CHARACTERIZING THE SPATIAL PATTERNS AND SPATIALLY EXPLICIT PROBABILITIES OF POST-FIRE VEGETATION RESIDUAL PATCHES IN BOREAL WILDFIRE SCARS

YIKALO H. ARAYA

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN GEOGRAPHY
YORK UNIVERSITY
TORONTO, ONTARIO

September 2014

ABSTRACT

Wildfire is one of the main natural disturbances that consume a substantial amount of forest cover, influencing and reshaping the landscape mosaic of boreal forests. Wildfires do not burn the entire landscape; they rather create a complex mosaic of post-fire landscape structure with different degrees of burn severity. The resulting spatial mosaic includes fully burned, partially burned, and unburned areas. Even though the most visible components of a fire disturbed landscape are the completely burned areas, a considerable number of residual patches of various size, shape, and composition are retained following a fire. The residual patches refer to remnants of the pre-fire forest ecosystem that left completely unaltered within the fire footprint. Improved understanding of the patterns and characteristics of wildfire residuals provides insights for investigating the effects of fire disturbances, emulating forest disturbances in harvesting operations, and improving forest management planning. Knowledge about the post-fire residuals relies on how well we measure the patterns and characteristics of post-fire residuals, determine the factors that explain their occurrence and patterns, and what consistent measurement framework we use to understand the patterns and predict their likely occurrence. In this study, the patterns and characteristics of post-fire residuals was initially examined based on eleven boreal wildfire events within northwestern Ontario; each ignited by lightning and never suppressed. The wildfire events were occurred in ecoregion 2W during the fire seasons of 2002 and 2003. In order to design a consistent and repeatable method for measuring the patterns of residuals, an integrate approach has been designed. This involves assessing the spatial patterns where the composition, configuration, and fragmentation of residual patches were assessed based on selected spatial metrics; examining the importance of predictor variables that explain residuals and their marginal effects on residual patch occurrence using Random Forest (RF) ensemble method; and developing a spatially explicit predictive model using the RF method where the combined effects of the variables were examined. Finally, the three approaches are applied and evaluated using a recent and independent data from the extensive RED084 wildfire event that occurred in 2011 within the adjacent ecoregion (3S). The effects of analytical scale (i.e., spatial resolution) on characterizing the spatial patterns, determining the relative variable importance, and predicted probabilities of residual patches are assessed. The results show that the composition and configuration of wildfire residuals vary as a function of measurement, spatial resolutions, and fire event sizes, suggesting the variation in fire intensity and severity across the fire events. The patterns of wildfire residuals are also sensitive to changing scale, but the responses of the spatial metrics to changing spatial resolutions are grouped into three categories:

monotonic change and predictable response in which three shape related metrics (LSI, MSI, and FRAC) show a predictable responsible; monotonic change with no simple scaling rule; and non-monotonic change with erratic response. The results also reveal that the factors that are incorporated in this study interactively affect the occurrence and distribution of residual patches, but natural firebreak features (e.g., wetlands and surface water) were among the most important predictors to explain wildfire residuals. Furthermore, the model implemented to predict residual patches has a reasonable or high predictive performance ('marginal' to 'strong' model performance) when it was applied in wildfire events that occurred in the same ecoregion. However, the predictive power of the model is low for the independent fire event (RED084). The overall findings of this dissertation reveal that the 1) predictive model based on RF is robust enough to determine the relative importance of the predictors and their marginal effect; 2) the model was flexible enough to identify areas where wildfire residuals are likely to occur; and 3) there is a repeatable, robust measurement framework for characterizing residual patches and understanding their variability across different wildfire events.

DEDICATION

I dedicate this dissertation to:

The almighty God who gave me the strength and patience to complete this study and for walking with me through many trials in my life.

My parents Hayelom Araya and Hiwot Tesfamariam, and my brothers and sisters for their endless love, support, sacrifices, and encouragement.

ACKNOWLEDGMENTS

I have been very privileged to find inspiration in the academic and mentorship culture that my supervisor, Dr. Tarmo K. Remmel, provides to his graduate students. I would like, therefore, to express my warmest thanks and gratitude to him for his professional and personal support and for his energy and commitment to university values that enlightened my years as a graduate student.

My heartfelt appreciations are also extended to the work of my committee comprised of Dr. Ajith Perera (Ontario Forest Research Institute - OFRI), Dr. Taly Drezner (York University, Geography), and Dr. Baoxin Hu (York University, Earth and Space Science, and Engineering), for their guidance, support, and constructive feedback. I was privileged to work closely with these great individuals. My thanks also go to my external reviewers Dr. Dawn Bazely (York University, Biology), Dr. Marie-Josée Fortin (University of Toronto, Ecology & Evolutionary Biology), and Dr. Rick Bello (York University, Geography).

Several individuals have provided significant support towards the completion of this dissertation. I am indebted to Dr. Ajith Perera's research group: Marc Ouellette and Lisa Buse for their support during my stay at Sault Ste. Marie and during the fieldwork in the Red Lake. Particular mention is also due to Yvonne Yim (York University, Geography Graduate Program Secretary) for all her personal support and understanding during my stay at York University. I have been fortunate to have outstanding colleagues and friends upon which to rely for input and help with collection field data, and companionship. I especially would like to thank Mussie Tekeste for his input, support, and companionship through the years. I would also like to extend my thanks to Remmel's lab group: Doris Siu, Yuestas David, Connie Ko, Alex Singh, and Patrick Mojdehi.

