mentionné ci-haut, nous avons également identifié les frontières persistantes qui distinguent les changements de patrons les plus drastiques à l'échelle du Québec (chapitre 2) et puis du Québec méridional (chapitre 3).

Finalement, nous avons tiré avantage des multiples divergences de classification des paysages (chapitre 3) afin d'identifier l'emplacement géographique des incompatibilités et d'analyser les endroits où le MRNF, par exemple, n'est pas cohérent avec les conditions réelles des patrons de paysages.

1.7 Les données de base pour cette recherche

1.7.1 La couverture du sol de l'OTDD

Comme données de base dans les chapitres 2 et 3, nous avons opté pour une classification de la couverture du sol en format matriciel, soit une mosaïque d'images satellitaires classifiées selon la couverture végétale du sol des années 2000. Des équipes de recherche canadiennes ont produit et rendu disponible gratuitement ces données pour le bien de la recherche sur les forêts canadiennes. Depuis quelques années, le réseau SAFORAH (www.saforah.org) distribue publiquement les données d'Observation de la Terre pour le développement durable des forêts (OTDD). Plus précisément, il s'agit d'une couverture terrestre complète du Canada forestier à partir de l'imagerie Landsat-7 ayant résolution spatiale de 25 mètres et catégorisée en 23 classes de couverture du sol (Wulder et al. 2008a). Près de 80 % du pays a été couvert et catégorisé en classes de couverture végétale, là où la végétation dominante est principalement arborescente. La partie restante au nord de la toundra arbustive a été classifiée avec l'imagerie Landsat, mais dans un contexte et une philosophie de classification différente et avec des objectifs différents (Olthof et al. 2009). Pour des raisons d'incompatibilité entre les deux produits, nous avons écarté cette dernière classification de la présente recherche.

Les données matricielles de l'OTDD ont été sélectionnées pour plusieurs raisons. En date de rédaction de ce mémoire, le produit de L'OTDD représentait toujours la seule source d'information sur la couverture terrestre pertinente et récente disponible pour notre aire d'étude. Ce jeu de données a été conçu pour étudier les paysages forestiers du Canada (Wulder et al. 2008a, Wulder et al. 2008c). D'autres auteurs ont déjà utilisé avec succès le

même type de données pour identifier les paysages représentatifs des États-Unis (Cardille & Lambois 2010) et du Canada forestier (Cardille et al. 2012), et pour caractériser la fragmentation des paysages forestiers du Canada (Wulder et al. 2008c). De plus, les données satellitaires de type Landsat sont reconnues comme étant bien adaptées pour l'étude des couvertures nationales et leur résolution moyenne permettent de les utiliser à plus grandes échelles, soit celle du paysage, de même que régionales (Lillesand et al. 2008).

Néanmoins, des transformations ont dû être effectuées sur les données originales de l'OTDD afin de circonscrire celles-ci à l'intérieur de la zone d'étude, de pouvoir travailler sur une mosaïque homogène unique (et non plusieurs qui s'entrecoupent) et de travailler avec des classes de couverture du sol pertinentes. Suite au téléchargement des tuiles couvrant grossièrement la zone d'étude (débordements), nous avons généré une mosaïque unique avec le logiciel ArcGIS Desktop 10 (ESRI 2011). Ensuite, nous avons extrait la partie de la mosaïque nouvellement créée située à l'intérieur des limites continentales de la Province du Québec (opération « extract by mask » dans ArcGIS). Dès lors, nous avons la mosaïque couvrant la zone d'étude. À des fins ultérieures dans la méthodologie, nous avons reclassé toutes les classes impertinentes sous une seule classe afin de générer un seul et unique arrière-plan (nuage, ombre, non classifiée, no data → no data). Conséquemment, les étapes subséquentes présentées dans les chapitres 2 et 3 sous-entendent que la classification utilisée était catégorisée selon cette reclassification.

1.7.2 Limites d'utilisation des données de l'OTDD

Plusieurs limites sont à prendre en considération avec la classification de l'OTDD comme avec toutes autres données de ce type. Tout d'abord, il s'agit de données secondaires, c'est-à-dire qu'elles ont été préparées et produites par un tiers parti. Conséquemment, nous nous devons de reconnaître certains risques qui pourraient être associés, à la qualité ou à la précision de la classification, ou encore à la représentativité de celle-ci pour le contexte québécois. Il n'en demeure pas moins que cette classification d'envergure a été produite avec des standards scientifiques élevés et qu'elle présente des

pourcentages d'exactitudes des classifications acceptables estimés à 77 % (Wulder et al. 2008a). Une autre limite des données de l'OTDD réside dans le fait que ses classes de couvertures du sol ne caractérisent que la couverture végétale et font abstraction de l'utilisation humaine du sol. Cela apporte nécessairement un biais quant au réalisme de certains patrons de paysage observé, notamment dans la partie sud du Québec. Par exemple, les champs agricoles ne pourront pas être catégorisés de la sorte, mais plutôt comme des parcelles d'arbrisseaux ou de sol à nu où la culture n'en est qu'à ses débuts de culture annuelle. Conséquemment, la signature humaine telle qu'observée par Cardille et Lambois (2010) pour les paysages des États-Unis ne pourra être détectée avec la même assurance.

Une autre limite est que les images qui ont servi à la classification de l'OTDD datent des années 2000 (entre 1999 et 2002) et que plus de dix ans se sont écoulés depuis leur capture (Wulder et al. 2008a). Par conséquent, les résultats et conclusions tirées dans ce mémoire de recherche ne représentent que de cette époque et ainsi, toute interprétation contemporaine des résultats serait risquée. Toutefois, le renouvellement des données de l'OTDD pour l'année décennale 2010 ne peut pas s'effectuer, car le capteur Landsat-7 est défectueux depuis 2003 (Wulder et al. 2011b). Par contre, le renouvellement du projet d'observation de la Terre Landsat (appelé « LDCM » pour Landsat Data Continuity Mission), qui tarde à démarrer, offrira aux scientifiques un renouveau de données de cette résolution et de cette étendue spatiales comparable. Son lancement est prévu pour décembre 2012 (NASA 2012a).

1.8 Objectifs généraux du mémoire

Dans ce projet de maîtrise, les paysages québécois ont été étudiés afin de mieux comprendre la diversité et la distribution spatiale des différents patrons qui caractérisent la province. Nous avons entre autres exploré la capacité qu'ont les indices de patrons à différencier quantitativement les paysages à l'échelle du Québec, localiser l'emplacement des plus grandes démarcations entre les différents types de patrons, et évaluer les résultats par rapport aux catégorisations existantes. Pour ce faire, nous avons adapté quelques

approches prolifiques proposées par des auteurs dans le domaine et nous en avons développé d'autres afin de voir émerger un maximum d'information géographique pertinente à l'étude des paysages du Québec. Voici les objectifs spécifiques présentés dans les chapitres suivants :

- Régionaliser les différents types de paysages québécois en se basant sur des mesures de patrons, c'est-à-dire de les grouper en fonction de la similitude des patrons puis de les cartographier (chapitres 2 et 3).
- Synthétiser la diversité des paysages en identifiant les paysages les plus représentatifs de chacune des régions formées (chapitre 2).
- Évaluer la centralité spatiale des paysages représentatifs, c'est-à-dire vérifier si la similitude des patrons est perceptible avec l'utilisation d'indices de patrons, et ce, non seulement dans l'espace « donné », mais aussi dans l'espace géographique (chapitre 2).
- Localiser les frontières les plus marquées entre les différents types de patrons de paysages et en représenter la variabilité spatiale (chapitres 2 et 3).
- Évaluer la capacité d'un nombre réduit d'indices afin de reproduire une classification existante et d'extraire les variables explicatives (indices) qui permettent cette classification (chapitre 3).
- Identifier les divergences entre les classifications écologiques existantes et les patrons réels observés; c'est-à-dire de localiser les paysages qui, dans les sous-zones de végétation du Québec méridional, ne correspondent pas aux paysages observés (chapitre 3).

Les motivations derrière ces objectifs de recherche et les contributions potentielles sont diverses. La plupart des éléments motivants et innovateurs de ce mémoire ont été présentés dans les sections précédentes et seront développés dans les chapitres qui suivent. Les chapitres suivants seront soumis à des revues scientifiques et sont écrits dans un format conséquent. Alors, chacun d'eux présente une problématique et des objectifs qui

correspondent à des contributions respectives. Par contre, nous avons tenté de peindre les patrons des paysages québécois de manière peu habituelle, afin de produire des résultats jamais présentés auparavant. Tout au long des chapitres suivants, nous avons fait des comparatifs avec le système hiérarchique de classification écologique du MRNF dans l'optique de bien montrer que l'information tirée d'une base de données sur les patrons des paysages permet de discerner les patrons réels ou observés plutôt que des patrons prédits basés sur des conditions théoriques.

2. Premier article

Présentation du premier article

Le chapitre deux présente un premier article scientifique dans lequel nous avons exploré l'utilisation à l'échelle du Québec des indices de patrons en tant qu'attribut biogéographique pour le groupement des paysages à l'échelle du Québec. Grâce à ces mesures quantitatives des patrons, nous avons portraituré les paysages représentatifs de la province, et nous avons approfondi nos connaissances sur les patrons réels observables à l'échelle de la province. Les conclusions de l'article mettent en lumière le potentiel des indices de patrons dans les futures classifications écologiques d'envergure nationale, et nous discutons du caractère innovateur des résultats qui contrastent avec les cartographies typiques.

Accord du coauteur

Ce premier article a conjointement été écrit par Kevin Partington, premier auteur, et par son directeur de maîtrise Jeffrey A. Cardille. L'article est soumis à la revue scientifique « *Canadian Journal of Forest Research* ».

À titre de coauteur, j'autorise mon étudiant Kevin Partington à publier l'article « *Landscape patterns of Quebec : representative landscapes, spatial clusters, and fences* » dans son mémoire de maîtrise.

Jeffrey A. Cardille

Landscape patterns of Quebec: representative landscapes, spatial clusters, and fences

By Kevin Partington & Jeffrey A. Cardille

2.1 Introduction

For nearly 30 years, landscape ecologists have used landscape metrics to study relationships between ecological processes and the variation in composition and configuration of landscapes (O'Neill et al. 1988, Gustafson 1998, Turner 2005). Still, the use of landscape metrics targeting the quantification of patterns for a better understanding of landscape pattern behaviour across very large areas remains rare (O'Neill et al. 1988, Wickham & Norton 1994, Riitters et al. 1995, Riitters et al. 2000). Even when considering their capacity to differentiate patterns, landscape metrics have not been commonly used as attributes for large-scale ecological mapping (Herzog & Lausch 2001, Long et al. 2010).

With the continuing collection and interpretation of regional- and continental-scale data, maps displaying land cover and/or land use across large areas have become increasingly accessible and adapted for diverse needs. Typically, they use data taken directly from satellites or other remote-sensing platforms (e.g., Landsat, Moderate Resolution Imaging Spectroradiometer -MODIS mapping products, etc). With the increase in the longevity and number of these systems, these now include basic land-cover maps (Loveland et al. 2000, Latifovic et al. 2004) biophysical stratifications (Mackey et al. 2008, Coops et al. 2009), and eco-regions delineation (Hargrove & Hoffman 1999, Prince & Steininger 1999). A relatively new and rapidly growing trend in mapping geographical patterns across large scales is the integration of ancillary data sets into wide-scale satellite-based products. This has recently included blends of satellite- and census-derived agricultural data (Girvetz et al. 2008, Monfreda et al. 2008, Ramankutty et al. 2008), forest fragmentation regimes mapped from a combination of landscape metrics, selected natural and human fragmentation drivers (Wulder et al. 2011a), and the delineation and description of anthropogenically-driven regions (Ellis & Ramankutty 2008).

Nearly all of these maps have focused on measures that are rooted in, or exclusively linked to, the proportion of a phenomenon in a given area. To date, little attention has been paid to the question of how patterns are spatially distributed across large areas, at a provincial or national scale for instance. Studies are few and typically targeted to a given metric or two mapped across large areas for a specific purpose or research question – mainly forest fragmentation (Riitters et al. 2002, Wade et al. 2003, Wulder et al. 2008c, Wulder et al. 2011a). In a similar vein, the *Metaland* tool allows users to map patterns of a given metric for large, pre-defined data sets across large areas (Cardille et al. 2005). Thus far, limited consideration has been paid to the question of the total content of information about spatial patterns across large areas, often because of the challenges of interpreting such large datasets over large spatial extents (Long et al. 2010).

To better understand the information in large geographic data sets, researchers commonly group similar items together using a variety of quantitative clustering methods (Hargrove & Hoffman 1999). This allows the variability in the data to be summarized, and because of the geographic context, group membership can be mapped (Hargrove & Hoffman 2004). Recently, some authors have proposed to identify the best representatives of each group. Such representatives have the ability to capture and summarize the data into a small set of items that may often be easier to understand. These can take various forms depending on the technique involved; some have used ‘centroids’ (Hargrove & Hoffman 1999, 2004), ‘medoids’ (Van der Laan et al. 2003, Long et al. 2010), and ‘exemplars’ (Frey & Dueck 2007, Cardille & Lambois 2010, Cardille et al. 2012), which all are obtained by different methods. Those methods are quantitative and flexible, and they potentially have relevant application for a better understanding large-scale landscapes pattern studies.

In Quebec, a large province of Canada spanning nearly 20 degrees of latitude, there are existing maps of forest characteristics and bioclimatic zones, but no comprehensive interpretation of the observable land-cover patterns. In this research we explore the information content of an extensive set of landscape metrics for the province: we asked (1) whether patterns in the set of landscape metrics were consistent with existing partitioning studies into ecological zones and (2) where are differences in landscape patterns the most

and least pronounced between adjacent landscapes in the province. This combination of a quantitative multivariate method and a clustering algorithm applied to landscape pattern metrics could help ecologists and geographers understand landscape patterns across regional- and national-scale study areas and provide objectively determined sampling study sites.

2.2 Methods

2.2.1 Study Area

This research focused on the forested landscapes of Quebec, Canada (figure 3). The province of Quebec is a vast 1.3 million km² (ISQ 2011) region geographically located between 45°N and 63°N and 57°W & 79°30' W (figure 3). Landscape patterns have been created by the interplay of a wide array of natural drivers and regimes (MRNF 2010), many different perturbation sources such as fires and logging (Pan et al. 1999, Bergeron et al. 2001, Boucher et al. 2006, Boucher et al. 2009); patterns are composed of diverse vegetation cover types appearing throughout the Province. At this scale, the main drivers are the wide climatic gradient occurring from south to north, and, to a lesser extent, topography and variation in precipitation regimes (MRNF 2010). These drivers divide the province into vegetation subzones (Robitaille & Saucier 1998) and, in southern Quebec, the Ministère des Ressources naturelles et de la Faune du Québec (MRNF) delineated five of them (figure 3) (MRNF 2010): (A) the northern deciduous forest (sugar maple, white and yellow birch, American linden, hickory and balsam fir); (B) the mixed-wood forest (balsam fir, yellow birch), (C) the continuous boreal forest (black and white spruce, aspen, balsam fir, white birch); (D) the taiga (black spruce, lichens); (E) the forested tundra (dwarf black spruce, lichens).

Disturbances substantially affect the patterns in some subzones. Patterns in the boreal forest (C) are influenced by fire regimes, forestry activities (Bergeron et al. 2001), and insect outbreaks (Payette et al. 2000). The mixed-wood forest (B), closer to population centers, is also affected by forest industries and fires (Boucher et al. 2009). The northern

deciduous forests (A) have been modified by the human activities of farming, urban settlements and forestry for hundreds of years (Pan et al. 1999, Boucher et al. 2006).

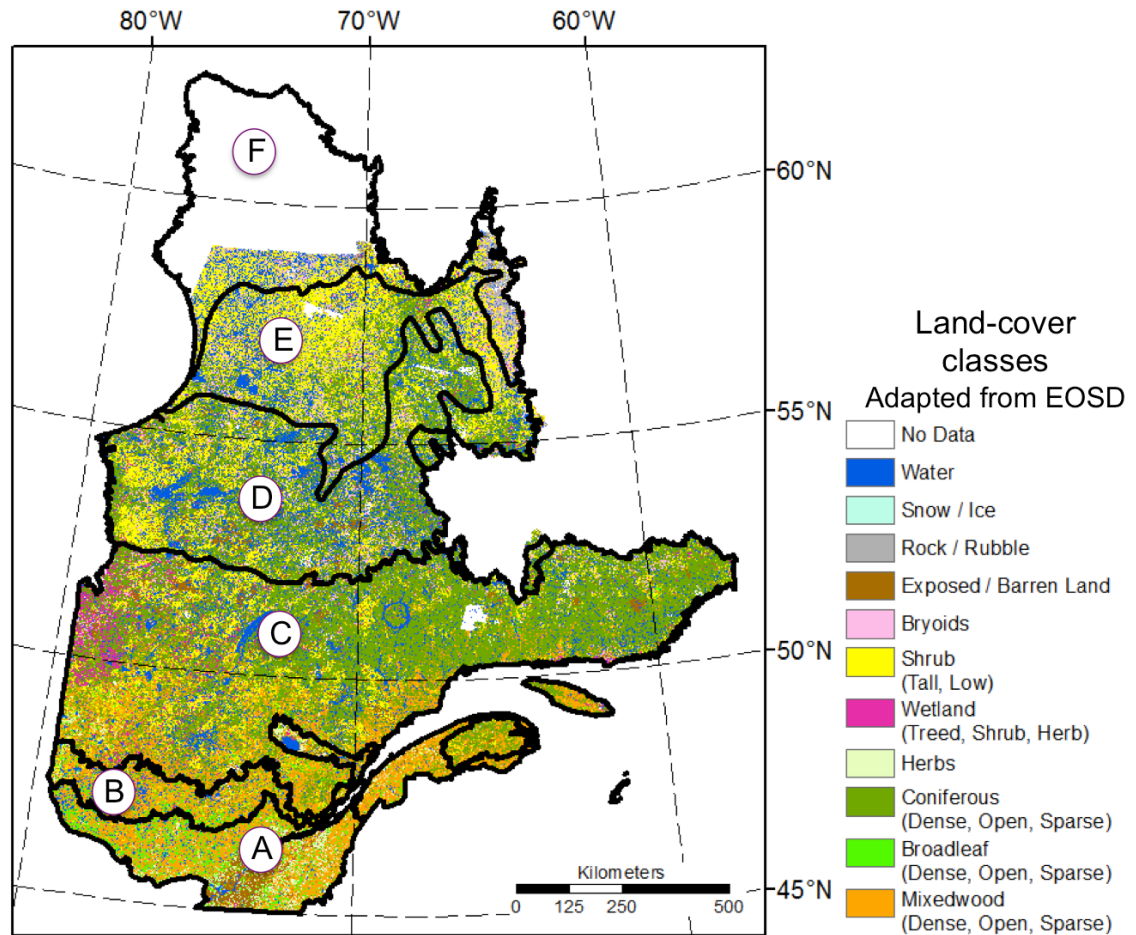


Figure 3. The province of Québec boundaries and, in background, the EOSD land cover characterizes forested landscapes. The MRNF subzones are delineated by black lines and identified by letters. Subzones are (A) northern deciduous forest; (B) mixed-wood forest; (C) continuous boreal forest; (D) taiga; (E) forested tundra. Subzone F is non-forested and not part of the EOSD forest classification.