I appreciate the financial support I received from NSERC, through my supervisor Dr. Tarmo K. Remmel. I was fortunate to have outstanding support for collecting field data and travelling to different parts of the country for presenting my work in various conferences (CAG and SKI).

Finally, my deepest appreciation goes to my wife, Kokob Afewerki, for her love, sacrifices, and understanding that allowed me to achieve my dreams. Most importantly, I would like to thank my parents, Hayelom Araya and Hiwot Tesfamariam, by brothers (Dawit, Isaias, and Kessete) and sisters (Yirgalem, Jerusalem, and Ruth), for everything they meant for me throughout my life.

I thank you from all my heart!

Table of contents

List of Tables

List of Figures

Acronyms and abbreviations

AB:	Alberta
AML:	Arc Macro Language
AS:	Alder Shrub
AOU:	Area of Undertaking
AUC:	Area Under Curve
BC:	British Columbia
BV:	Bedrock and non-Vegetated
BU:	Burn Area
CA:	Class Area
CART:	Classification and Regression trees
CB:	Complete Burn
CBI:	Canadian Boreal Initiative
CS:	Cloud and Shadow
DC:	Dense Conifer
DE:	Deciduous
DEM:	Digital Elevation Models
EL:	Elevation
END:	Emulation of Natural Disturbance regimes
ESA:	Ecological Society of America
FAO:	Food and Agricultural Organization
FRAC:	Fractal Dimension
GAM:	Generalized Additive Model
GIS:	Geographic Information Systems
GLM:	Generalized Additive Model
GOC:	Government of Canada
IFFN:	International Forest Fire News
LC:	Land Cover
LPI:	Largest Patch Index
LPM:	Landscape Pattern Metrics
LS:	Low Shrub
LSI:	Landscape Shape Index
MA:	Marsh
MAUP:	Modifiable Areal Unit Problem
MB:	Manitoba
MNN:	Mean Nearest Neighbour
MNR:	Ministry of Natural Resources
MPS:	Mean Patch Size
MSI:	Mean Shape Index
NDPE:	Forest Management Guide for Natural Disturbance Pattern Emulation
NP:	Number of Patches
NRC:	Natural Resource Canada
NT:	Northwest Territories
OB:	Old Burn
ON:	Ontario
OMNR:	Ontario Ministry of Natural Resources
OP:	Open Wetland
PA:	Producer’s Accuracy
PB:	Partial Burn

PCC:	Percent Correctly Classified
PD:	Patch Density
PDP:	Partial Dependency Plot
PSCV:	Patch Size Coefficient of Variation
PSSD:	Patch Size Standard Deviation
QC:	Quebec
RI:	Ruggedness Index
RF:	Random Forest
ROC:	Receiver Operating Characteristics
SC:	Sparse Conifer
SL:	Slope
TW:	Treed Wetland
UNEP:	United Nations Environment Program
%LAND:	Percent of Landscape
UA:	User's Accuracy
WA:	Water
WL:	Wetland variable

1. Introduction: background and context

1.1. Context

A forest may be defined as a biological community dominated by trees and other woody vegetation, but the way a forest is defined depends a lot on who is defining it. Forests has been defined differently by foresters, forest managers, and ecologists based on various physiographic criteria (e.g., area, crown cover, tree height, and tree density and proportion) (Malmberg and Miljoanalys 2001; Lund 2011). Based on canopy cover, for example, Pretorious (2013) defined forests as areas consisting of trees with 76-100% crown cover while Fisher et al. (2013) described forests as areas covered with dense tree growth (70-100%), > 20 m height. Forests are also described as land with trees reaching a minimum height of 5 m and crown cover of more than 10%, with an area of more than 0.5 ha, excluding land predominantly used for agriculture (FAO 2011). A recent study of the various definitions of forests found that more than 800 different definitions for forests and wooded areas were in use in the world, with some countries adopting several such definitions at the same time (Lund 2011). In Canada, for example, a minimum area of1 ha, 25% of canopy cover, and minimum tree height of 5 m are used to describe forests (GOC 2007). All of the definitions invariably describe an extensive plant community with a high proportion of tree cover. These extensive plant communities cover a large portion of the Earth's surface (4 billion ha, accounting for 9.4% of the planet and 31% of the total land area) (UNEP 2009; FAO 2011). The area of forest cover is unevenly distributed in the world, with the five most forest rich countries (Russia, Brazil, Canada, the United States of America, and China) accounting for more than half of the total forest area (53%) (FAO 2011). The geographical distribution of the world's forest covers is shown in Figure 1.1, with Europe (including the Russian Federation) accounting for 25% of the world's total forest area, followed by South America (21%), and North and Central America (17%).

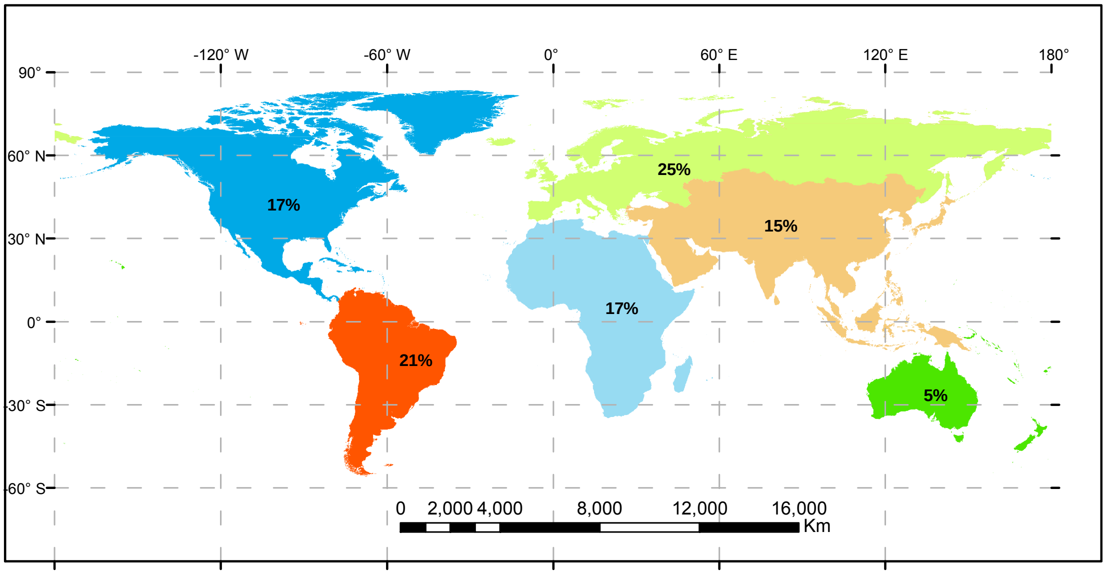


Figure 1.1. Geographical distribution of forest cover by region (Source: FAO 2011).

A typical forest is composed of different vegetation layers such as the overstory (upper tree layer of the canopy) and the understory (shrub, herb, and moss layer). Forested areas are also classified differently depending on the biome in which they exist (e.g., whether they are evergreen or deciduous) or based on species composition (i.e., whether the forests are composed predominantly of broadleaved trees, needle leaved coniferous trees or mixed). Some of the forest types include tropical and subtropical forests, temperate broadleaf and mixed forests, temperate coniferous forests, Mediterranean forests, Mangroves, boreal forests (Taiga). However, the focus of this study is placed on boreal forests, which are predominantly evergreen and coniferous, and mainly occupy the subarctic zone. Deciduous species such as aspen (*Populus tremuloides)* and birch (*Betula papyrifera)* are also common in the boreal forests, but becoming less frequent further north.

The boreal forest is largest intact forest ecosystem covering over 11% of the Earth's terrestrial surface (Bonan and Shugart 1989; Engelmark 1999), with temperature being the most important environmental factor determining its geographic location (Kuusela 1992). As shown in Figure 1.2, the boreal forests, which account for 29% of the world's total forest area, forms a "green-belt" of various width stretching through Russia, Canada, Alaska, and the Nordic countries (Finland, Norway, and Sweden) roughly between latitudes 45° and 70° N (Kuusela 1992; Olsson 2009). The region, specifically the boreal forests in Canada and Russia, is home to more than half of the world's remaining intact forest ecosystems (Olsson 2009). The region also plays an important role in maintaining biological diversity, controlling soil erosion, and promoting soil formation (Wells et al. 2010). This extensive forest area has an important influence on global, continental and regional climate over a short and long timescales (Weber and Stocks 1998; Olsson 2009). The boreal forests, comprising trees, wetlands, and peatlands, store a considerable amount of global carbon (i.e., over one trillion tons of carbon) and their biomass is so huge and vital that when they are in their maximum growth phase during the northern spring and summer, the worldwide levels of carbon dioxide fall and the levels of oxygen rise (Dale et al. 2001; Runesson 2011).