2.2.2 EOSD land-cover data and landscape metrics

To characterize the land-cover composition of landscapes in Québec, we used the Earth Observation for Sustainable Development of Forests land-cover (EOSD) data set. The EOSD data set is a freely available product derived from Landsat-7 imagery resampled at 25-meter pixels for circa 2000 (background in figure 3) (Wulder et al. 2008a, SAFORAH

2012). The EOSD project categorizes Canada's forested regions with 23 different land-cover classes (Wulder et al. 2008a). The EOSD product did not categorize northern non-forested regions above the forested tree line (the blank F zone in figure 3); so, they were not included in the classification.

We used a seamless EOSD mosaic of Quebec (background in figure 3) to generate regular sized individual landscapes for which we calculated a host of landscape metrics. The 23 possible classes were reduced to 20 usable classes, by recategorizing unusable categories “shadow”, “cloud”, and “unclassified” as an unused “No Data” class. Individual study units, here called “landscapes”, were then extracted from the mosaic using the Canadian National Topographic System 1:50,000 scale individual sheet borders making 1398 rectangular shaped raster landscapes of 900 km² having 25-meter pixels. This particular cutting frame was performed to be consistent with existing nation-wide studies (Wulder et al. 2008a, Wulder et al. 2008c). For each landscape, *Fragstats v3.3* (McGarigal et al. 2002) was used to calculate 90 pattern metrics at the landscape level and 90 at the class level (for each of the possible 20 classes found in a landscape), resulting in a set of 2000 landscape metric values for each unit. Landscape metrics are pixel-based spatial statistics that quantify spatial arrangements and composition of a given landscape and for a given land-cover category (O'Neill et al. 1988, Turner et al. 2001).

2.2.3 Extracting information from metric values, clustering and determining representatives

Because landscape pattern metric values are numbers with different orders of magnitude, and may well be correlated (Li & Wu 2004, Cushman et al. 2008), we scaled them and performed a principal component analysis on the metric set in *R* (R Development Core Team 2011) to form a set of orthogonal meta-metrics (Riitters et al. 1995, Cain et al. 1997, O'Neill et al. 1999, Cardille & Lambois 2010). This eliminated redundancy among metrics and permitted the comparison of the suite of principal components values among landscapes. Because there were 1398 landscapes, 1398 principal components were retained.

Although the metric set contained values for 1398 metrics for each landscapes, a vector of 85 principal components was sufficient to capture 90% of the variation.

We used the values of 1398 principal components to calculate the Euclidean distance between each pair of landscapes as an estimate of the difference between their landscape patterns. This matrix was used as the input to the affinity propagation algorithm (Frey & Dueck 2007). That algorithm clusters more efficiently than the *k-means* algorithm, while it simultaneously identifies an item from the set that best represents each clusters (Frey & Dueck 2007). We used this protocol to determine landscapes that best represent landscape patterns of Quebec. We also developed cluster solutions with a range of cluster counts to better interpret large-scale patterns of forested Quebec. Using the *apcluster* package (Bodenhofer et al. 2011), in *R* (R Development Core Team 2011), we clustered the landscape set to 54 groupings of a specified number of clusters ($k = 2$ to 50, 55, 60, 65, 70, 75, and 100) and corresponding exemplars.

From the 54 clustering outputs, we select certain iterations to compare with hierarchies of the existing MNRF classification. One of the outputs, a five-cluster set, is just one possibility and should not be considered as the optimum clustering answer. One may prefer, for example, more or fewer sampling sites over the province and, we could tune the algorithm to identify the wanted number of cluster along with its matching exemplar-representative landscapes.

2.2.4 Mapping centrality of exemplars

To understand the distribution of landscape patterns in space and assess the spatial pattern of differences within a cluster of landscapes, we developed a map of the nearness of every cluster to its exemplar (Hargrove & Hoffman 2004). For each group, landscapes were mapped and coloured according to their estimated similarity to the respective exemplar. The map, for a given cluster of landscapes, reveals the spatial aspect of similarity among landscape patterns.

2.2.5 Spatial fences

From the 54 clustering levels, segments of reoccurring spatial cluster borders – “spatial fences” (Hargrove & Hoffman 2004) - were drawn to locate the most pronounced variation among adjacent landscapes in forested Quebec. To illustrate the degree of importance of each fence, fences were mapped with varying thickness (Hargrove & Hoffman 2004). Thicker segments represented more noteworthy fences, and the thickness was set using two criteria. First, we added an inversely proportional weight on the number of cluster making borders between spatial clusters thicker (more important) if the “ k ” was small. As a result, the borders of the clustering with two groups were the strongest and thus drawn more thickly. Secondly, thickness was also determined by the existence of a division between cluster types of multiple clustering iterations. Hence, thickest fences present the location of the most drastic changes in landscape patterns we can monitor using the data.

2.3 Results

2.3.1 Composite image of PCA values

The main characteristics of landscape patterns change gradually across Quebec (figure 4). Across the wide spectrum of bioclimatic conditions in the province, areas with similar conditions are grouped together in space. At this scale, some landscape patterns can be often distinguished by a single principal component (visible here as one of the three colours), while some other patterns are better explained by a combination of two, three, or more principal components. For instance, the most northern patterns would be highly related to PC1 in red, while the darker-blue landscapes located in the south-central part of Quebec represent a blend of the 3 principal components.

Imaging the first three principal components (figure 4) reveals some of the ways that the MRNF-delineated zones (black lines) are consistent with the land-cover patterns in landscapes in the EOSD data. The forested tundra subzone (figure 4, E) contains relatively homogeneous land-cover patterns that differ strongly from those of southern Quebec. That is visible because they vary greatly in the first principal component of all the metrics. The

northern deciduous forest subzone (figure 4, area A) is also characterized by relatively homogeneous land-cover patterns, varying substantially in only the third principal component. In contrast, the mixed-wood forest, the continuous boreal forest and the taiga subzones reveal distinct patterns, as variations in the three first principal components are perceptible. Additionally, they are presenting a clear latitudinal structure. The second principal component explains mostly the patterns in the eastern part of the continuous boreal forest (figure 4, C), while patterns are distinguished by the third principal component in the west. This is consistent with the land-cover data where a clear distinction in landscape composition divides the boreal forest and the taiga with dense coniferous landscapes in the east.

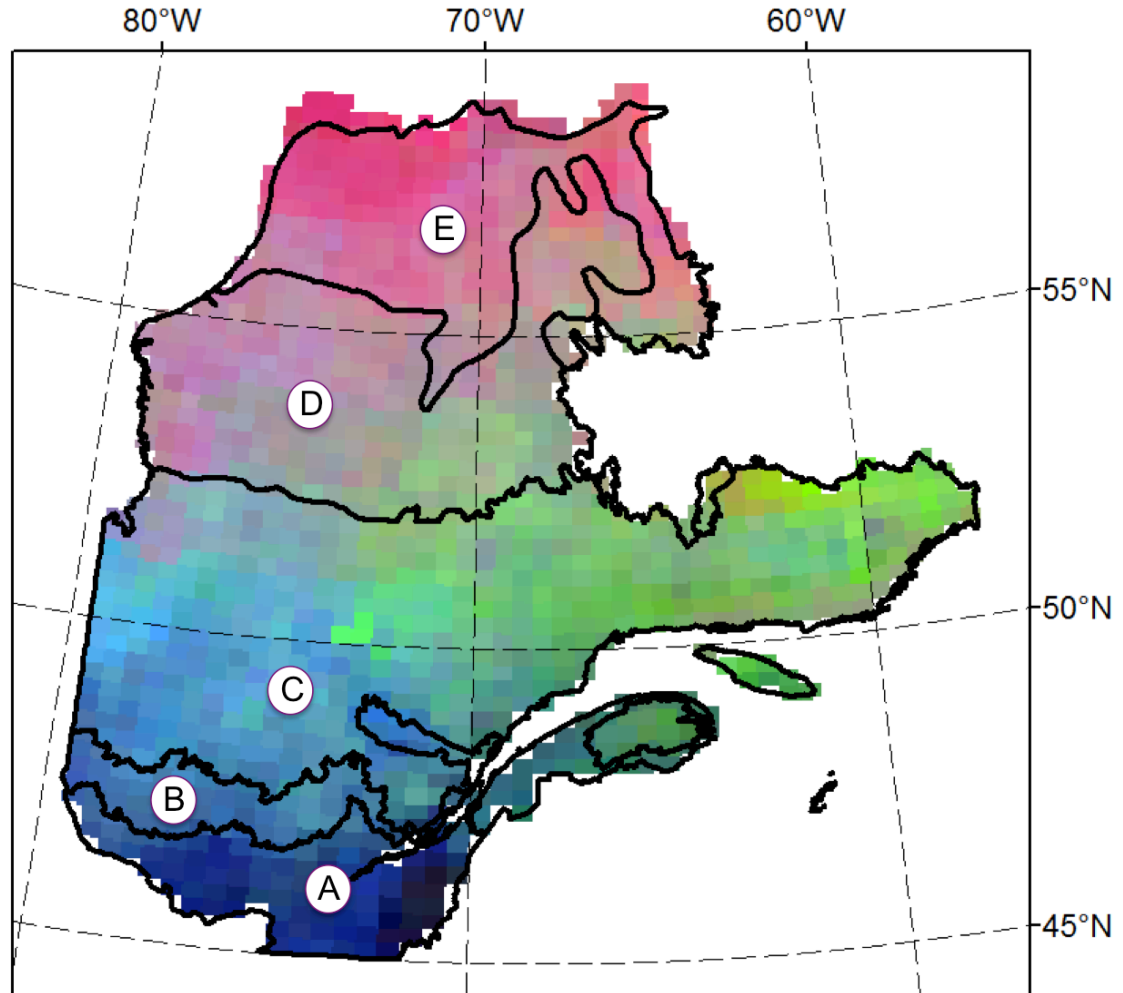


Figure 4. Composite image constituted of the three first principal components as optical bands (PC1 in red, PC2 in green, PC3 in blue). Each square cell represents an individual landscape. Overlaid on top are the MRNF vegetation subzones (black lines), which reveal partly the same structure presented by the three first principal components. MRNF subzones are (A) northern deciduous forest; (B) mixed-wood forest; (C) continuous boreal forest; (D) taiga; (E) forested tundra.

2.3.2 Representative land-cover patterns of forested Quebec

Clusters of land-cover patterns generally conform to the latitudinal bioclimatic gradient, and the patterns within representative landscapes are, with little exception, highly comparable with existing description. The exemplars represent the landscape patterns found within clusters (figure 5). The spatial distribution of representative landscapes (figure 5) is consistent with the expected latitudinal spatial distribution of the five vegetation subzones

(described in section 2.2.1, illustrated in figure 3). Our interpretation of patterns – based on the EOSD classification – is that the exemplars closely correspond with the descriptions of vegetation composition based on the bioclimatic drivers foreseen by the MRNF (2010) with exception of southern landscapes being marked with humanized patterns:

- **“Exemplar 1 - pasture lands and broad-leaf forest”** (figure 5, #1) is a landscape dominated by a mix of pasture-lands (herbs) and croplands (exposed lands), and patches of dense broad-leaf forest. This landscape corresponds to a typical southern Quebec countryside’s pattern influenced by human activities - square shaped patches - rather than natural land cover patterns. Exemplar 1 fits the MRNF the northern deciduous forest (figure 3, A) with the exception that our results show more disturbed patterns due to human activities, which is consistent with Pan et al. (1999).
- **“Exemplar 2 - mixed-wood forest”** (figure 5, #2) is represented by dense mixed-wood and dense coniferous forest, and logging patterns are also very noticeable with networks of square-shaped ‘shrub’ patches linked by thin straight forest roads. Even though this representative is geographically located within the MRNF corresponding vegetation subzone, its patterns remain similar to the description of the mixed-wood forest (figure 3, B). However, human induced patterns such as typical logging patterns and highly linear patches (roads, power transmission, rail roads) diverge from the descriptions. These patterns match expectations based on previous studies (Boucher et al. 2006, Boucher et al. 2009).
- **“Exemplar 3 - eastern boreal forest”** (figure 5, #3) is characterized by dense and continuous coniferous forest with some sparse patches of bryoids, shrubs, and exposed lands. This landscape presents the same composition as the MRNF’s description of the continuous boreal forest (figure 3, C).
- **“Exemplar 4 - taiga & northern transition forest”** (figure 5, #4) is comprised of very large patches of sparse-open coniferous forest and shrubs, and ubiquitous water bodies. That relates to the MRNF taiga description of open coniferous forests paired with lichen and sphagnum soil covers (figure 3, D).
- **“Exemplar #5 - shrubby tundra”** (figure 5, #5) is mainly represented by shrubs and lakes, and composed of sparse coniferous patches along water patches sides, typical of northern Quebec and similar to the MRNF forested tundra description (figure 3, E).

Despite the fact that clusters did not fit perfectly the MRNFs vegetation subzones, clustering similar landscapes into metric space did not generate arbitrarily partitioned

clusters in geographic space. Clusters are spatially organized and this structure corresponds to the provincial latitudinal gradient. Clustering of landscapes into five groups revealed cohesive patterns in metric space as well as also being homogeneously grouped in space.

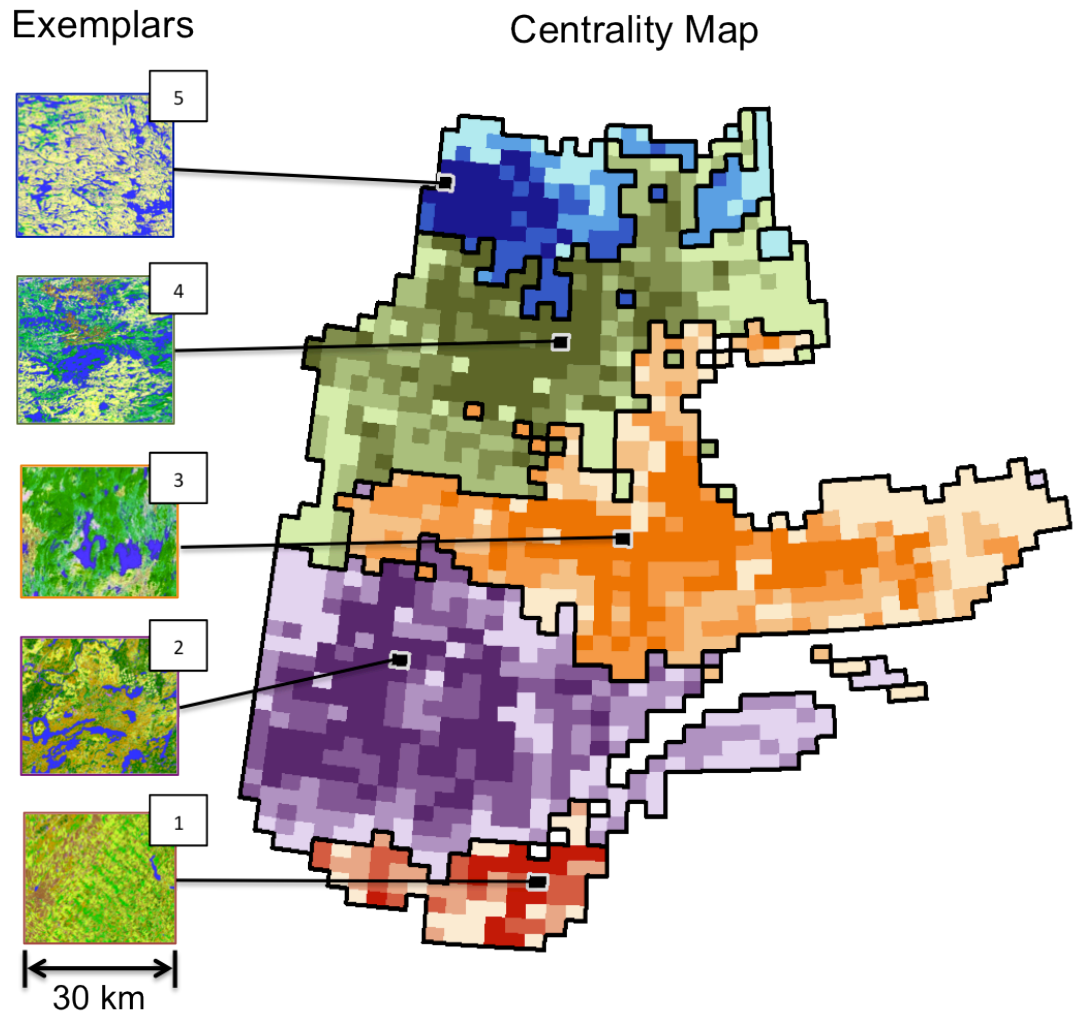


Figure 5. Clustering set of five with representative landscapes (exemplars) located within their clusters in Quebec. In the centrality map, measures of centrality are represented with colour ramps (darker colours represent higher similarity of landscapes to their exemplar, lighter colours represent lower similarity values).

2.3.3 Spatial clusters

With little exception, landscapes that were near each other in metric-space were also near each other in geographic space. Landscapes that were the most similar to their exemplar also strongly tended to be those that were among those geographically nearest to it (figure 5). The centrality map (figure 5) presents the spatial distribution of the similarities between grouped landscapes and their exemplar. As observable, landscapes near exemplars are mostly similar to them (darker colours), while distant landscapes are generally less similar (lighter colours near cluster boundaries). This representation of the exemplars' geographic and numeric centrality reveals the intra-group homogeneity as well as the inter-group heterogeneity. For example, within-group homogeneity can be seen in cluster #3 where similarity values are mostly uniform across the cluster. On the other hand, inter-group heterogeneity can be observed in at least two ways on geographic space. One example is a gradual transition of patterns exemplified by a distinctive strip of light coloured cells on common frontiers between clusters #1 and #2. A much different example is found at the border of clusters #4 and #5 where sharper distinction among patterns is seen.

Despite the fact that no positional information (i.e., coordinates, zones, etc.) was included in the similarity and clustering processes, classifications of landscapes of Quebec display a strong and consistent spatial cohesion (figure 6, top row). As in figure 5, other cluster solutions, made using only landscape metric values, were largely homogeneously grouped when mapped in geographic space, rather than randomly located within the province. With few exceptions, clusters are spatially contiguous, especially for runs with lower number of clusters.