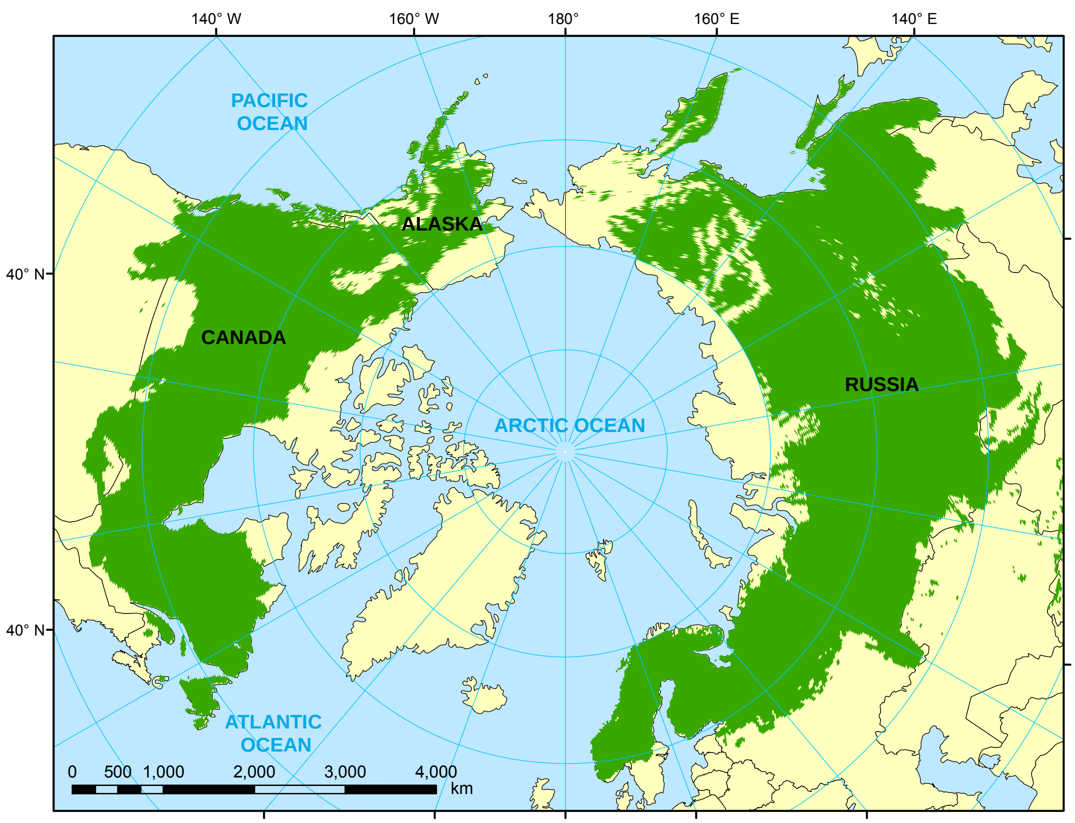


Figure 1.2. The boreal forest in the northern hemisphere occurs in broad band across northern America, Russia, and Nordic countries (Source: NRC 2014).

The North American boreal forest constitutes the largest biome in most Canadian provinces (and territories), and in the state of Alaska, USA. In Canada, the boreal forest forms a transcontinental band (stretching over 5,000 km) from Newfoundland in the east to the Yukon Territory in the west, comprising approximately 290 million ha (i.e., 30% of the circumpolar boreal zone) (Figure 1.3) (Rowe and Scotter 1973). The boreal region covers about 60% of the country's land area, and three-quarters of Canada's forest and other woodlands (CBI 2005). The region also forms one of the world's largest intact forest ecosystem; even larger than the remaining Brazilian Amazon (CBI 2005). The boreal forest exhibits similarities from the Atlantic to the Pacific coast, mainly in the composition of tree species (Parisien et al. 2011). A small number of needle leaved coniferous tree species such as black spruce (*Picea mariana* [Mill.] B.S.P), white spruce (*Picea glauca* (Moench) Voss), jack pine (*Pinus banksiana* Lamb), tamarack/larch (*Larix laricina* [Du Roi] K. Koch.), balsam fir (*Abies balsamea* [L.] Mill.), a limited number of broadleaved trees such as trembling aspen (*Populus tremuloides* Michx.) and birch (*Betula papyrifera* Marshall) and shrub species including willow (*Salix* spp.) and alder (*Alnus* spp.) dominate the region. The region is also characterized the abundance of freshwater (around 1.5 million lakes), a high concentration of wetlands or peatlands, and some of the world's richest deposits of natural resources (Wells et al. 2010).

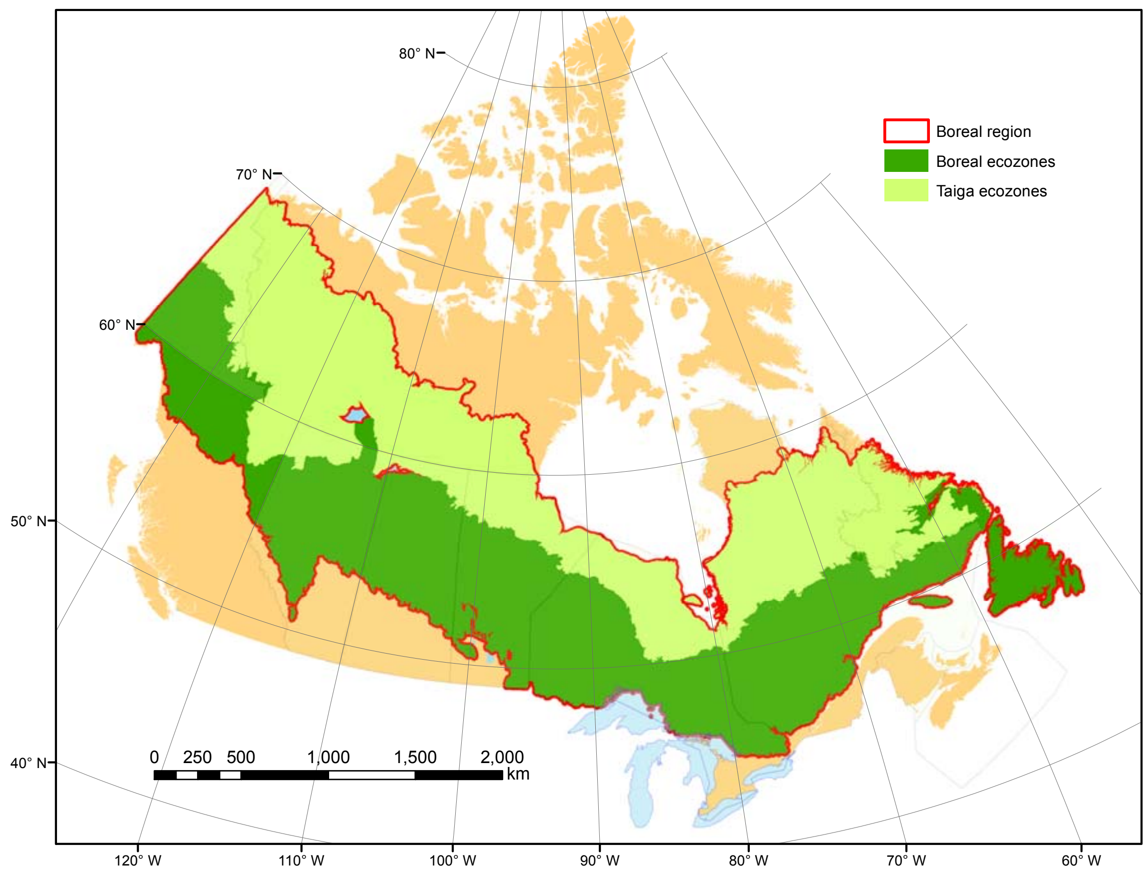


Figure 1.3. The Canadian boreal forest cover map: the boreal forest forms a broad band stretching from Newfoundland in the east to the Yukon Territory in the west.