The classified pattern values (figure 6, top row) are roughly consistent with Quebec's major bioclimatic landscape pattern drivers (figure 6, bottom row). First, clustering outputs illustrate a latitudinal structure in spatial distribution of land-cover patterns based on the bioclimatic gradient, which the MRNFs classifications represent (figure 6, bottom row). When looking at maps A-1, B-1 and C-1, the north-south structure

is clear even though it is not as clearly defined as in the MRNFs corresponding maps A-2, B-2, and C-2. Then, for runs with higher numbers of clusters (maps C-1 and D-1), clusters tend to partition the geographic space from west to east along 70°W. That is also coherent with more local- or regional-scale factors such as hydrological regimes and other edaphic factors having consequences on vegetation growth and, presumably, landscape patterns (Robitaille & Saucier 1998, Sarr et al. 2005).

Although our cluster results generally match the main bioclimatic drivers, the MRNF classification (figure 6, bottom row) and the partitioning here (figure 6, top row) differ in the spatial structure of clusters. The clustering of landscape pattern values detected a greater level of detail, or a number of different landscape pattern types, above latitude 50°N than did the MRNF did with its classifications. In contrast, the MRNF classification delineates more regions below 50°N. For example, the eight-cluster map (C-1) has five different clusters above the 50°N, while the MRNF has three (C-2). Below this latitude however, the MRNF map has six different regions (C-2) where there are four clusters in our case (C-1). In general, clustering of landscape pattern values generated more uniformly distributed clusters, as they are more diverse in the north and reflect the observed land-cover conditions.

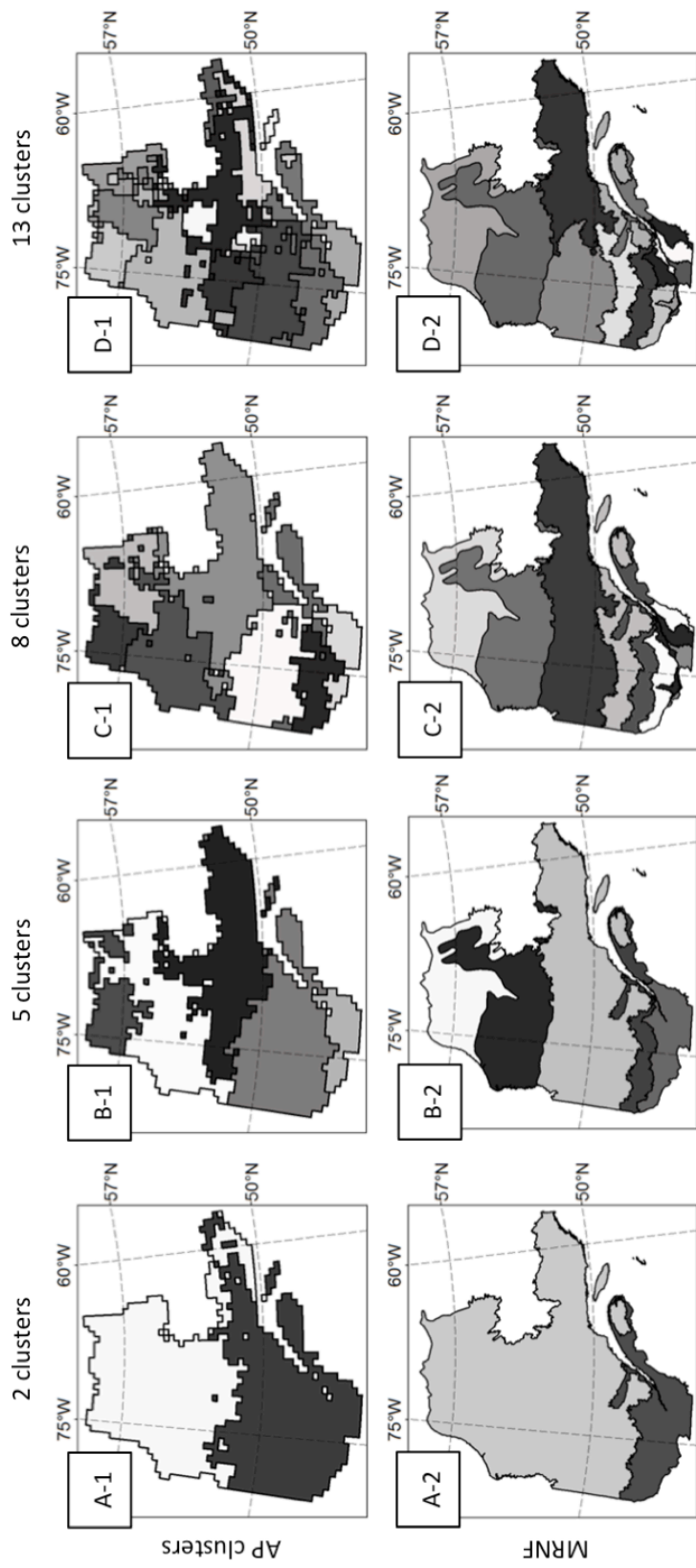


Figure 6. Clustering outputs (top row) compared with hierarchical divisions of the MRNF ecological classifications (bottom row). Each column shows (below) a given hierarchical MRNF classification level and (above) an unconstrained clustering, into the same number of groups of landscapes using landscape metrics (number of clusters = 2, 5, 8, and 13).

2.3.4 Spatial fences

Spatial fences identify the most pronounced variation in land-cover patterns among adjacent landscapes across forested Quebec (figure 7). The degree of magnitude in pattern differences is represented by the thickness of fences: thickest fences identify the most drastic changes (figure 7, box G), while the thinnest fences represent smoother changes (figure 7, box H), and gaps between segments represent areas where changes across space are too gradual, or inexistent, to be noticeable (box H). The thinnest fences are divisions between clusters that reoccurred at least 5 times (box H) across the 54 clustering levels, and the thickest fences reappeared more than 16 times (box G).

Since fences delineate sharp pattern changes in geographic space, they consequently expose at least two interesting facts about Quebec landscape patterns. A first thing to note is the major fence that occur along the 52°N line (box I), where the proportion of dense forest decreases. These segments are among the most pronounced in the study area and they divide the province into two regions from South (where patterns are mainly dominated with higher forest densities) to North (where patterns are mostly characterized by sparsely forest patches) (MRNF 2010). This important fence generally matches the current northern commercial logging limit, which was delimited, partly, by biophysical and stand productivity constraints (MRN 2000). A second feature to note is that fences identify at least two kinds of geographic regionalization in landscape patterns of Quebec; there are “pure” regions, and “mixed” regions. Box J presents an example of pure region where no fences were found more than 5 times out of 54 possibilities. Pure regions are therefore areas where landscape are characterized by strongly homogeneous patterns compared with nearby landscapes. In contrast, mixed regions are areas characterized by a high fence density (box K). That is, mixed regions are those where considerable differences among landscapes occur in a relatively concentrated area.

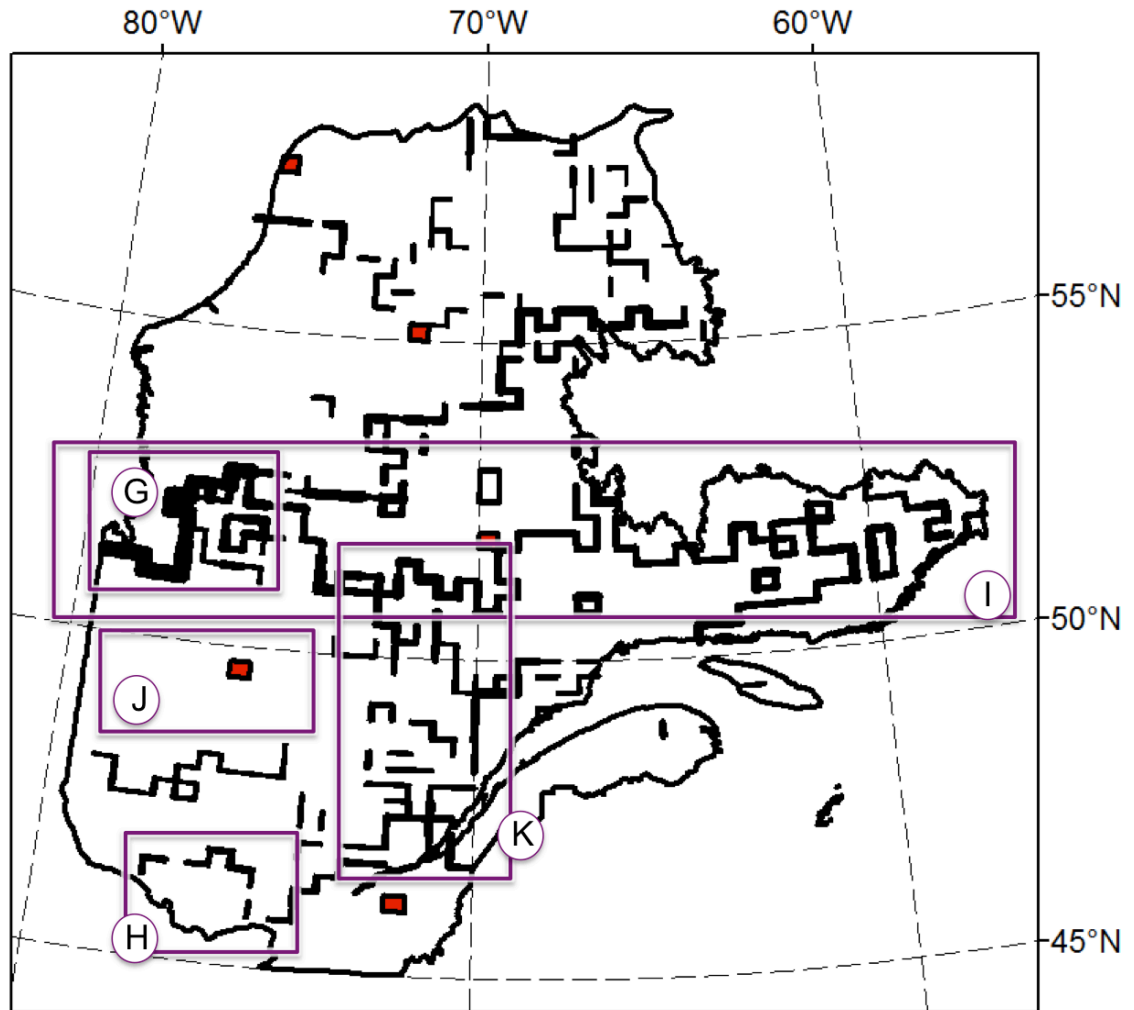


Figure 7. From the 54 clustering levels, segments of reoccurring spatial cluster borders were drawn to locate the most pronounced variation among adjacent landscapes in forested Quebec. Shown are fences that appeared five times or more at the same location. Thickness of fences represents the degree of magnitude in pattern dissimilarities: thicker fences locate more important changes, and gaps locate gradual changes in landscape patterns. In red are shown the representative landscapes presented in figure 5.

2.4 Discussion

2.4.1 Landscape pattern information and clustering

In assessing the total information content of the set of landscape metrics for forested Quebec, we revealed that landscape patterns are largely cohesively structured in space, and that clusters of patterns are broadly consistent with what can be seen in the EOSD data. Although multiple principal component axes were used for understanding differences among landscapes, the three first principal components contained considerable information on pattern spatial distribution across our large study area. This image map also helped us to envision how the clustering algorithm would partition landscapes throughout Quebec.

Clustering results suggest that landscape metrics are powerful enough to reveal general structure in landscape pattern geographic distribution over Quebec. In this case, spatial clusters were spatially aggregated and, surprisingly to us, quite homogeneous throughout the province. They revealed that landscape patterns vary principally along a south-north axis in the province of Quebec and, for greater numbers of clusters, in more complex ways longitudinally. This varying spatial structure, which differs from the MRNF, brought new insights on how we can understand Quebec landscapes.

The spatial patterns of clusters did not entirely match the MRNF ecological classification, which suggests that the MRNF expected patterns do not represent well the current land-cover patterns. Using landscape pattern metrics help to identify the “real” ground landscape patterns, unlike other classifications that are largely based on expected or interpolated conditions and not necessarily meant to capture current conditions. In our opinion, this explains in part the reason why spatial clusters did not neatly coincide with the boundaries in the MRNF classification. The metrics quantified landscape patterns according to the EOSD land cover product: patterns were regionalized considering the ground patterns as captured by the satellite, and therefore not reflecting the theoretical patterns expressed by the MRNF. Consequently, clustering results represent a new type of biogeographical information that could contribute to future broad scale ecological mapping efforts. The quantitative and flexible approach presented here might be adapted for tracing

new ecological regions by combining other types of relevant ecological ancillary data for specific needs.

In contrast to the regionalization of British Columbia of Long et al. (2010), our metric-based clusters formed groupings that were spatially contiguous. This may be due to several factors. First, because Long et al. used substantially fewer metrics (6 versus 1398 in this study), they may have had less information available with which to successfully discriminate among landscapes. Additionally, because their study was on a much smaller area ($5.5 \times 10^4 \text{ km}^2$ vs. $1.3 \times 10^6 \text{ km}^2$) using much smaller study units (1 km^2 vs. 900 km^2), their processing might have searched for much finer-scale variation than in our provincial-scale coarse partitioning to cluster major landscape types. Third, their study region (mountainous British Columbia) may simply contain much more spatial variation than Quebec. The intriguing differences and similarities suggest a need for deeper research on the use of landscape pattern metrics to delineate regions over large areas.

2.4.2 Representative Landscapes

Representative landscapes, available through the clustering algorithm used here, provide an efficient way of objectively summarizing the information content of a large and complex landscape pattern metric data set (Cardille & Lambois 2010, Cardille et al. 2012). Here, exemplars followed, with few but noticeable exceptions, the comparative descriptions of the MRNF vegetation subzones. Exceptions were highlighted in the two southernmost exemplars (#1- pasture lands and broad-leaf forest and #2-mixed-wood forest) and revealed clear patterns of the human land-uses of farming and logging. Further research would be needed in northern deciduous and the mixed-wood forest subzones to investigate whether seeking additional exemplars only within these MRNF subzones would also indicate human-driven patterns. In this setting, where representations of real-world landscapes can be seen, small sets of exemplars are likely easier to quickly understand than summary tables of metric values (Cardille & Lambois 2010, Cardille et al. 2012). In our opinion, this could be a very useful way of generalizing complex scientific research results on landscape patterns and help managers and stakeholders.

Representative landscapes may be useful in a variety of applications for studying forested landscapes in Quebec among many other places. Geographers, landscape ecologists and resource managers could reasonably use exemplars as representative sampling landscapes for broad scale ecological studies. For example, these exemplars could help managers identify a representative provincial-scale network of study sites (Keller et al. 2008), or to verify the spatial representativeness of existing study sites (Hargrove & Hoffman 2004). Cardille et al. (2012) have demonstrated an example of using exemplars for exploring landscape patterns inside and outside Ontario's parks and protected areas. In short, many possibilities remain to be considered, but it is still with great potential that exemplars determined with landscape metrics may inform future broad-scale ecological studies.

2.4.3 Representativeness of exemplars and homogeneity of clusters

Because clustering processes necessarily result in a simplification of the data, assessing the spatial aspects of clusters with centrality maps can be useful in certain ways. With this geographic representation of the similarity values, we can quickly assess the clustering intra-group homogeneity and the inter-group geographic frontiers (figure 5). The centrality map (figure 5) identifies where the similarities or representativeness within groups are the strongest (or the weakest). From this information, field ecologists wishing to avoid unusual or non-representative landscapes in their campaigns could then use such information. The homogeneity of clusters could be useful to locate landscape pattern transition zones over Quebec, for example, where boundaries between clusters are characterized by patterns that are less similar to their exemplar.

The reasonable results obtained with the centrality map suggest that landscape pattern metrics successfully express the spatial distribution of landscape patterns across large areas. By illustrating this, we reinforced that our clustering is not arbitrarily distributed in geographic space – between clusters and within clusters. Despite all the potential issues related to the use of landscape pattern metrics (Li & Wu 2004), using

hundreds of properly treated metrics for quantitatively characterizing and grouping diverse landscapes works well to summarize landscape patterns across Quebec.

2.4.4 Spatial fences locate important landscape pattern differences

Fences synthesize results of multiple clustering iterations seeking different numbers of groups. Here, spatial fences succeeded to display in geographic space the degree of dissimilarity between regions. In figure 7, this variability was shown through the thickness of fences, in which the thickest segments demonstrated the most pronounced differences, while gaps between segments presented the most gradual changes in patterns. This synthesis of multiple clustering solutions through fences delineates pattern changes as a gradient that contrasts with common discrete polygon-based classifications (Hargrove & Hoffman 2004).

2.4.5 Limits of research and future work

It is worth recalling that we used the EOSD land cover product circa 2000, and results shown here are ultimately limited by its classification and quality. In the interim since the satellite data were first collected, ground reality may have changed and our results should be interpreted only in this light. Additionally, since landscape pattern metrics may behave differently depending on the classification, the spatial resolution and the extent of study units, direct comparison or conclusion with similar studies should be performed carefully. In future work, trying a different study unit as the basis for calculating metrics might be fruitful. Notably, the MRNF is using a much smaller unit for resource management – ecological districts. These ecological districts vary both in shape and size since they were delineated using natural features such as relief, geology, geomorphology and regional vegetation (Robitaille & Saucier 1996, Robitaille & Saucier 1998).

2.5 Conclusion

In this research, we demonstrate that the proper use of landscape metrics can capture the broad-scale structure of patterns across a very large study area. To begin, the principal component image composite visually demonstrated a useful capacity of metrics to

express the variation of patterns in geographic space. Clustering of metrics revealed the distribution of landscape patterns for comparison with the main bioclimatic drivers in Quebec. These results present the differences between the observed landscape patterns and the projected theoretical patterns. In short, results have shown the potential of landscape pattern metrics to successfully differentiate landscape patterns across national-scale areas. As the development and use of large-scale products increases, efficient tools will be needed to summarize and represent the data, especially when dealing with national-scale landscape metric sets. As demonstrated by Cardille et al. (2010, 2012), one way of portraying the variability of landscape patterns is through objectively defined representatives. Here, Affinity Propagation was found to be one flexible and efficient algorithm by its twin operation of clustering similar landscapes while objectively identifying representatives within groups. Such exemplars aid the understanding of patterns and our interpretation of clustering outputs.