In Ontario, the boreal forest region occupies around 50 million ha, which is two-thirds of the forest cover in Ontario or 46% of the province's land area. The boreal forest in Ontario, which comprises 15% of Canada's boreal forest, is more than 1.5 times larger than the size of France. This vast forested space has the potential to contribute to climate change through its influence on the global carbon cycle (global warming) (Weber and Stocks 1998). The boreal forest in Ontario alone stores 49 billion tC in its soils, peat, and forests; an amount equivalent to 249 years of Canada's annual carbon emission (CBI 2005). The area has also a wide range of socioeconomic importance (e.g., recreation, landscape and community protection, timber production, and employment). The region however has been shaped and changed by various forms of disturbance, mainly wildfire. Generally, the northern boreal forest has a shorter fire season than the south, but the greater summer daylight period and the dominance of conifers species, which are more flammable than deciduous species, make the region vulnerable to various wildfire incidences during fire seasons (Parisien et al. 2011). The boreal forest also has a broad-scale longitudinal moisture gradient, where by areas of central Ontario experience more frequent and more intense droughts than in eastern Canada, which results in greater fire weather severity (Parisien et al. 2011). Studies have also indicated that northwestern Ontario has experienced more fire than the rest of the province; thus my study focuses on fire incidents that have been recorded in northwestern Ontario in different wildfire seasons.

1.2. Boreal forest fire disturbances

1.2.1. Boreal forest fire behaviour

Boreal forests are inherently dynamic; this dynamism is attributed to different agents of disturbance including wildfire, extreme weather (wind and ice-storms), insect infestation, and harvesting (Bergeron et al. 1998). For example, extreme weather events affect forest conditions by blowing down large swaths of trees or causing snow and ice damage while forest insects (e.g., spruce budworm or forest tent caterpillar) defoliate vast areas of forest (OMNR 2009). Although all forms of natural disturbances play a role to shape the boreal landscape, wildfire is the primary agent of disturbance that affects the boreal landscape and its biodiversity over the long term (Johnson 1995; van Wagtendonk 2004; McKenzie et al. 2011).

The ignition and occurrence of wildfire can generally be triggered by different natural factors (e.g., lighting and sparks from rock falls) (Ainsworth and Doss 1995), but in the boreal region lighting is the primary source of natural ignition. Many wildfires are also attributed to human sources such as land conversion burning or agricultural activities, recreation (i.e., careless campfires), and industrial activities (forest industry). However, the seasonality, frequency, size, and behaviour of human-induced fires are different from naturally caused fires (Johnson 1995). Compared with human-induced fires, naturally ignited wildfires are more extensive, frequent, and seasonal in the boreal region (Olsson 2009). Nonetheless, for many forms of disturbance, there is a gradient from relatively minor (e.g., damage on individual trees) to relatively major events (damage to thousands hectares forest cover) (White 1979). Lighting in boreal forests, for example, can cause a fire that can damage from scales ranging from a wildlife tree to landscapes (OMNR 2010).

For fire to play a role in landscape change, a source of ignition, sufficient fuel to burn, and favourable weather conditions for burning must be present (van Wantendonk 2004). These conditions are all met frequently in the boreal forest, and fires occur annually throughout the region. Boreal forest fires specifically occur when 1) a colder and stable arctic airstream is replaced by a warmer and unstable air streams, and 2) the mean air temperature is above 0ºC (Johnson 1995; Flannigan and Wotton 2001). These conditions trigger temperature increases and fire ignition; the fire spreads and grows in size and intensity when the conditions are favourable (Viegas 1993).

Boreal forest fires can be classified based on their physical fire behaviour into three general categories: ground fires, surface fires, and crown fires (Nelson 2001). A ground fire burns or smoulders materials on the ground surface including duff, tree or shrub roots (Butler 2007; Max et al. 2010). Surface fires burn the upper litter layer and small branches that lie on or near the ground. Surface fires produce flaming fronts that consume needles, moss, lichen, herbaceous vegetation, shrubs, small trees, and saplings (Max et al. 2010). The fire sometimes propagates as a crown fire when it grows vertically and reaches the trees' crowns (Viegas 1993; Flannigan and Wotton 2001; Johnson 1995). Crown fires can be either intermittent (trees torching individually) or active (with solid wall of flame development in the crowns) but active crowns are the most common type of fire in the boreal forest (Flannigan and Wotton 2001). Boreal forest fires are specifically characterized by high fire intensity (crown fires with intensities from 8,000 to > 100,000 KW/m; i.e., when flames extend into and ignite the tree crowns), high flame length (> 5 m) (Johnson 1995), and frequent cyclic fire behaviour or return interval (< 100 years) (Cui et al. 2009).

Fuel consumption and spread rates can also vary both within and between boreal fires, but generally crown fires consume 20-30 tonnes/ha of fuel with roughly two-thirds of this total associated with consumption of forest floor (litter, moss, and humus layer) and dead woody surface fuels while crown fuels (needles and fine twigs) account for the remaining one-third of the total fuel consumed (IFFN 2004). The spread rates in the boreal forests can also vary from ~5 m/min in intermittent crown fires to > 100 m/min in fully developed crown fires (IFFN 2004). Such wildfires have been responsible for the burning of millions of ha of forests annually throughout the region. Forest fire statistics from northern circumpolar countries, for example, indicate that an estimate of 5-15 million ha burns annually in the boreal region (IFFN 2004).

1.2.2. Wildfires in the boreal forests of Ontario

The boreal forest in Canada is a mosaic of species and stands, varying in composition from coniferous to deciduous. The diversity of the forest mosaic is largely attributed to the recurring wildfires over a long period of time. The wildfires have been recorded in every part of the Canadian boreal forest, but the number of occurrences and area burned vary temporally and spatially (Parisien et al. 2011). The variation can be attributed to the localized weather patterns (e.g., those produced by large water bodies) or to the variation in the source of ignition, fire intensity, and vegetation patterns (OMNR 2009; Parisien et al. 2011). The annual fire occurrence in Canada has increased from approximately 6,000 in the 1930-1960s to around 9,000 fires during the 1980s and 1990s (IFFN 2004). The increase in fire occurrence can be explained by a growing population, increased forest use, or due to development in fire detection capability (IFFN 2004). Besides, the NRC report indicated that an average of 8,300 forest fires have occurred over the last 25 years; with the total area burned averaging 2.3 million ha in Canada (NRC 2014). However, only 3% of the wildfires burn an area larger than 200 ha; these fires account for 97% of the total area burned across the country (NRC 2014).

Moreover, the number of wildfires and area burned vary spatially across the country, with British Columbia (BC) experiencing the highest number of wildfires record in the last decade followed by Alberta (AB) and then Ontario (ON) (Figure 1.4). The number of fires recorded in Quebec was not as high as the fires recorded in BC, AB, and ON, but Quebec (QC) experienced the largest area burned followed by Manitoba (MB) and Northwest Territories (NT) (Figure 1.5).

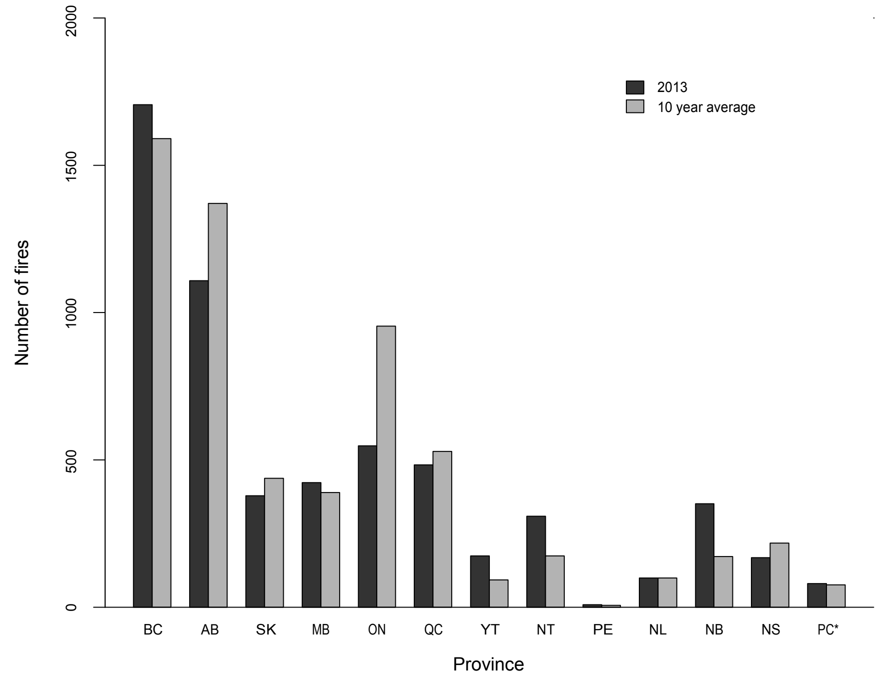


Figure 1.4. Total number of fire incidents in 2013 and 10 year average, by province; PC* = Parks Canada (Source: NRC 2014)