The centrality map and spatial fences shown here are a new form of biogeographical information about the forested landscapes of Quebec. These contrast with commonly used ecological classifications, in which it is quite difficult to adequately express the spatial homogeneity within groups and the spatial variability of patterns among groups (Hargrove & Hoffman 1999, 2004). Here, the centrality map revealed where patterns are highly similar to representatives, allowing us to locate patterns that can be confidently viewed as spatially cohesive. Spatial fences illustrated where pattern differences are the sharpest, as well as where changes are too gradual to form a discrete spatial entity. This contrasts with typical region classifications, in which contiguous closed polygons are most often assumed to contain relatively homogenous elements within their borders. The methods and results presented in this paper are, we believe, only a starting point for regionalization of large-scale ecological classifications using landscape metrics. The quantitative and flexible approach presented here can be used directly by resource managers and can easily be adapted with different data sets as well as integrating other types of ancillary data for other specific needs.

3. Deuxième article

Présentation du deuxième article

Le chapitre trois présente le deuxième article scientifique contenu dans ce mémoire qui est composé en deux volets. Dans le premier, nous avons examiné les capacités des indices de patrons à distinguer les paysages dans un contexte géographique prédéfini à l'échelle du Québec méridional. L'analyse des résultats a suggéré que les indices ont cette capacité, mais que pour certains types de paysages, les patrons prédits par la classification existante pouvaient être incohérents comparativement aux patrons réels observés. Pour le deuxième volet, nous avons repris essentiellement la même méthodologie que pour le premier article afin de montrer les différences majeures que l'on obtient à partir d'une classification libre *a priori* de toutes contraintes géographiques. Nous concluons en faisant une mise en garde sur l'interprétation du découpage discret souvent retrouvé dans les classifications écologiques de références puisqu'elles n'informent pas sur la variabilité des phénomènes ni sur l'homogénéité des patrons dans l'espace géographique.

Accord du coauteur

Ce premier article a conjointement été écrit par Kevin Partington, premier auteur, et par son directeur de maîtrise Jeffrey A. Cardille. L'article est soumis à la revue scientifique « *Écoscience* ».

À titre de coauteur, j'autorise mon étudiant Kevin Partington à publier l'article « *Constrained and unconstrained classification of landscape patterns in southern Quebec* » dans son mémoire de maîtrise.

Jeffrey A. Cardille

Constrained and unconstrained classification of landscape patterns in southern Québec

By Kevin Partington & Jeffrey A. Cardille

3.1 Introduction

It is a longstanding challenge to appropriately map large areas. For this, scientists typically use a variety of maps that classify broad regions according to a range of criteria. In ecology, examples include terrestrial eco-region maps (Olson et al. 2001), environmental domain classifications (Robitaille & Saucier 1998, Coops et al. 2009, MRNF 2010), and continental biogeographical regionalizations (Hargrove & Hoffman 1999, Mackey et al. 2008). Many scientists rely on these products for accomplishing their research (Olson et al. 2001, Hutchinson et al. 2005), while resource managers need them for better planning of resource uses (Loveland & Merchant 2004). In southern Quebec, for example, the delineation of bioclimatic domains reflects largely stable biophysical factors, such as climatic variables and end-of-succession vegetation (MRNF 2010).

Despite their clear utility, typical broad-scale classification maps are limited in ways that can have important implications. Regionalization of environmental factors into contiguous polygons inherently implies a generalization of the phenomenon, thus, they comprise inconsistencies or heterogeneity within regions (Loveland & Merchant 2004). In other words, landscape patterns vary locally within regions in ways that the heterogeneity is unknown for whole regions. Because most polygon-based maps do not represent variability within regions (Hargrove & Hoffman 1999), it is hard to know whether or not areas that are really representative of the conditions predicted on the map.

Since broad-scale classification maps are mostly built on slowly changing biophysical attributes, it is difficult to say whether or not broad-scale ecological maps match today's observable land-cover conditions. In a given area, landscape patterns are not only a result of the interaction between environmental drivers such as bioclimatic regimes but also the consequences of disturbances such as windthrows, insect outbreaks, wild fires,

and anthropogenic land-uses such as logging activities, just to name a few (Turner et al. 2001, Turner 2005). Hence, the local variations in environmental drivers may cause the patterns to change across space (Bailey 2004), which may lead to potential misinterpretation of predicted homogeneous conditions.

In addition to ecological classification maps, scientists have access to an increasing number of high quality land-cover data sets to study large areas (e.g. MODIS and AVHRR). Land-cover maps present a pixel per pixel characterisation of the ground surface composition at a much finer scale than that of most polygon-based regional maps (Lillesand et al. 2008). Such raster land-cover maps classify each pixel with its corresponding vegetation cover, rather than grouping similar objects into closed polygons. To understand the variability of land-cover data within large regions, tables of statistics are often constructed to summarize the land-cover conditions (Lillesand et al. 2008, Wulder et al. 2008a, Cardille & Lambois 2010). For those interested in landscape patterns, land-cover maps typically quantify landscape composition and mostly avoid related questions of land-cover configuration.

For an improved understanding of current land-cover patterns, the large set of landscape metrics in landscape ecology literature has the ability to measure both landscape composition and configuration of ground patterns from categorical land-cover maps. Typically, they are used in studies seeking relationships between spatial characteristics of landscapes and ecological processes (Turner 2005). More recently, landscape metrics have been integrated as ancillary data for large-scale ecological investigations, such as for national forest fragmentation assessments (Kupfer 2006, Wulder et al. 2008c). Long et al. (2010) have regionalized metrics to create spatial pattern regions of forested landscapes in a region of British Columbia, Canada. Cardille et al. (2010, 2012) have used landscape metrics to identify the most representative landscapes for the USA (Cardille & Lambois 2010) and for forested regions of Canada (Cardille et al. 2012). Landscape metrics have the potential to characterize the observed land-cover patterns across large areas.

It remains unclear to what extent the southern Quebec vegetation subzones match today's ground conditions. The MRNF have delineated three vegetation subzones in southern Quebec as part of a hierarchical system of ecological classification (MRNF 2010). This ecological classification was produced for scientists and resource managers for a better understanding and use of natural resources, especially for forest interests (Robitaille & Saucier 1998). In Quebec, landscape patterns at any moment are a function of many abiotic, biotic, and human factors (Bergeron et al. 2001, Turner et al. 2001, Boucher et al. 2006, Boucher et al. 2009), only some of which are directly incorporated into the MRNF classification. Because they categorize the current conditions of landscape patterns, current satellite-based assessments of Quebec's forested land-cover can be used to better understand the differing goals and results of these two approaches to understanding large-scale patterns on the earth's surface.

We had two main objectives for this research. First, we examined the spatial structure of landscape patterns within the MRNF vegetation subzones in southern Quebec, asking whether and where patterns of landscape metrics correspond with established bioclimatic boundaries. Second, we asked what were the current landscape patterns of southern Quebec free of the constraints of existing classifications. The results will allow resource managers to quantify the degree of fit between today's patterns and existing classifications, and suggest a reclassification of southern Quebec's landscapes according to land-cover patterns. Scientists could be interested in knowing if land-cover patterns are consistent with the expected conditions, and where inconsistencies are located. Answers to these questions may provide a clearer understanding of the opportunities and limits of describing and mapping land-cover patterns across large areas.

3.2 Methods

3.2.1 Study area

This study focused on the southern region (figure 8) of the Province of Quebec, Canada corresponding with the three southern-most vegetation subzones delineated by the *Ministere des Ressources naturelles et de la Faune du Quebec* (MRNF) (MRNF 2010). The study area is located between 45°N and 53°N, and from 57°06'W to 79°30'W, covering nearly 760,000 km². Human settlements and farming activities are established along the St-Laurent river shores and generally in the south of the study area (Z11 in figure 8) while the northern regions (Z12 and Z21 in figure 8) of the study area remain relatively unsettled by humans with the exception of substantial logging activities.

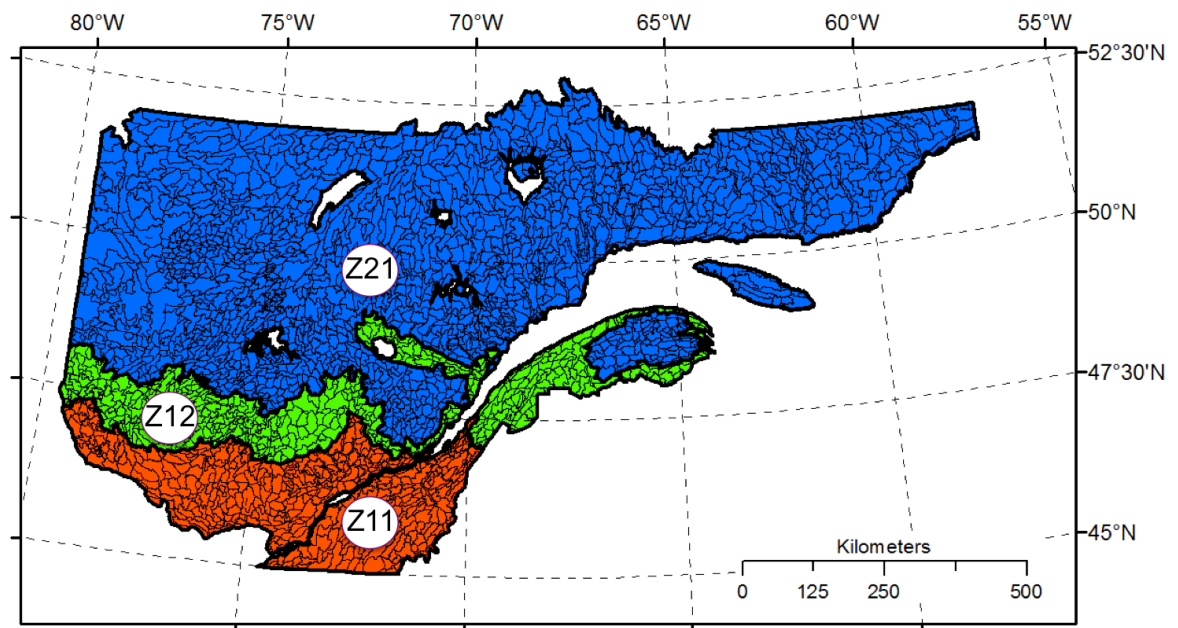


Figure 8. Southern Quebec categorized by three vegetation subzones (Z11-Orange: deciduous forest; Z12-Green: mixed-wood forest; Z21-blue: continuous boreal forest), and by ecological districts (dark gray contour lines polygons). Sources: MRNF (2005, 2010).

The MRNF vegetation subzones delineate contiguous regions characterized by homogeneous vegetation assemblages of landscapes at the end of ecological succession (Robitaille & Saucier 1998). The MRNF identified three vegetation subzones in Southern Quebec (figure 8): Z11, the northern deciduous forest (sugar maple, white and yellow birch, lime tree, hickory and balsam fir), Z12, a transition zone represented as the mixed-wood forest (balsam fir, yellow birch), and Z21, the continuous boreal forest (dense stands of black and white spruce, aspen, balsam fir, white birch).

3.2.2 Land-cover data and landscape metrics

The primary data used in this study was the Earth Observation for Sustainable Development of Forests (EOSD) (Wulder et al. 2008a, Wulder et al. 2008c). Derived from Landsat imagery and resampled at 25-meter pixels covering forested regions of Canada (Wulder et al. 2008a, SAFORAH 2012), the data set classifies each pixels into one of 23 land-cover classes and represents Canada's forest land cover circa 2000 (Wulder et al. 2008a, Wulder et al. 2011b). Using the full set of 23 available classes, EOSD mosaics were cut along our study area borders and merged into a seamless land-cover mosaic of southern Quebec. The mosaic was then reclassified into 20 categories to make the unusable categories Shadow, Cloud, and Unclassified to be NoData.

In the hierarchical system of ecological classification, the MRNF divides vegetation subzones into smaller multiple irregular polygons called ecological districts. Ecological districts are delineated according to multiple criteria: relief, geology, geomorphology and regional vegetation. In the three zones of Southern Quebec, there are 2489 ecological districts (figure 8). The contours of each ecological district were used to extract individual raster landscapes from the land-cover mosaic. A total of 2489 raster landscapes varying in shape, size and land-cover composition were produced. For each landscape, landscape metrics were computed using *Fragstats v3.3* (McGarigal et al. 2002). We followed the protocol of Cardille and Lambois (2010) to calculate 90 landscape level metrics and 90 class level metrics for each of the 20 possible categories.

3.2.3 Constrained clustering: interpreting characteristics of landscapes in vegetation subzones

To better understand landscape patterns in MRNF vegetation subzones, we explored the spatial and statistical aspects of landscape metrics in southern Quebec. We asked which landscapes best fit their respective MRNF vegetation subzones with landscape metrics using classification trees (De'ath & Fabricius 2000). We called this clustering “constrained” because the fit of landscape characteristics to the pre-defined MRNF geographical zones was under study. Using classification tree classifiers, we forced the landscape classification to fit the existing spatial boundaries of vegetation subzones established by the MRNF in southern Québec. The goals of using classification trees were to identify a small set of landscape pattern metrics that properly distinguish patterns in the respective vegetation subzones, and to classify landscapes with the identified metrics (Breiman et al. 1984). Trees were constructed with the *rpart* package in the *R* statistical software (R Development Core Team 2011) by processing the 1825 pattern metrics of each landscape. We produced 13 different tree sizes, where the smallest tree had three leaves and the largest had 25 leaves.

For each classification tree, we produced a confusion matrix (Lillesand et al. 2008) to identify differences between the MRNF classification and current land-cover patterns. Here, misclassified landscapes were neither good nor bad, but simply were those whose land-cover patterns were most similar to a subzone other than that in which they were located. In these tables, the User's accuracy is defined by the ratio of the number of correctly classified landscapes to the total number of landscapes fitted in that category. The Producer's accuracy is the ratio of the number of correctly classified landscapes to the original MRNF reference classification. For interpretation, the Producer's accuracy was preferred over the User's since it presents the fit with respect to the constraining MRNF subzones. Overall accuracy is calculated by dividing the sum of all three correct numbers of landscapes by the total number of landscapes.

Maps of the predicted fits were produced to locate similarities and differences with MRNF vegetation subzones. To illustrate the overall difficulty of categorizing a landscape

into its accepted subzone, we counted and superposed all of the mismatches obtained across the range of 13 constructed trees. The frequency of misclassified landscapes was calculated—that is, the numbers of times each landscape mismatched MRNF vegetation subzones. From this, a map was created to locate reoccurring inconsistent landscapes.

3.2.4 Unconstrained classification of landscapes: understanding and mapping broad-scale landscape pattern

To explore the current landscape patterns of southern Quebec, we grouped and mapped landscape patterns free of the MRNF subzone geographical constraints. We called this clustering “unconstrained” because landscapes were grouped in metric space free of *a priori* geographic limitations. The unconstrained clustering analysis allowed us to categorize landscapes based only on their observed land-cover patterns.

To allow landscapes to be clustered according to landscape patterns, we reduced redundancy correlation among metrics, using principal components rotation (Riitters et al. 1995, Cain et al. 1997). Following Cardille and Lambois (2010), the similarity of patterns between landscapes was estimated as the negative Euclidean distance between the values on the principal components axes. These similarity values were employed by the affinity propagation algorithm (Frey & Dueck 2007, Mezard 2007, Bodenhofer et al. 2011) to cluster similar landscapes and simultaneously determine the best representative of each cluster. For direct comparison with the three vegetation subzones established by the MRNF classification, we clustered the landscapes into three groups. Then, to understand the sensitivity of the results to the requested number of clusters, we used the aspect of the clustering algorithm that allows the user to specify a desired number of clusters. For this process, we repeatedly ran the clustering and generated 23 sets of clusters, for 2, 3, 5, 6, ... 25 clusters.

3.2.4.1 Creating Spatial Fences

We explored the geographic borders between clusters in the unconstrained classification by locating the most important borders among landscape types in southern

Quebec. For this, the idea of “spatial fences” was adapted from Hargrove and Hoffman (2004), which consisted of identifying cluster boundaries that appeared repeatedly across the iterations with different numbers of groups. The spatial fences between types of land-cover assemblages occurred from 1 to 23 times, and a colour ramp was applied to express the degree of importance of fences: fences between very different land-cover assemblages are distinguished, and were not intended to form closed polygons as opposition to typical ecological regionalization work (Hargrove & Hoffman 2004). Instead, gaps in fences occur where the transition between pattern types is gradual.

3.2.4.2 Creating a spatial pattern diversity map

To better understand the spatial trends and the variability of the different pattern types across the range of clustering levels, we produced a spatial pattern diversity map. Block statistics were processed on each clustering level to calculate the different number of different cluster types. The number of the different cluster types was calculated for block windows of 30 km by 30 km. Each 30 km by 30 km block was given the number of different cluster types, with results varying from 1 to a maximum of 5. Then, the diversity calculations were summed to obtain a final diversity map with scores between 23 and 103. The final output was then categorized into three classes by using the Jenks natural breaks algorithm in ArcGIS (ESRI 2011).

3.3 Results

3.3.1 Constrained clustering of landscape patterns in Southern Quebec

3.3.1.1 Classification trees

With respect to MRNF subzones, the majority of landscapes in southern Quebec can be properly distinguished using only two metrics: the percentage of coniferous-open forest and the percentage of coniferous-sparse forest. A minimal three-leaf tree indicates that two metrics can be used to correctly fit 82% of landscapes within the original MRNF subzones (figure 9). Landscapes having a relatively high (>6.035%) proportion of coniferous-open forest are associated with the continuous boreal forest subzone. Landscapes characterized

with lower proportion of coniferous-open forest were associated with either the northern deciduous forest or the mixed-wood forest. Of landscapes with a low proportion of coniferous-open forest, those with a higher percentage ($>0.145\%$) of coniferous-sparse forest were tagged as mixed-wood forest; the remainder were classified as northern deciduous forest. The proportions of these two coniferous forest categories were powerful enough to distinguish landscapes of three subzones with a relatively high matching rate.