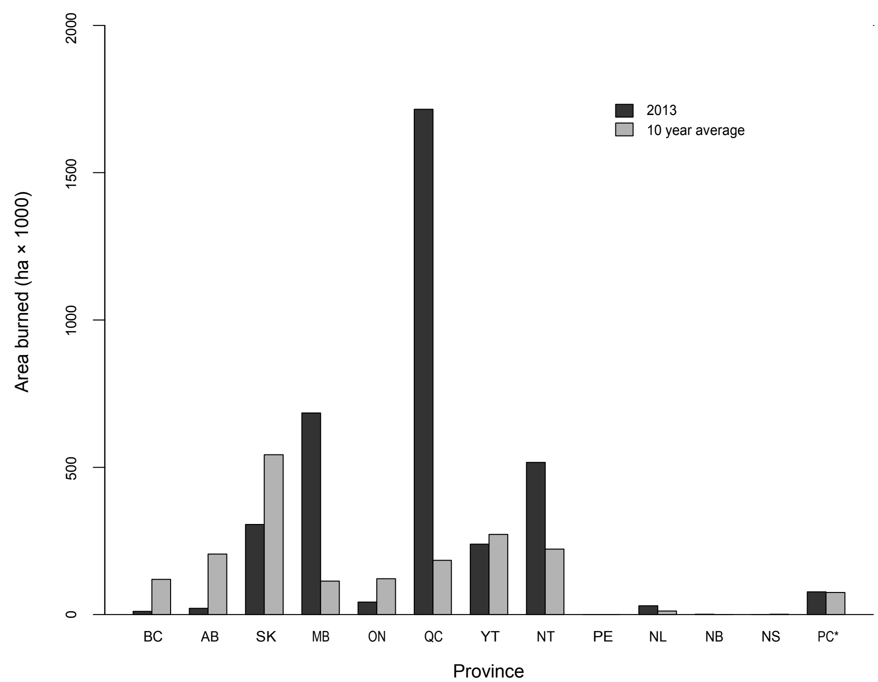


Figure 1.5. Total area burned in 2013 and 10 year average, by province (Source: NRC 2014).

The boreal region is characterized by short growing seasons, low temperatures, long summer daylight hours, low biological productivity (Engelmark 1999), and relatively low and variable annual rainfall (Bonan and Shugart 1989) between 600 to 900 mm (Runesson 2011). Owing to these favourable conditions for burning, wildfires have occurred throughout the boreal forest landscape of Ontario, but the number of fires and area burned varies spatially within the boreal forests in Ontario, where highest wildfire activity has concentrated in northwestern Ontario (Figure 1.6).

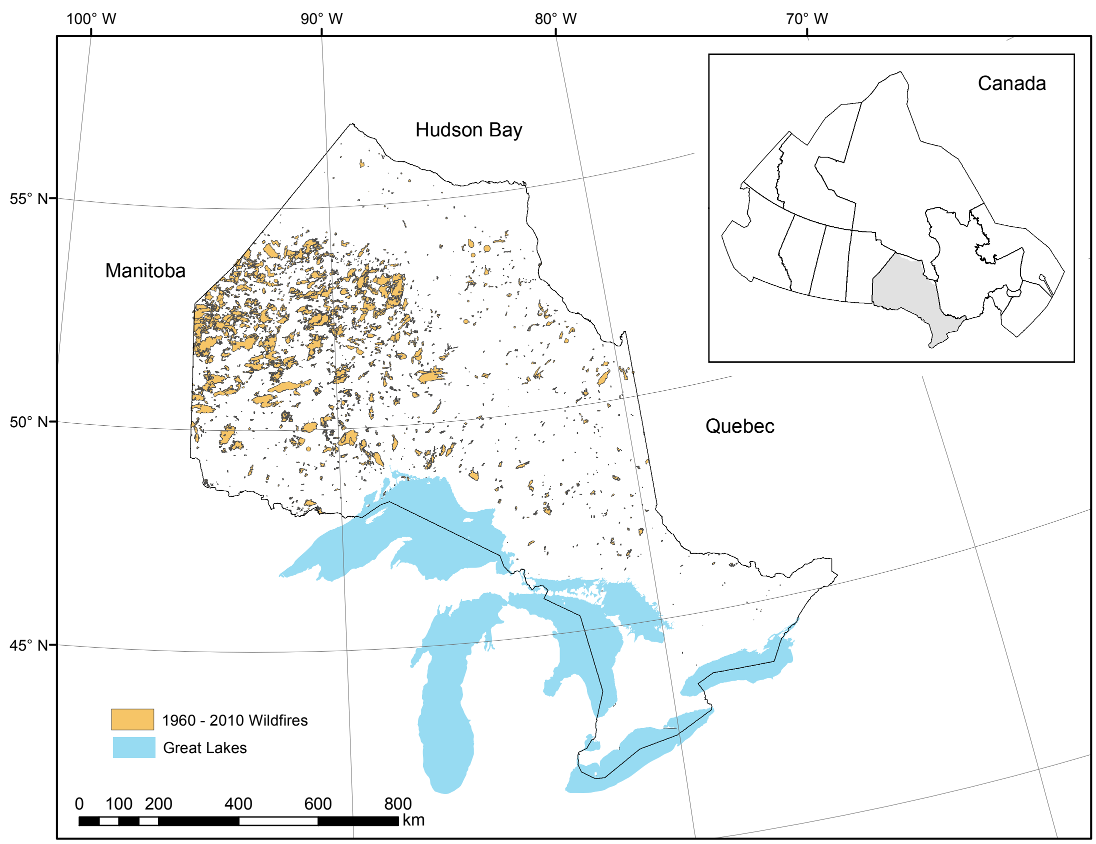


Figure 1.6. The occurrence and extent of wildfires in Ontario during the last 50 years: northwestern Ontario had the highest concentration of large wildfires.

The occurrence of wildfires in Ontario has also been inconsistent during the past 50 years. The 2008 fire season is recognized as the lowest year of fire record in the past 50 years followed by the 2009 fire season (OMNR 2011). The total number of fires during the fire seasons of 2008 and 2009 was 341 and 384, burning 1,316 and 20,656 ha respectively. The fire season in 2011 recorded 1,334 of fires in Ontario affecting more than 635,374 ha of land; the most area burned in the past 50 years. As shown in Figure 1.7, the number of fires and the area burned in Ontario during the last decade varies from year to year, but wildfires in Ontario burn approximately 1% of the boreal forest each year (Remmel and Perera 2009).