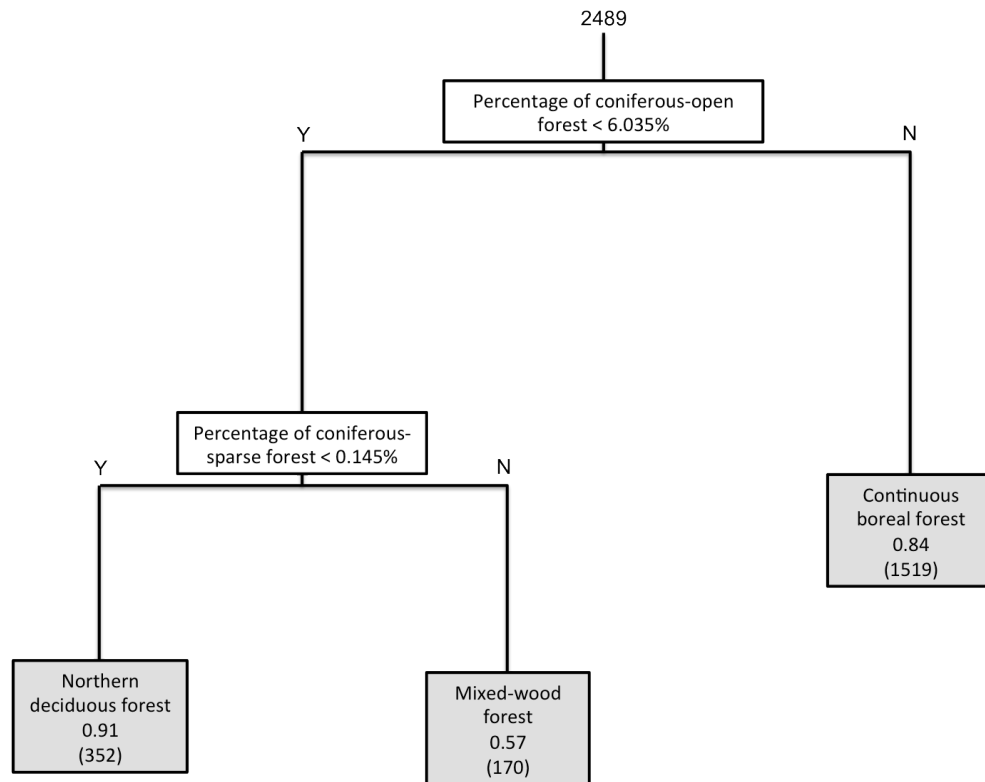


Figure 9. A three-leaf tree for 2489 landscapes constrained to MRNF vegetation subzones (northern deciduous forest, mixed-wood forest, continuous boreal forest). This tree is the simplest that contains two nodes. Presented on each leaf: the fitted subzone type, the percentage of correctly categorized landscapes in that leaf, and the number of correctly categorized landscapes in that leaf (in parentheses). This tree has 448 miscategorized landscapes, which represent 18% of the total.

The three vegetation subzones required different numbers of metrics to properly distinguish patterns among them. When looking at the nine-leaf tree (figure 10), the northern deciduous forest was distinguished well from other subzones with the use of three different metrics (percentage of coniferous-open forest, percentage of coniferous-sparse

forest, and the mean area of water patches). Patterns were harder to differentiate between mixed-wood forest and boreal forest landscapes: eight metrics were needed to distinguish these subzones. Yet, the eight metrics are easy to interpret, and involve land-cover classes (coniferous-open and –sparse forest, mixed-wood dense forest, broad-leaf open forest, water) that are plausibly linked to membership in these different MRNF subzones.

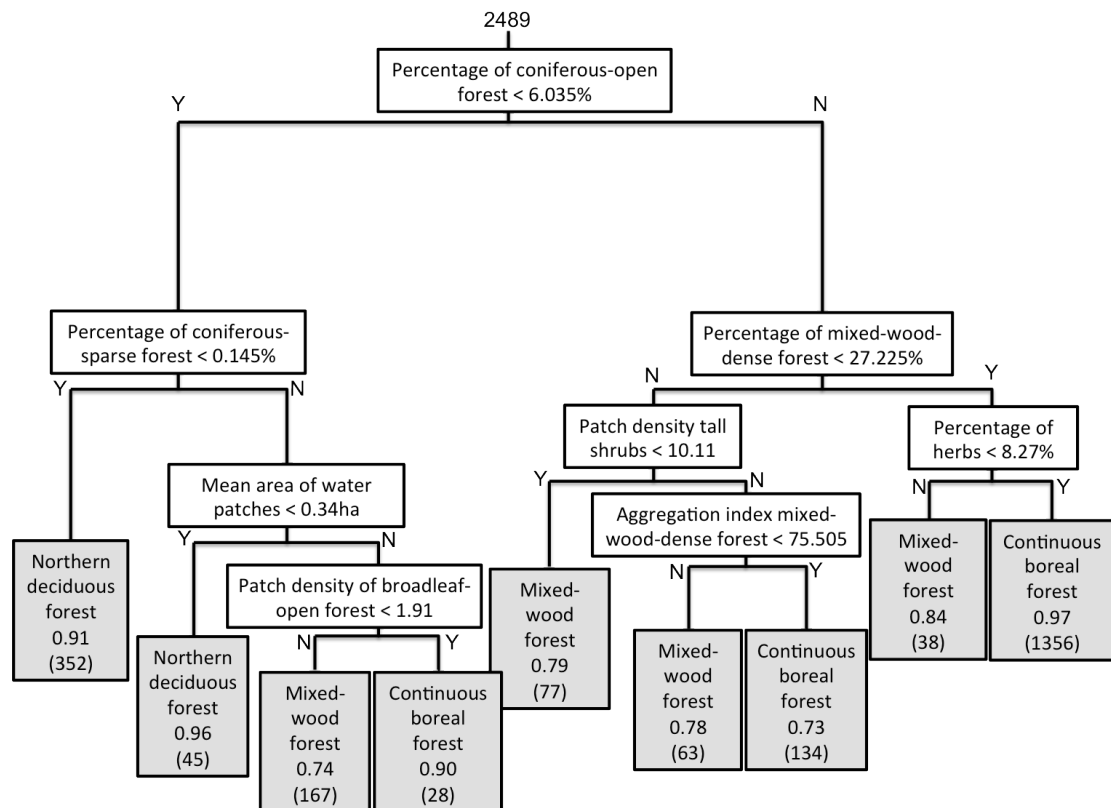


Figure 10. Resulting classification of the maximal nine-leaf tree for 2489 landscapes constrained to MRNF vegetation subzones (northern deciduous forest, mixed-wood forest, continuous boreal forest). Presented on each leaf: the fitted subzone type (one of three), the percentage of correctly categorized landscapes in that leaf, and the number of correctly categorized landscapes in that leaf (in parentheses). This tree had an inconsistency rate of 9%. This tree size was selected because of its relatively limited number of splitting nodes, good fit, and interpretable and understandable set of metrics.

3.3.1.2 Relationship between land-cover patterns and MRNF vegetation subzones

Although the three-leaf tree was able to approximately classify patterns with two metrics (figure 9), it offers relatively less power for classification compared to larger trees

(figure 10). There were limits in using such a minimal tree: for landscapes that are located in the mixed-wood forest subzone, the three-leaf tree correctly placed only 37% of landscapes (table 1), while the two other subzones were strongly fit with the two metrics (96% and 78%). Not all landscapes were equally well classified with these two metrics according to the MRNF constraints, especially for landscapes of the mixed-wood forest.

Table 1. Misclassification matrix for a classification based on the three-leaf tree.

Three-leaf tree		Reference Classification (MRNF)			Prediction Total	User's Accuracy
		Northern Deciduous Forest (Z11)	Mixed-wood Forest (Z12)	Boreal Continuous Forest (Z21)		
Classification Results	Northern Deciduous Forest (Z11-orange)	352	23	10	385	0.91
	Mixed-wood Forest (Z12-green)	71	170	56	297	0.57
	Continuous Boreal Forest (Z21-blue)	26	262	1519	1807	0.84
Reference Total		449	455	1585	2489	Overall
Producer's Accuracy		0.78	0.37	0.96	Overall	0.82

Misclassifications represent landscapes whose MRNF subzone was difficult to reconcile with observed land-cover patterns. The misclassification matrix for the three-leaf tree demonstrates that the mixed-wood forest has patterns that are the least uniform and unique, making it the most challenging to fit of the three MRNF subzones (table 1). Using the three-leaf tree, a total of 448 landscapes were different from the MRNF reference classification. With the three-leaf classification tree, the Producer's accuracy was high for the continuous boreal forest (96%), and for the northern deciduous forest (78%). Misclassified landscapes from the three-leaf tree implied that the MRNF mixed-wood forest subzone was too large and included many landscapes that were more similar to the boreal forest representative. The three-leaf tree classification indicated that of the 455 MRNF mixed-wood forest landscapes, 262 were better described as boreal forest landscapes and 23 as deciduous forest landscapes.

Table 2. Misclassification matrix for a classification based on the nine-leaf tree.

	Nine-leaf tree	Reference Classification (MRNF)			Prediction Total	User's Accuracy
		Northern Deciduous Forest (Z11)	Mixed-wood Forest (Z12)	Continuous Boreal Forest (Z21)		
Classification Results	Northern Deciduous Forest (Z11-orange)	397	23	12	432	0.92
	Mixed-wood Forest (Z12-green)	43	345	55	443	0.78
	Continuous Boreal forest (Z21-blue)	9	87	1518	1614	0.94
	Reference Total	449	455	1585	2489	Overall
	Producer's Accuracy	0.88	0.76	0.96	Overall	0.91

Despite the additional power given by the nine-leaf tree compared to the three-leaf tree, the mixed-wood forest subzone remained relatively difficult to explain within its accepted geographic constraints. The nine-leaf tree produced a total fit of 91%, making 229 different landscapes (table 2). The northern deciduous forest and the continuous boreal forest were easily predicted with Producer's accuracy of 88% and 96%, respectively. In both cases, mismatches appeared mostly within mixed-wood forest with 43 deciduous forest landscapes and 55 boreal forest landscapes classified as mixed-wood forest. The northern deciduous forest Producer's accuracy was 10% higher with the nine-leaf tree, while remaining the same for the continuous boreal forest. The mixed-wood forest was harder to fit, and had a 76% Producer's accuracy. There were 87 mixed-wood misclassifications seen as being more similar to landscapes in the continuous boreal forest. To a lesser extent, the mixed-wood forest had 23 inconsistent landscapes classified as northern deciduous forest. This highlights the fact that the MRNF mixed-wood forest border is difficult to reconcile with respect to the land-cover patterns observed on the ground.

3.3.1.3 Predictions based on classification trees

As observed for the classification trees, differences between MRNF subzones and land-cover patterns were unevenly distributed among the vegetation subzones. In both predictions (figure 11), trees produced reasonable fits for both northern deciduous forest and continuous boreal forest subzones, indicating that these two subzones were both relatively homogeneous and different from each other. Most differences were located within and along the original MRNF mixed-wood forest subzone limits. Landscapes located near the MRNF mixed-wood boundaries tended to be difficult to classify with these trees.

Prediction maps show that parts of the MRNF subzones are composed of a number of landscapes whose patterns more closely match other subzones. In the case of the three-leaf prediction (figure 11.a), two large clusters of mixed-wood forest (figure 11.a, boxes A and B) are misclassified in the original northern deciduous forest subzone just south of the borders drawn by the MRNF. The classification of continuous boreal forest landscapes is generally well predicted except the large cluster of mixed-wood forest landscapes (figure 11.a, box D) located just north of the MRNF mixed-wood and boreal forest border in the mid-west. Mismatches between land-cover patterns and MRNF subzones were mainly related to the eastern part of the original MRNF mixed-wood forest (figure 11.a, box C)

In the case of the nine-leaf tree prediction (figure 11.b), there are fewer misclassifications but, as in the three-leaf tree, they are located near and within the mixed-wood forest. Most categorized northern deciduous forest landscapes are placed within the proper MRNF subzones using the nine-leaf tree. Nevertheless, there are some mismatches located in the MRNF northern deciduous subzone, which were classified as both mixed-wood forest and continuous boreal forest (box F). Most MRNF continuous boreal forest landscapes are clustered with others rightly, with exceptions of a few inconsistencies being classified as mixed-wood forest (figure 11.b, box H). In landscapes classified as mixed-wood forest by the MRNF, the nine-leaf tree placed most correctly. Many landscapes, however, appeared more similar to those found in the continuous boreal forest subzone (figure 11.b, boxes E, F, G). More importantly, almost all misclassifications of the three-

vegetation subzones classification occurred along and within the mixed-wood forest subzone.

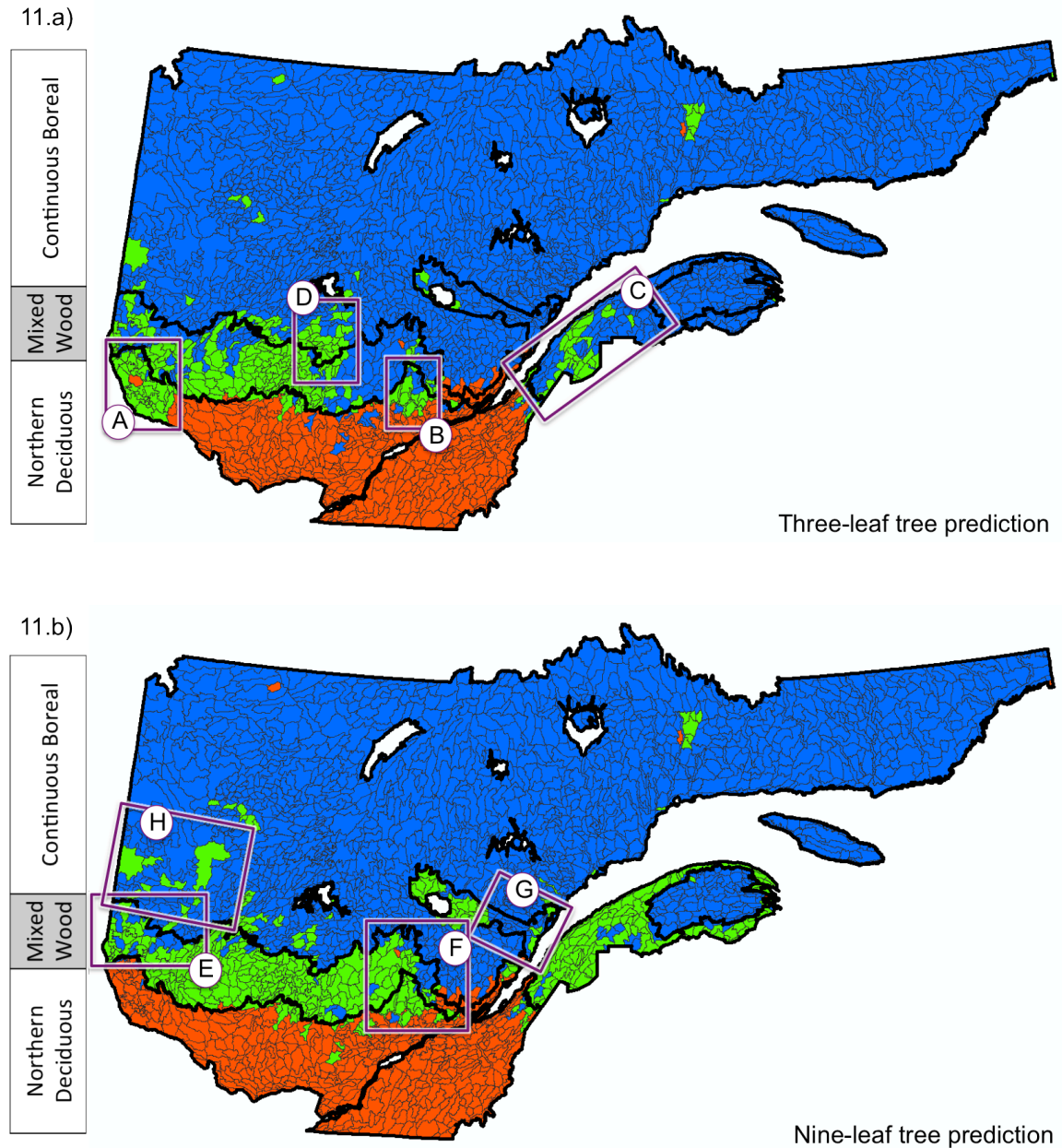


Figure 11. Spatial characteristics of the constrained classifications. Landscapes are mapped with different colours according to fitted groups from table 1 and table 2: in orange are northern deciduous forest landscapes, in green are mixed-wood forest landscapes, and in blue are continuous boreal forest landscapes. The black thick lines represent the established MRNF vegetation subzone borders (as seen in figure 8). Inconsistencies locate where the MRNF

classification is not well matched by current landscape patterns.

Figure 11.a) Spatial fit of the three-leaf tree constrained classification. The overall accuracy for this prediction is 82%. Differences between the two assessments are located mostly in the original MRNF mixed-wood forest (box C). The border between existing MRNF mixed-wood forest and continuous boreal forest is also very different. Exception to the green cluster in the middle of this border (box D), landscapes of boreal forest (coloured in blue) is still very well predicted.

Figure 11.b) Spatial fit of the nine-leaf tree constrained classification. The classification results match 90% of the original subzones. Both the northern deciduous forest (orange) and the continuous boreal forest (blue) are well fit. In contrast, the mixed-wood classification results (green) are not as well reproduced, with inconsistencies of both northern deciduous forest (orange) and continuous boreal forest (blue) categorized landscapes (boxes E, F, G).

3.3.1.4 Misclassification map of constrained classification

In Southern Quebec, several landscapes were identified whose patterns were consistently different from those of their nearby landscapes that had been classified into the same zone by the MRNF. These landscapes were consistently challenging to constrain, for classification trees of a range of sizes, into the pre-defined subzones using data from landscape metrics (figure 12). They possessed landscape patterns that tended to identify with landscapes in a different zone; for example, a landscape in northern deciduous forest would be better placed in mixed-wood forest by the nine-leaf classification tree. This phenomenon can be interpreted in one of at least two ways: these are landscapes whose patterns were formerly similar to others in their zone, but which changed in intervening years; or these landscapes were difficult to classify in the original MRNF, but were drawn within a zone for other reasons—for example, to keep the zone contiguous.

Misclassifications occurred mainly within and along borders of the mixed-wood forest. Across the range of classification tree sizes, the frequency of inconsistent landscapes is expressed with a colour ramp (figure 12). Less frequent misclassifications—i.e., landscapes that were easier to predict in a constrained context—are shown in light colours (boxes I, J and K), and they were inconsistent between 1 and 8 times. Darkly coloured landscapes denote different fits across the range of classifications from 9 to 13 times (boxes L, M and N), and are mostly located along the MRNF mixed-wood forest borders. Most

landscapes were consistently classified with just a few landscape metrics into their proper MRNF zone (white areas in figure 12)

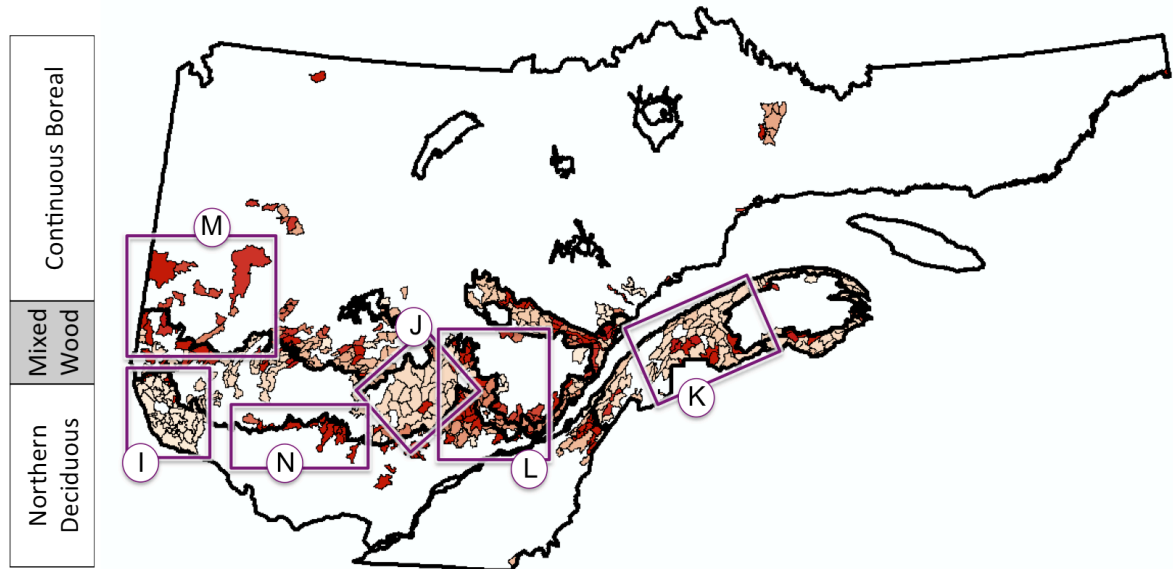


Figure 12. The misclassification map of constrained classifications identifies where the reference MRNF classification is incompatible with observed land-cover patterns. Most landscapes were consistently classified by using a few landscape metrics into their proper MRNF zone (in white). Yet, reoccurring different landscapes were identified within and along the mixed-wood forest originally defined by the MRNF (coloured from light pink to dark red). Different landscapes were harder to reconcile with the reference zones.

3.3.2 Unconstrained clustering of landscape patterns in Southern Quebec

3.3.2.1 Spatial clusters – spatial pattern regions

Unconstrained clustering of the 2489 landscapes into three groups revealed that the nearby metric space is very highly related to nearness in geographic space in southern Quebec. The clusters calculated using only landscape metric values were largely homogeneously grouped in geographic space, rather than randomly located within the province (figure 13). With very few exceptions, nearby landscapes were characterized by similar patterns.

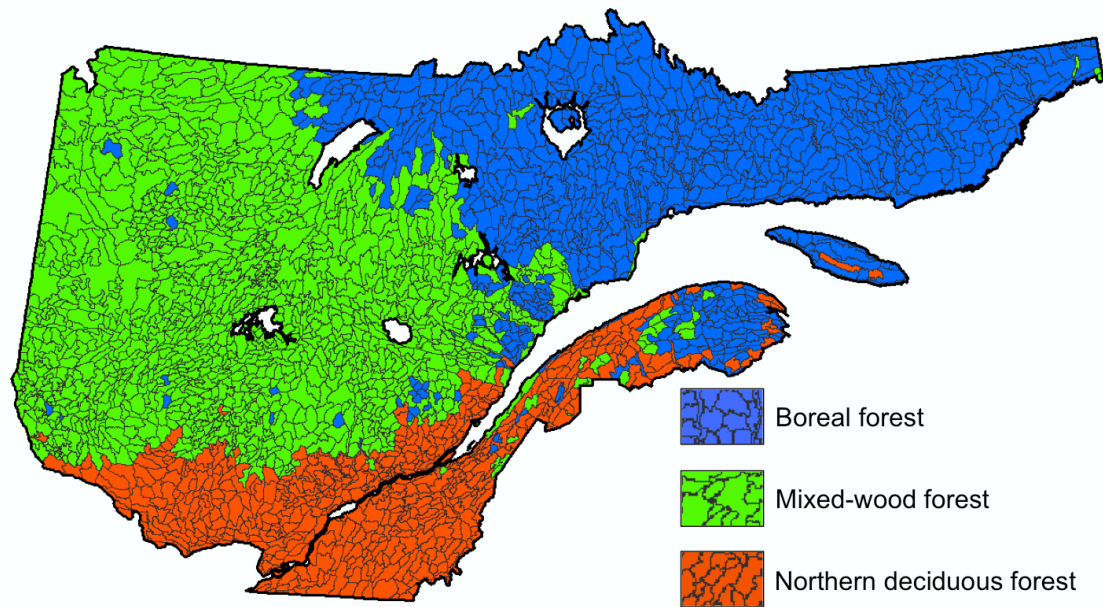


Figure 13. Unconstrained clustering of three clusters calculated with affinity propagation using landscape pattern metrics. Clusters are homogeneously grouped in geographic space. Results are clearly not demonstrating the same spatial structure as the MRNF vegetation subzones. It is revealing that northernmost landscape patterns differ from west to east and not along the MRNF latitudinal gradient presented in figure 8. Although, this latitudinal transition was visible in the most detailed MRNF map (figure 6, H) in chapter 2.

The results of the unconstrained classification of landscape pattern metrics differ substantially from the established MRNF classification. Whereas the MRNF classification is largely partitioned along a latitudinal gradient (figure 8) (Robitaille & Saucier 1998, MRNF 2010), an unconstrained classification reveals that land-cover patterns of the northernmost landscapes of Southern Quebec primarily differ from west to east (visible as green and, blue clusters in figure 13). Also, the cluster representing the southernmost patterns in the unconstrained clustering (in orange, figure 13) varies from the southernmost vegetation subzones in the MRNF classification (Z11 in figure 8): it is more extended north and, especially, in the east.

Although the west-to-east division in the northernmost landscapes does not appear in the vegetation subzone divisions of figure 8, the MRNF ecological classification is comprised of finer levels of classification that contain a comparable structure. The

bioclimatic subdomains map (figure 6, map H in chapter 2) partitions the northern part of our study area (south of 53°N) into multiple distinctive regions from west to east (notably, the separation from 47°N to 52°N along 70°W line). This land-cover pattern division was seen in our unconstrained clustering as being a fundamental division of Quebec landscape patterns, rather than a fine-grained subdomain-scale feature. In the context of three unconstrained clusters, the division of observed land-cover patterns of the northern part of southern Quebec is as important from west to east as it is with the MRNF classification of eleven bioclimatic subdomains.

3.3.2.2 Spatial fences – Pattern region barriers

Spatial fences identify transition zones of land-cover patterns across Southern Quebec (figure 14). Adjacent landscapes can differ by varying degrees of similarities in patterns; in maps of clusters (figure 13), a border between landscapes that belong to two different clusters represents a division between landscapes of differing types. Darker fences identify strong transition zones in landscape pattern, while the lightest fences represent smoother changes in patterns (figure 14). Gaps between segments represent areas where differences are too slight to be noticeable relative to others in the province.

Spatial fences reveal important geographic information about the similarities and differences in nearby landscape patterns (figure 14). First, there are relatively “pure” regions (figure 14, boxes O and Q), large areas whose land-cover patterns are relatively homogenous, marked by only faint fences. Pure regions constitute areas where, from across the range of clustering outputs, adjacent landscapes were usually or always in the same cluster. In the area surrounding and within box O, many pure regions are located side-to-side and separated by dark fences, meaning that patterns are highly distinctive between pure regions, yet are relatively homogeneous within each region. Such fences correspond with a structure of stark boundaries and homogeneous polygons, similar to the MRNF vegetation subzones (figure 8). Second, “mixed” regions are characterized by a high variety of pattern types over small areas (figure 14, boxes P, R, and S). Regions may be mixed for several reasons. For instance, box P presents a very mixed region constituted of highly concentrated dark fences, which means not only that adjacent landscapes are different, but

also they are highly dissimilar in a relatively small space. Adjacent landscapes in boxes R and S are also different, but their dissimilarities are less stark—that is, they are more gradual in geographic space. We can also note that the whole northern part corresponding to the continuous boreal forest (Z21 in figure 8) is not a large uniform area as depicted by the MRNF, since dark fences are partitioning it from west to east (compare figure 14, box R with figure 13).

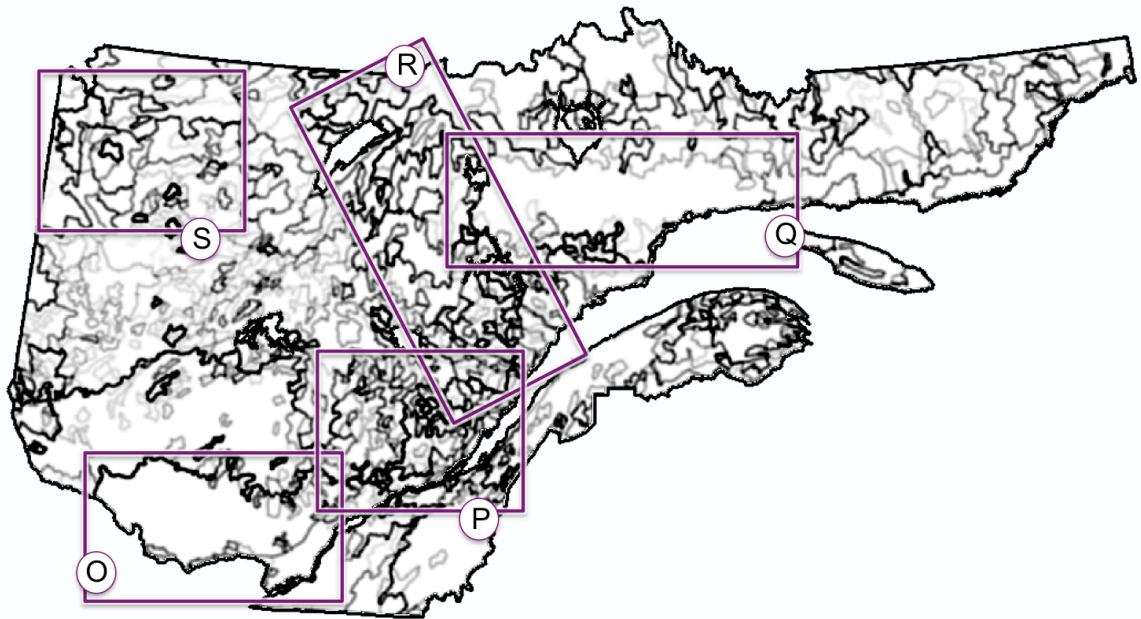


Figure 14. Spatial fences generated from repeated clustering iterations. Spatial fences represent persistent geographic borders between spatial clusters across multiple levels of clustering. Spatial fences need not form closed polygons since barriers are not necessarily continuously spread in space. Darker prominent fences are the most important, for instance segments that reoccurred at each level, while lighter gray segments represent less important differences between adjacent landscapes.

3.3.2.3 Spatial cluster diversity map

Spatial fences revealed variability in the differences between adjacent landscapes, and the diversity map (figure 15) shows that the number of adjacent pattern types cannot be easily associated with any broad-scale geographic structure in southern Quebec. The density of spatial fences is variable in space, which means the diversity of pattern types are

variable and figure 15 illustrates a “purity” classification of spatial clusters throughout multiple clustering levels in Southern Quebec. Pure regions present areas where scores were between 23 and 36. Typically, these scores meant only one or two types of spatial cluster were within 30 km by 30 km blocks for each clustering level. In contrast, mixed regions characterized areas where 4 or 5 spatial cluster types were found within 30 km by 30 km blocks for each clustering level (scores 57 to 103); and in between, blocks locate “semi-pure” areas (scores 37 to 56). Landscapes in pure regions are likely to be surrounded by neighbour landscapes of the same cluster type – i.e., by similar patterns. On the other hand, landscapes in heavily mixed regions are geographically near many other different landscape types. There is no major trend except that the majority of mixed regions are mostly located in the central part of the study area.

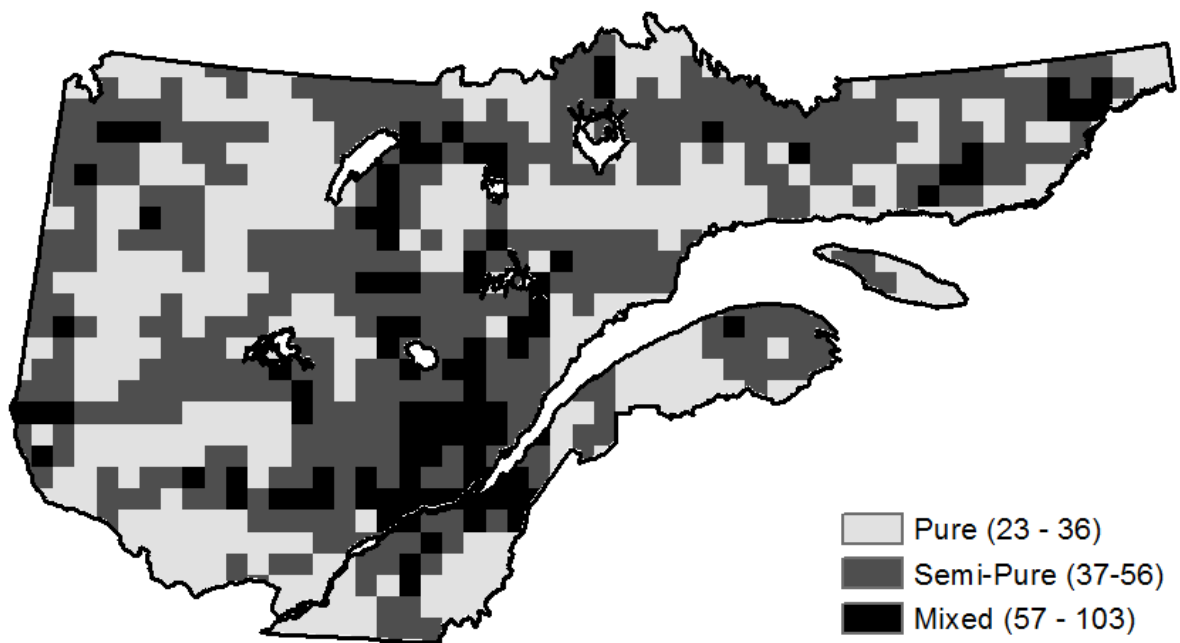


Figure 15. The spatial cluster diversity map characterizes the purity of the landscapes within 30 km by 30 km blocks. Diversity scores were classified into 3 classes according to the Jenks natural breaks. It depicts how many different types of spatial clusters are in each cell. Lighter cells represent regions where adjacent landscapes tend to be placed into the same cluster. Darker cells illustrate where adjacent landscapes were often placed into different clusters.

3.3.3 Comparing constrained and unconstrained classifications of Quebec landscapes

Comparing the constrained classification inconsistencies and the pureness of landscape patterns obtained with the unconstrained clustering can shed additional light on the patterns of landscapes seen in the EOSD and captured by the MRNF. There is no clear relationship between the inconsistencies of the MRNF classification and the diversity of adjacent landscape pattern types (figure 16). Results in figure 16 present the diversity of landscape pattern types (figure 15) masked with the constrained classification inconsistencies (figure 12). The first thing to note is that no inconsistent landscape with the predicted MRNF zones is exclusively pure or mixed. No relationship can therefore be derived from the degree of pureness in unconstrained classification and the results of constrained classification inconsistencies. “Pure inconsistencies” are consequently not random classification mismatches from our prediction but rather regions with neighbouring landscape patterns that were consistently seen as differing from the MRNF definition. “Mixed inconsistencies” imply a high number of different landscape pattern types within small area, and consequently results may capture a higher level of details than the MRNF does at this scale.

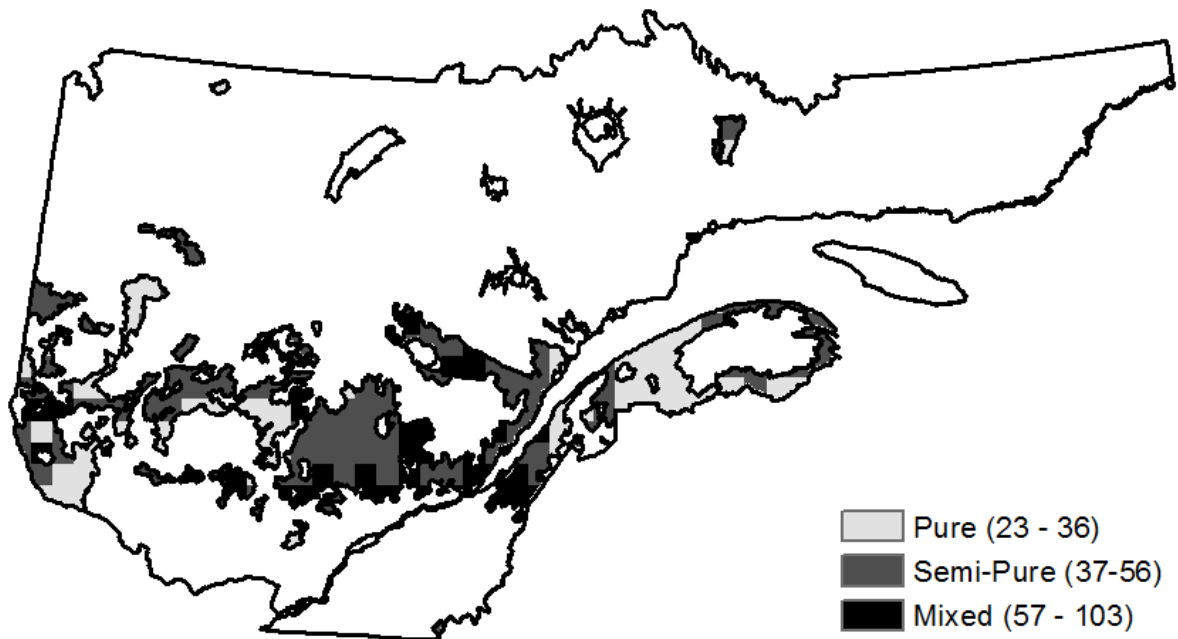


Figure 16. Diversity of unconstrained spatial pattern regions within constrained classification inconsistencies. As seen in figure 15, pure regions, semi-pure regions, and mixed regions are exposed but here in the context of constrained classification inconsistency (mask, figure 12). Inconsistencies do not have any relation with the relative diversity of landscape pattern types. All pure, semi-pure and mixed regions are found within the inconsistency polygon.

3.4 Discussion

3.4.1 Constrained clustering

Using classification trees to constrain land-cover patterns into the MRNF classification is a useful way to identify the metrics that can most clearly distinguish the land-cover conditions being mapped. Classification trees produced reasonable and illustrative fits (table 1 and table 2) while keeping the number of metrics relatively low and easy to interpret (figure 9 and figure 10). For example, the three-leaf tree classified the majority of landscapes in the predicted subzones with two metrics of coniferous forest density. With the nine-leaf tree, the number of different landscape metrics remained low, with four different indices necessary to differentiate the land-cover patterns of the three

vegetation subzones of southern Quebec (patch density, percentage of landscape, aggregation index and mean patch area). Classification trees showed that the land cover categories distinguish MRNF vegetation subzones: seven of the twenty categories were highlighted as predictors for the nine-leaf classification (coniferous-open, coniferous-sparse, mixed-wood dense, water, shrubs tall, herbs, and broadleaf-open).

Constrained classification showed that a small set of landscape metrics could reveal new relevant information on where the observed land-cover patterns and the expected vegetation subzones patterns differ consistently. At first glance, misclassifications could have been considered as reoccurring classification “errors” rather than actual relevant information. In this context, however, misclassifications identified the landscapes in southern Quebec that are the most unlike others in their MRNF subzone (figure 12). These uncertain landscapes indicate places where the conditions are likely far from the typical landscape patterns within a subzone according to their land-cover patterns. The spatial distribution of inconsistencies is not random: they are almost all within and along the MRNF mixed-wood forest zone (figure 11 and figure 12). The mixed-wood region is a natural transition zone between the northern deciduous forest and the continuous boreal forest (MRNF 2010), which appears to relate closely to the results. Landscapes of the transition zone were more frequently unlike their neighbours because they are part of a zone where two different pattern types, the northern deciduous and the continuous boreal forests, are mixed under the same macro-environmental conditions (Bailey 2004). Moreover, this zone is disturbed by major logging activities which will have impacts on metric values, notably in diversity and complexity measures (Boucher et al. 2009). It is an interesting future question to explore whether misclassifications in the constrained classification are more a consequence of the natural environment of the transition zone (a mixed of both surrounding regions), or rather a product of current or historical land-use.

Although the clustering processes performed here gave identical labels to all items in the clusters (table 1), landscape patterns inside a given cluster can vary considerably. Generally, landscapes that were situated far from the MRNF mixed-wood forest tended to be easier to classify (higher accuracy with less metrics) while landscapes located near or

within the mixed-wood subzone were harder to classify. The capacity of distinguishing landscapes with metrics is variable in geographic space. As a result, not all landscape pattern types can be as readily distinguished by using the few same metrics.

3.4.2 Unconstrained clustering

An extensive set of landscape metrics and the capacity of clustering algorithms to group similar landscapes without reference to existing classifications can help to improve our understanding of landscape patterns across southern Quebec. Spatial clusters exhibit a snapshot of circa 2000 landscape patterns, which substantially differ from the more idealized patterns of the MRNF bioclimatic subzones. Since landscape patterns at any point in time are not necessarily in stable climax conditions, this could explain why we detect a major difference between western and the eastern parts of the study area, which differs from the MRNF assessment of a fundamental latitudinal gradient (Robitaille & Saucier 1998). Departure from theoretical conditions is not surprising where substantial anthropogenic disturbances have occurred, as in the deciduous and mixed-wood forest vegetation subzones (Pan et al. 2001, Boucher et al. 2006, Boucher et al. 2009).

A quick comparison between a more detailed MRNF classification level (figure 6, map H in chapter 2) and our unconstrained clustering set of three (figure 13) showed the differences of classifying landscapes in the context of predicted or observed patterns. When considering observed land-cover patterns, the distinction between west and east in the northern part of the study area was stronger than the equivalent bioclimatic factors that theoretically drives the patterns to be divided in a latitudinal gradient. This means that the factors used in the more detailed bioclimatic subdomains to separate landscapes in the west to east manner may be worth exploring since they potentially have a substantial impact on observed patterns, perhaps at the scale of the bioclimatic latitudinal gradient.

Unconstrained clustering indicates that an extensive set of landscape metrics processed with the affinity propagation clustering algorithm has great potential for broad-scale ecological studies. Transforming metrics into the set's principal components and processing them with AP were efficient methods in eliminating many of the known issues

of working with multiple landscape metrics (Riitters et al. 1995, Cain et al. 1997). The mapping of clusters in geographic space adds to the approach of Cardille *et al.* (2010, 2012) and demonstrates the ease and success with which clustering of landscape metrics can give context to land-cover patterns across a large area.

The unconstrained clustering was successful in classifying landscape patterns with metrics; this flexible technique could be adapted for other attributes and needs. As pointed out by Hargrove and Hoffman (2004), and Cardille *et al.* (2012), a clear strength of unconstrained clustering techniques is their objectivity and their independence of geographic constraints. Since this unconstrained classification method is quantitative, flexible and repeatable, extensions of this approach would be straightforward. For example, a variable number of clusters could be computed for other research purposes, while incorporating other sources of ancillary data for particular research needs.

3.4.3 Spatial fences and the diversity map

The spatial fences map reveals the spatial variation in landscape patterns across southern Quebec (figure 14), which can further improve our understanding of land-cover patterns distribution. In contrast to many ecological classifications where patterns are delineated with discrete and contiguous polygons, the spatial fences map expresses landscape patterns as a phenomenon that can vary with different intensity across space. The fences identify transition points in landscape patterns, and could be used to find study areas where there is a concentrated mix of landscape patterns, to prioritize and choose among candidate sampling sites.

Meanwhile, spatial fences (figure 14) and diversity maps (figure 15) illustrated the homogeneity and heterogeneity of regional landscape patterns. Pure regions, representing where neighbouring landscapes are similar, are not uniformly spread throughout southern Quebec. Rather, many regions in the study area have differing adjacent landscape patterns. This showed how complex landscape patterns are structured at this scale: specific degrees of diversity are not unique to any of the vegetation subzones. To help understand this

spatial variability further research should be conducted to isolate potential drivers of pattern diversity.

There was no clear relationship between pattern diversity (figure 15) and the MRNF inconsistencies (figure 12), but at least two different structures might explain inconsistencies. First, ecological districts delineation on the MRNF map was much coarser than satellite observed patterns. So, broader-scale factors drove patterns of inconsistent landscapes and their adjacent landscapes to be all out of the expected geographic expected MRNF limits and were not captured as is at the time of the map creation, these were referred to as pure inconsistencies (figure 15 and figure 16). Secondly, highly heterogeneous landscape pattern types found in adjacent landscapes were hard to fit in discrete and contiguous polygons; such areas were referred to as mixed inconsistencies (figure 15 and figure 16). Finer-scale factors may have induced the patterns to drastically change patterns locally. Further investigations, and potentially field surveys, are needed to elucidate the reasons of such variability in landscape patterns diversity.

3.5 Conclusion

The two clustering approaches in this study successfully employed landscape metrics to classify patterns across southern Quebec, showing the degree to which the MRNF subzones match observed land-cover patterns. Results in this research suggest where and how this well-known broad-scale ecological classification can contain considerable variability with respect to the observed patterns measured by landscape metrics and exclusively forest cover data. These related approaches contribute to the growing literature on large-region classifications by extending the use of landscape pattern indices as quantitative environmental descriptors. As discussed by Cardille *et al.* (2010, 2012), it further demonstrates the power of pattern indices to successfully discriminate patterns for broad-scale studies.

The constrained clustering showed that two metrics were enough to distinguish most landscapes of the three subzones of southern Quebec. Classification trees were capable of extracting meaning from and greatly simplifying the large and complex landscape metric

data set. Indeed, relatively simple trees were capable of correctly reproducing an existing ecological classification map. Inconsistencies between these two assessments revealed fundamental differences between the two classifications that can be explored in future work. Prediction with the fitted trees revealed certain landscapes that were frequent outliers with respect to other landscapes in the same established bioclimatic subzone.

The unconstrained clustering showed that the full set of landscape metrics can reveal regional land-cover patterns that differ fundamentally from the bioclimatic subzones frequently used in science and management in this region. This can be due to several reasons, including human and perhaps natural disturbances and the limitations of polygon-based cartography. This research extended the clustering capacity of affinity propagation as outlined in Cardille *et al.* (2010, 2012) by presenting and exploring the resulting clusters in geographic space. Through spatial fences, results gave a new perspective on how we can delineate patterns over large areas, especially for landscape patterns. It expressed the spatial variability and complexity present in the different landscape patterns across the study area.

Comparison of the results with other similar studies should be made with caution since landscape metric values are dependent, for example, on the land cover product classification scheme, the spatial resolution, and the spatial extent of study units. Results presented in this study characterized landscape patterns of circa 2000 according to the EOSD land cover product and should not be considered as representing exact conditions of the present day. Nevertheless, the EOSD is still the most detailed product for our study area, and it will remain a relevant data set to use until the next national-scale updates are made. The upcoming replacement of the actual Landsat-7 platform, the Landsat Data Continuity Mission (NASA 2012a), should eventually offer quality data for the renewal of this representation of the forested landscape cover of Canada.

The methods of both unconstrained and constrained classification could be adapted for other needs and applications. For example, other existing types of attributes (e.g., soil types, topography, precipitation regimes) for landscapes could be integrated to extend this data set and to diversify the classification needs or goals. By identifying relevant

discriminating factors with classification trees, our flexible unconstrained approach could easily be used to group landscapes for any given subset of relevant metrics. Comparing both southern Quebec land-cover and vegetation subzones clarified the differences between observed patterns and late-successional vegetation stages in southern Quebec. Results of both classification methods open to a new perception of broad-scale maps in southern Quebec.

4. Conclusion générale du mémoire

Le Québec en entier autant que sa partie méridionale en soit est immense et les résultats présentés dans les chapitres 2 et 3 montrent conséquemment que les paysages québécois sont d'une complexité et d'une diversité sans équivoque. Ce manuscrit n'a fait qu'effleuré cette réalité, mais il a toutefois permis de mettre en lumière que les paysages, particulièrement dans les régions les plus méridionales, démontrent des divergences entre les descriptions existantes basées sur les prédictions de conditions bioclimatiques et physiques et nos résultats observés. De plus, les patrons des paysages septentrionaux se sont révélés plus complexes et beaucoup plus diversifiés comparativement à ce qui est communément entendu dans la classification de référence. Effectivement, le portrait détaillé des paysages du Québec présenté dans ce mémoire met en évidence l'importance de la variabilité spatiale des patrons et donc, de la variabilité spatiale des facteurs qui sculptent les patrons de paysages – peu importe les processus écologiques, l'origine de perturbations et l'emplacement géographique.

Un nouveau genre d'information géographique a été généré à travers les résultats et analyses de cette recherche et ceux-ci contrastent avec les classifications écologiques classiques. Comme discuté dans les chapitres précédents, les paysages sont caractérisés par une grande variabilité autant par leurs patrons que par la distribution spatiale de ceux-ci. Les cartes de représentativité, de frontières persistantes et de complexité en sont des démonstrations claires. Ainsi, la définition typique de zones en polygones discrets cache une réelle facette de la transition des paysages sur un territoire aussi complexe que celui du Québec. L'ajout d'information – encore marginal jusqu'à ce jour - sur l'homogénéité et sur les divergences des classifications apporte une autre dimension aux résultats. De plus, comme la plupart des classifications ne mettent pas autant d'accent sur les patrons anthropiques, et ce, même s'il s'agit de la réalité de certaines régions, nous pensons qu'il y a un manque potentiel de représentativité des conditions réelles. En ce sens, cette recherche contribue à une connaissance géographique représentative de la variabilité des patrons à l'échelle du Québec.

Les indices de patrons ont démontré un énorme potentiel d'utilisation pour les futures classifications écologiques à petite échelle et tant que les phénomènes environnementaux aux échelles nationales et continentales continueront de captiver l'intérêt des scientifiques, des méthodes facilement adaptables et efficaces comme celles présentées dans ce mémoire seront nécessaires. Malgré le développement des indices de patrons depuis les années 1980, leur potentiel scientifique a été peu investigué pour régionaliser les paysages sur de vastes territoires comme celui du Québec, mais les résultats de cette recherche ont révélé leurs capacités et utilités. Au départ, les mises en garde par rapport à leur utilisation de même que la complexité qu'engendre l'analyse des paysages à cette échelle auront sûrement freiné, en partie, ce type de recherche. Par contre, les méthodes efficaces proposées dans les chapitres précédents et, surtout, les résultats convaincants démontrent bien le grand potentiel de ces indices comme attributs biogéographiques pertinents. N'en demeure pas moins que l'usage des indices tels que présentés dans ce mémoire devrait être réitéré afin de renforcer sa robustesse. Cette application vient s'ajouter à la liste grandissante d'études d'envergure utilisant les indices de patrons comme données.

Ce projet a, tout d'abord, été inspiré par d'autres recherches et pourrait ouvrir la porte à plusieurs autres. Dans la même lignée que les frontières persistantes, des « super-exemplars » pourraient être identifiés parmi les multiples niveaux de groupages obtenus, c'est-à-dire de sélectionner les paysages représentatifs qui sont identifiés de fois en fois, peu importe le nombre de groupes calculés. Les super-exemplars identifiés pourraient être ajoutés à la carte des frontières persistantes pour donner une synthèse complète de l'information contenue par les indices de patrons. Ensuite, la création d'un réseau de paysages représentatifs pourrait être établie à l'échelle du Québec, soit avec les super-exemplars, ou encore avec un nombre préétabli d'exemplars. Pour demeurer cohérentes avec le système de classification écologique déjà en place, la régionalisation des patrons et l'identification simultanée des paysages représentatifs pourraient être effectuées pour chaque sous-zone de végétation du MRNF (ou tout autre niveau de la classification) individuellement afin de peindre les patrons réels à l'intérieur même de chaque zone.

Finalement, nous croyons qu'une campagne de terrain devrait être mise de l'avant afin de valider nos résultats et, par exemple, pour comparer les relevés de faune et de flore entre nos zones résultantes et les zones du MRNF.

Certains auteurs ont présenté les avantages qu'offrent les méthodes quantitatives de régionalisation des paysages telles qu'employées dans ce mémoire et nous appuyons ces observations. Même si nous n'avons utilisé que des indices de patrons, plusieurs autres types de variables pertinentes pourraient bien évidemment être intégrées et servir pour des recherches spécifiques, il s'agit d'un avantage de la flexibilité de cette approche. La reproductibilité des exercices de ce mémoire ouvre la porte à des études menées sur les autres provinces du Canada ainsi qu'un suivi temporel des paysages dans le cas d'un éventuel renouvellement d'une couverture du sol mise à jour avec le lancement du nouveau capteur LDCM (Landsat Data Continuity Mission).

Bibliographie

- Bailey, R.G. 2004. Identifying ecoregion boundaries. *Environmental Management*, **34**: S14-S26. doi: 10.1007/s00267-003-0163-6.
- Bergeron, Y., Gauthier, S., Kafka, V., Lefort, P. & Lesieur, D. 2001. Natural fire frequency for the eastern Canadian boreal forest: consequences for sustainable forestry. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, **31**(3): 384-391. doi: 10.1139/cjfr-31-3-384.
- Bodenhofer, U., Kothmeier, A. & Hochreiter, S. 2011. APCluster: an R package for affinity propagation clustering. *Bioinformatics*, **27**: 2463-2464. doi: 10.1093/bioinformatics/btr406.
- Boucher, Y., Arseneault, D. & Sirois, L. 2006. Logging-induced change (1930-2002) of a preindustrial landscape at the northern range limit of northern hardwoods, eastern Canada. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, **36**(2): 505-517. doi: 10.1139/x05-252.
- Boucher, Y., Arseneault, D., Sirois, L. & Blais, L. 2009. Logging pattern and landscape changes over the last century at the boreal and deciduous forest transition in Eastern Canada. *Landscape Ecology*, **24**(2): 171-184. doi: 10.1007/s10980-008-9294-8.
- Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J. 1984. Classification and regression trees. Chapman & Hall/CRC, Belmont, California, USA. pp.358.
- Cain, D.H., Riitters, K. & Orvis, K. 1997. A multi-scale analysis of landscape statistics. *Landscape Ecology*, **12**(4): 199-212.
- Cardille, J., Turner, M., Clayton, M., Gergel, S. & Price, S. 2005. METALAND: Characterizing spatial patterns and statistical context of landscape metrics. *Bioscience*, **55**(11): 983-988.
- Cardille, J.A. & Clayton, M.K. 2007. A regression tree-based method for integrating land-cover and land-use data collected at multiple scales. *Environmental and Ecological Statistics*, **14**(2): 161-179.
- Cardille, J.A. & Foley, J.A. 2003. Agricultural land-use change in Brazilian Amazonia between 1980 and 1995: evidence from integrated satellite and census data. *Remote Sensing of Environment*, **87**(4): 551-562.
- Cardille, J.A., Foley, J.A. & Costa, M.H. 2002. Characterizing patterns of agricultural land use in Amazonia by merging satellite classifications and census data. *Global Biogeochemical Cycles*, **16**(3): 17pp. doi: 10.1029/2000GB001386.

- Cardille, J.A. & Lambois, M. 2010. From the redwood forest to the Gulf Stream waters: human signature nearly ubiquitous in representative US landscapes. *Frontiers in Ecology and the Environment*, **8**(3): 130-134. doi: 10.1890/080132.
- Cardille, J.A., White, J.C., Wulder, M.A. & Holland, T. 2012. Representative landscapes in the forested area of Canada. *Environmental Management*, **49**(1): 163-173.
- 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.
- Cushman, S.A., McGarigal, K. & Neel, M.C. 2008. Parsimony in landscape metrics: strength, universality, and consistency. *Ecological Indicators*, **8**(5): 691-703.
- de Blois, S., Domon, G. & Bouchard, A. 2001. Environmental, historical, and contextual determinants of vegetation cover: a landscape perspective. *Landscape Ecology*, **16**(5): 421-436. doi: 10.1023/a:1017548003345.
- De Clercq, E.M., Vandemoortele, F. & De Wulf, R.R. 2006. A method for the selection of relevant pattern indices for monitoring of spatial forest cover pattern at a regional scale. *International Journal of Applied Earth Observation and Geoinformation*, **8**(2): 113-125. doi: 10.1016/j.jag.2005.07.002.
- De'ath, G. & Fabricius, K.E. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology*, **81**(11): 3178-3192. doi: 10.1890/0012-9658(2000)081[3178:cartap]2.0.co;2.
- Dueck, D., Frey, B.J. & Ieee. 2007. Non-metric affinity propagation for unsupervised image categorization. Pages 198-205 in 11th IEEE International Conference on Computer Vision, Rio de Janeiro, BRAZIL.
- Ellis, E.C. & Ramankutty, N. 2008. Putting people in the map: anthropogenic biomes of the world. *Frontiers in Ecology and the Environment*, **6**(10): 439-447.
- ESRI. 2011. ArcGIS Desktop, release 10. Environmental Systems Research Institute, Redlands, CA.
- Fahrig, L. 2003. Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology Evolution and Systematics*, **34**: 487-515. doi: 10.1146/annurev.ecolsys.34.011802.132419.
- Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N. & Snyder, P.K. 2005. Global consequences of land use. *Science*, **309**(5734): 570-574. doi: 10.1126/science.1111772.

- Fortin, M.J., Boots, B., Csillag, F. & Rimmel, T.K. 2003. On the role of spatial stochastic models in understanding landscape indices in ecology. *Oikos*, **102**(1): 203-212. doi: 10.1034/j.1600-0706.2003.12447.x.
- Fortin, M.J. & Edwards, G. 2001. Delineation and analysis of vegetation boundaries. *Spatial Uncertainty in Ecology*: 158-+.
- Frey, B.J. & Dueck, D. 2007. Clustering by passing messages between data points. *Science*, **315**(5814): 972-976. doi: 10.1126/science.1136800.
- Friedlaender, A.S., Halpin, P.N., Qian, S.S., Lawson, G.L., Wiebe, P.H., Thiele, D. & Read, A.J. 2006. Whale distribution in relation to prey abundance and oceanographic processes in shelf waters of the Western Antarctic Peninsula. *Marine Ecology-Progress Series*, **317**: 297-310. doi: 10.3354/meps317297.
- Girvetz, E.H., Thorne, J.H., Berry, A.M. & Jaeger, J.A.G. 2008. Integration of landscape fragmentation analysis into regional planning: A statewide multi-scale case study from California, USA. *Landscape and Urban Planning*, **86**(3-4): 205-218. doi: 10.1016/j.landurbplan.2008.02.007.
- Grimm, N.B., Foster, D., Groffman, P., Grove, J.M., Hopkinson, C.S., Nadelhoffer, K.J., Pataki, D.E. & Peters, D.P.C. 2008. The changing landscape: ecosystem responses to urbanization and pollution across climatic and societal gradients. *Frontiers in Ecology and the Environment*, **6**(5): 264-272. doi: 10.1890/070147.
- Gustafson, E.J. 1998. Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems*, **1**(2): 143-156. doi: 10.1007/s100219900011.
- Hansen, M.C., Defries, R.S., Townshend, J.R.G. & Sohlberg, R. 2000. Global land cover classification at 1km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, **21**(6-7): 1331-1364. doi: 10.1080/014311600210209.
- Hargrove, W.W. & Hoffman, F.M. 1999. Using multivariate clustering to characterize ecoregion borders. *Computing in Science & Engineering*, **1**(4): 18-25. doi: 10.1109/5992.774837.
- Hargrove, W.W. & Hoffman, F.M. 2004. Potential of multivariate quantitative methods for delineation and visualization of ecoregions. *Environmental Management*, **34**: S39-S60. doi: 10.1007/s00267-003-1084-0.
- Herzog, F. & Lausch, A. 2001. Supplementing land-use statistics with landscape metrics: Some methodological considerations. *Environmental Monitoring and Assessment*, **72**(1): 37-50. doi: 10.1023/a:1011949704308.

- Holling, C.S. 1992. Cross-scale morphology, geometry, and dynamics of ecosystems. *Ecological Monographs*, **62**(4): 447-502. doi: 10.2307/2937313.
- Homer, C., Huang, C.Q., Yang, L.M., Wylie, B. & Coan, M. 2004. Development of a 2001 National Land-Cover Database for the United States. *Photogrammetric Engineering and Remote Sensing*, **70**(7): 829-840.
- Hutchinson, M.F., McIntyre, S., Hobbs, R.J., Stein, J.L., Garnett, S. & Kinloch, J. 2005. Integrating a global agro-climatic classification with bioregional boundaries in Australia. *Global Ecology and Biogeography*, **14**(3): 197-212. doi: 10.1111/j.1466-822x.2005.00154.x.
- ISQ. 2011. Le Québec chiffres en main - Québec Handy Numbers, Edition 2011. Gouvernement du Québec, Institut de la statistique du Québec, 1987, Québec. pp.72.
- Keller, M., Schimel, D.S., Hargrove, W.W. & Hoffman, F.M. 2008. A continental strategy for the National Ecological Observatory Network. *Frontiers in Ecology and the Environment*, **6**(5): 282-284.
- Kelling, S., Hochachka, W.M., Fink, D., Riedewald, M., Caruana, R., Ballard, G. & Hooker, G. 2009. Data-intensive Science: A New Paradigm for Biodiversity Studies. *Bioscience*, **59**(7): 613-620. doi: 10.1525/bio.2009.59.7.12.
- Kie, J.G., Bowyer, R.T., Nicholson, M.C., Boroski, B.B. & Loft, E.R. 2002. Landscape heterogeneity at differing scales: Effects on spatial distribution of mule deer. *Ecology*, **83**(2): 530-544. doi: 10.2307/2680033.
- King, R.S., Richardson, C.J., Urban, D.L. & Romanowicz, E.A. 2004. Spatial dependency of vegetation-environment linkages in an anthropogenically influenced wetland ecosystem. *Ecosystems*, **7**(1): 75-97. doi: 10.1007/s10021-003-0210-4.
- Kupfer, J.A. 2006. National assessments of forest fragmentation in the US. *Global Environmental Change-Human and Policy Dimensions*, **16**(1): 73-82. doi: 10.1016/j.gloenvcha.2005.10.003.
- Latifovic, R., Zhu, Z.L., Cihlar, J., Giri, C. & Olthof, I. 2004. Land cover mapping of north and central America - Global Land Cover 2000. *Remote Sensing of Environment*, **89**(1): 116-127. doi: 10.1016/j.rse.2003.11.002.
- Lausch, A. & Herzog, F. 2002. Applicability of landscape metrics for the monitoring of landscape change: issues of scale, resolution and interpretability. *Ecological Indicators*, **2**(1-2): 3-15.

- Lawler, J.J. & Edwards, T.C. 2002. Landscape patterns as habitat predictors: building and testing models for cavity-nesting birds in the Uinta Mountains of Utah, USA. *Landscape Ecology*, **17**(3): 233-245. doi: 10.1023/a:1020219914926.
- Leone, M., Sumedha & Weigt, M. 2007. Clustering by soft-constraint affinity propagation: applications to gene-expression data. *Bioinformatics*, **23**(20): 2708-2715. doi: 10.1093/bioinformatics/btm414.
- Li, H.B. & Wu, J.G. 2004. Use and misuse of landscape indices. *Landscape Ecology*, **19**(4): 389-399.
- Lillesand, T.M., Kiefer, R.W. & Chipman, J.W. 2008. Remote sensing and image interpretation, 6th Edition. John Wiley & Sons, Inc. pp.756.
- Long, J., Nelson, T. & Wulder, M. 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-S13. doi: 10.1007/s00267-003-5181-x.
- 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**(6-7): 1303-1330. doi: 10.1080/014311600210191.
- Mackey, B.G., Berry, S.L. & Brown, T. 2008. 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.
- Mackey, B.G., Nix, H.A., Hutchinson, M.F., Macmahon, J.P. & Fleming, P.M. 1988. Assessing representativeness of places for conservation reservation and heritage listing. *Environmental Management*, **12**(4): 502-514.
- McGarigal, K., Cushman, S.A., Neel, M.C. & Ene, E. 2002. FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps, version 3.3. Computer software program produced by the authors at the University of Massachusetts, Amherst. <http://www.umass.edu/landeco/research/fragstats/fragstats.html>.
- McIntire, E.J.B. 2004. Understanding natural disturbance boundary formation using spatial data and path analysis. *Ecology*, **85**(7): 1933-1943. doi: 10.1890/03-0579.
- Mezard, M. 2007. Computer science - Where are the exemplars? *Science*, **315**(5814): 949-951. doi: 10.1126/science.1139678.

- Monfreda, C., Ramankutty, N. & Foley, J.A. 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, **22**(1): GB1022. doi: 10.1029/2007GB002947.
- MRN. 2000. La limite nordique des forêts attribuables. Ministère des Ressources Naturelles, Gouvernement du Québec. pp.101.
- MRNF. 2005. Système hiérarchique de classification écologique – Données spatiales numériques. Direction des inventaires forestiers. Ministère des Ressources naturelles et de la Faune du Québec.
- MRNF. 2010. Zones de végétation et domaines bioclimatiques du Québec [En ligne]. <http://www.mrnf.gouv.qc.ca/forets/connaissances/connaissances-inventaire-zones-carte.jsp - haut>. [Accédé le 9 juillet 2012].
- NASA. 2012a. National Aeronautics and Space Administration, LANDSAT Data Continuity Mission [En ligne]. <http://ldcm.nasa.gov/>. [Accédé le 30 avril 2012].
- NASA. 2012b. National Aeronautics and Space Administration, Landsat program [En ligne]. <http://landsat.gsfc.nasa.gov/about/>. [Accédé le 16 mai 2012].
- O'Neill, R.V., Krummel, J.R., Gardner, R.H., Sugihara, G., Jackson, B., DeAngelis, D.L., Milne, B.T., Turner, M.G., Zygmunt, B., Christensen, S.W., Dale, V.H. & Graham, R.L. 1988. Indices of landscape pattern. *Landscape Ecology*, **1**(3): 153-162. doi: 10.1007/bf00162741.
- O'Neill, R.V., Riitters, K.H., Wickham, J.D. & Jones, K.B. 1999. Landscape pattern metrics and regional assessment. *Ecosystem Health*, **5**(4): 225-233. doi: 10.1046/j.1526-0992.1999.09942.x.
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. & Kassem, K.R. 2001. Terrestrial ecoregions of the worlds: A new map of life on Earth. *Bioscience*, **51**(11): 933-938. doi: 10.1641/0006-3568(2001)051[0933:teotwa]2.0.co;2.
- Olthof, I., Latifovic, R. & Pouliot, D. 2009. Development of a circa 2000 land cover map of northern Canada at 30 m resolution from Landsat. *Canadian Journal of Remote Sensing*, **35**(2): 152-165.
- Pan, D., Domon, G., de Blois, S. & Bouchard, A. 1999. Temporal (1958 to 1993) and spatial patterns of land use changes in Haut-Saint-Laurent (Quebec, Canada) and their relation to landscape physical attributes. *Landscape Ecology*, **14**(1): 35-52.

- Pan, D., Domon, G., Marceau, D. & Bouchard, A. 2001. Spatial pattern of coniferous and deciduous forest patches in an Eastern North America agricultural landscape: the influence of land use and physical attributes. *Landscape Ecology*, **16**(2): 99-110. doi: 10.1023/a:1011167905097.
- Payette, S., Bhiry, N., Delwaide, A. & Simard, M. 2000. Origin of the lichen woodland at its southern range limit in eastern Canada: the catastrophic impact of insect defoliators and fire on the spruce-moss forest. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, **30**(2): 288-305. doi: 10.1139/cjfr-30-2-288.
- Peters, D.P.C., Bestelmeyer, B.T. & Turner, M.G. 2007. Cross-scale interactions and changing pattern-process relationships: Consequences for system dynamics. *Ecosystems*, **10**(5): 790-796. doi: 10.1007/s10021-007-9055-6.
- Peters, D.P.C., Groffman, P.M., Nadelhoffer, K.J., Grimm, N.B., Coffins, S.L., Michener, W.K. & Huston, M.A. 2008. Living in an increasingly connected world: a framework for continental-scale environmental science. *Frontiers in Ecology and the Environment*, **6**(5): 229-237. doi: 10.1890/070098.
- Peters, D.P.C., Pielke, R.A., Bestelmeyer, B.T., Allen, C.D., Munson-McGee, S. & Havstad, K.M. 2004. Cross-scale interactions, nonlinearities, and forecasting catastrophic events. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(42): 15130-15135. doi: 10.1073/pnas.0403822101.
- Prince, S.D. & Steininger, M.K. 1999. Biophysical stratification of the Amazon basin. *Global Change Biology*, **5**(1): 1-22. doi: 10.1046/j.1365-2486.1998.00220.x.
- Qian, Y., Yao, F. & Jia, S. 2009. Band selection for hyperspectral imagery using affinity propagation. *Iet Computer Vision*, **3**(4): 213-222. doi: 10.1049/iet-cvi.2009.0034.
- Questier, F., Put, R., Coomans, D., Walczak, B. & Heyden, Y.V. 2005. The use of CART and multivariate regression trees for supervised and unsupervised feature selection. *Chemometrics and Intelligent Laboratory Systems*, **76**(1): 45-54. doi: 10.1016/j.chemolab.2004.09.003.
- R Development Core Team. 2011. R: A Language and Environment for Statistical Computing. Vienna, Austria. Logiciel disponible à : <http://www.R-project.org>.
- Ramankutty, N., Evan, A.T., Monfreda, C. & Foley, J.A. 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles*, **22**(1): 1-19 (GB1003). doi: 10.1029/2007GB002952.
- Richard, P.J.H. 1995. Végétation du Québec-Labrador: Formations et grands domaines. Carte mise à jour à partir de la publication du même auteur datant de 1987: Le couvert végétal du Québec-Labrador et son histoire post-glaicière. Notes et

Documents n° 87-01. Département de géographie, Université de Montréal, Montréal, Québec, Canada.

- Riitters, K., Wickham, J., O'Neill, R., Jones, B. & Smith, E. 2000. Global-scale patterns of forest fragmentation. *Conservation Ecology*, **4**(2): 3.
- Riitters, K.H., Oneill, R.V., Hunsaker, C.T., Wickham, J.D., Yankee, D.H., Timmins, S.P., Jones, K.B. & Jackson, B.L. 1995. A factor-analysis of landscape pattern and structure metrics. *Landscape Ecology*, **10**(1): 23-39.
- Riitters, K.H., Wickham, J.D., O'Neill, R.V., Jones, K.B., Smith, E.R., Coulston, J.W., Wade, T.G. & Smith, J.H. 2002. Fragmentation of continental United States forests. *Ecosystems*, **5**(8): 815-822. doi: 10.1007/s10021002-0209-2.
- RNC. 2009. Ressources naturelles Canada: GéoGratis [Données en ligne]. <http://geogratias.cgdi.gc.ca/geogratias/fr/index.html>. [Accédé en 2011].
- Robitaille, A. & Saucier, J.P. 1996. Land district, ecophysiological units and areas: The landscape mapping of the Ministère des Ressources Naturelles du Québec. *Environmental Monitoring and Assessment*, **39**(1-3): 127-148. doi: 10.1007/bf00396141.
- Robitaille, A. & Saucier, J.P. 1998. Paysages régionaux du Québec méridional. Gouvernement du Québec, Ministère des ressources naturelles. pp.213.
- SAFORAH. 2012. System of Agents for Forest Observation Research with Automation Hierarchies [Données en ligne]. <http://www2.saforah.org:7700/>. [Accédé le 1^{er} mai 2012].
- Sarr, D.A., Hibbs, D.E. & Huston, M.A. 2005. A hierarchical perspective of plant diversity. *Quarterly Review of Biology*, **80**(2): 187-212.
- Serra, J.M., Cristobal, J. & Ninyerola, M. 2011. A Classification Procedure for Mapping Topo-climatic Conditions for Strategic Vegetation Planning. *Environmental Modeling & Assessment*, **16**(1): 77-89. doi: 10.1007/s10666-010-9232-4.
- Trzcinski, M.K., Fahrig, L. & Merriam, G. 1999. Independent effects of forest cover and fragmentation on the distribution of forest breeding birds. *Ecological Applications*, **9**(2): 586-593. doi: 10.2307/2641146.
- Turner, M.G. 2005. Landscape ecology: What is the state of the science? *Annual Review of Ecology Evolution and Systematics*, **36**: 319-344. doi: 10.1146/annurev.ecolsys.36.102003.152614.
- Turner, M.G., Gardner, R.H. & O'Neil, R.V. 2001. Landscape ecology in theory and practice : pattern and process, First Edition. Springer-Verlag, New York. pp.404.

- Urban, D.L., Oneill, R.V. & Shugart, H.H. 1987. Landscape Ecology. *Bioscience*, **37**(2): 119-127. doi: 10.2307/1310366.
- Van der Laan, M.J., Pollard, K.S. & Bryan, J. 2003. A new partitioning around medoids algorithm. *Journal of Statistical Computation and Simulation*, **73**(8): 575-584. doi: 10.1080/0094965031000136012.
- Vitousek, P.M., Ehrlich, P.R., Ehrlich, A.H. & Matson, P.A. 1986. Human appropriation of the products of photosynthesis. *Bioscience*, **36**(6): 368-373. doi: 10.2307/1310258.
- Vitousek, P.M., Mooney, H.A., Lubchenco, J. & Melillo, J.M. 1997. Human domination of Earth's ecosystems. *Science*, **277**(5325): 494-499. doi: 10.1126/science.277.5325.494.
- Wade, T.G., Riitters, K.H., Wickham, J.D. & Jones, K.B. 2003. Distribution and causes of global forest fragmentation. *Conservation Ecology*, **7**(2): 7. [En ligne] URL: <http://www.ecologyandsociety.org/vol7/iss2/art7/>.
- Wickham, J.D. & Norton, D.J. 1994. Mapping and analyzing landscape patterns. *Landscape Ecology*, **9**(1): 7-23.
- Wickham, J.D., Riitters, K.H., Wade, T.G. & Homer, C. 2008. Temporal change in fragmentation of continental US forests. *Landscape Ecology*, **23**(8): 891-898. doi: 10.1007/s10980-008-9258-z.
- Wu, J., Jelinski, D.E., Luck, M. & Tueller, P.T. 2000. Multiscale analysis of landscape heterogeneity: scale variance and pattern metrics. *Annals of GIS*, **6**(1): 6-19.
- Wulder, M.A., White, J.C. & Coops, N.C. 2011a. Fragmentation regimes of Canada's forests. *Canadian Geographer/Le Géographe canadien*, **55**(3): 288-300. doi: 10.1111/j.1541-0064.2010.00335.x.
- Wulder, M.A., White, J.C., Cranny, M., Hall, R.J., Luther, J.E., Beaudoin, A., Goodenough, D.G. & Dechka, J.A. 2008a. Monitoring Canada's forests. Part 1: Completion of the EOSD land cover project. *Canadian Journal of Remote Sensing*, **34**(6): 549-562.
- Wulder, M.A., White, J.C., Goward, S.N., Masek, J.G., Irons, J.R., Herold, M., Cohen, W.B., Loveland, T.R. & Woodcock, C.E. 2008b. 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., White, J.C., Han, T., Coops, N.C., Cardille, J.A., Holland, T. & Grills, D. 2008c. Monitoring Canada's forests. Part 2: National forest fragmentation and pattern. *Canadian Journal of Remote Sensing*, **34**(6): 563-584.

- Wulder, M.A., White, J.C., Masek, J.G., Dwyer, J. & Roy, D.P. 2011b. Continuity of Landsat observations: Short term considerations. *Remote Sensing of Environment*, **115**(2): 747-751. doi: 10.1016/j.rse.2010.11.002.
- Xie, X.-n., Wu, F. & Ieee Computer, S.O.C. 2008. Automatic Video Summarization by Affinity Propagation Clustering and Semantic Content Mining. *Proceedings of the International Symposium on Electronic Commerce and Security*: 203-208.
- Zhang, J. & Goodchild, M.F. 2002. *Uncertainty in geographical information*. Taylor & Francis. pp.266.