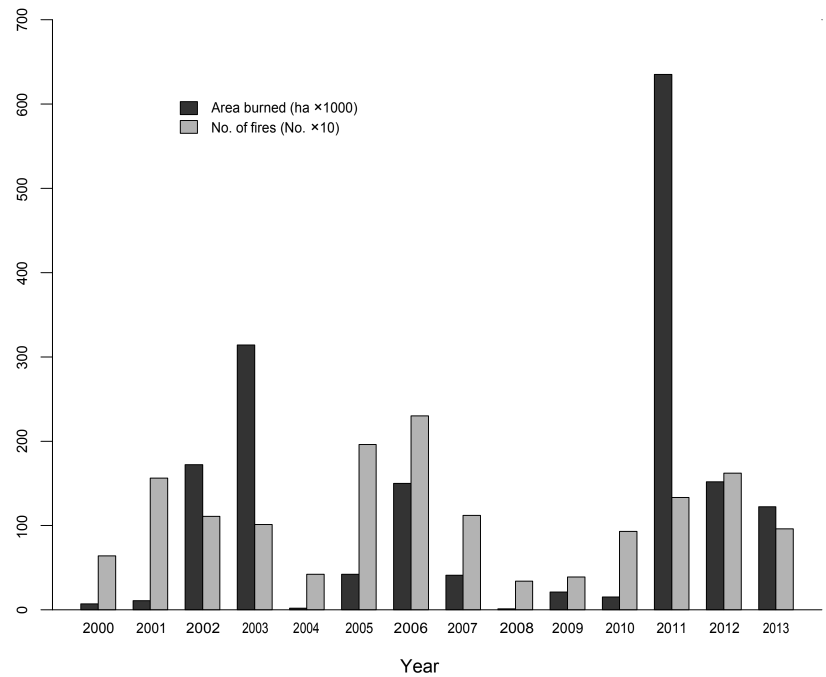


Figure 1.7. The total number fire incidents and total area burned in Ontario between 2000 and 2013 (Source: NRC 2014).

1.2.3. Factors affecting wildfire behaviour

The spatial and temporal dynamics of boreal forest fire is attributed to various factors such as weather and climate, topography, natural firebreak features, and forest fuel that interactively affect fire behaviour (Viegas 1993; Butler 2007; Cui et al. 2009). The primary weather variables that affect fire behaviour are temperature, wind, and relative humidity (Viegas 1993; Butler 2007). Of these factors, by far the most important is wind, defined by its velocity and direction (Rowe and Scotter 1973) and influenced by topography (van Wagtendonk 2004). Wind brings an additional supply of air to the fire, and causes spot fires by blowing sparks and embers ahead of the main fire into new sources of fuel. Generally, the stronger the wind, the faster the rate of fire spread. Wind blows through the forest, dries the foliage and makes it increasingly flammable; the fire is further intensified by the abundance of light, small, and fast-burning forest fuels or duff (e.g., dry grass, dead leaves and tree needles). Heavy fuels (e.g., logs, stumps, and branch wood) take longer to ignite and slow fire spreads slower relative to lighter fuels. Forest fuels have a varying degree of effect on fire behaviour as they have different properties depending on plant species type (whether they are alive or dead), fuel types (light or heavy fuels), and the amount of fuel available and its spatial distribution (Viegas 1993; ESA 2002; van Wagtendonk 2004). The moisture content of fuel particles is also of great importance for fire spread, as high values of moisture content slow the rate of burning even prevent fire spread (Nelson 2001). Owing to the greater abundance of fine fuels in the form of needles and twigs, canopy architecture, low foliar moisture and thin bark, the boreal forests are characterized by high fire intensity.

Fire behaviour is also affected by topography (e.g., slope, aspect and elevation). Terrain may control wind flow in a relatively large area as wind follows the direction of least resistance features (e.g., flat or nearly flat surface). A fire ignited at the bottom of a slope spreads rapidly, and gains momentum, as it burns uphill because warm air rises and preheats uphill fuels (Viegas 1993; ESA 2002). A fire ignited on the top of a slope, on the other hand, spreads slowly as it burns downhill. Furthermore, topographic features such as streams and lakes can create natural firebreaks, and hence influence fire spread and intensity, and distribution of burns.

1.2.4. Effects of wildfire on vegetation and wildlife

Fire is a disturbance that influences plant communities over time and serves as an important function in maintaining biological diversity in forest ecosystems. Wildfire facilitates the removal of old trees and clears dead (or decaying organic matter) within the forest; this enables new plants to flourish (Major 2005; Clark and Bobbe 2007; Marzano et al. 2012). An added effect of plant removal is an increase in sunlight, which can also allow seed germination. Additionally, wildfires play an important role in driving forest ecosystem dynamics by removing diseased trees along with the insects that are associated with those trees (Runesson 2011). The process eventually affects the physical and biological processes of forest ecosystems such as forest succession and biological diversity (Ahlgren and Ahlgren 1960).

In the boreal forest, naturally occurring fire often causes loss of vegetation or biomass and animal species richness (Dale et al. 2001; Hooper et al. 2004; Gorte 2006). However, there are some ecosystems that rely on naturally occurring fires to regulate growth because their organic matter following fire is converted to available nutrients that support new growth (Ahlgren and Ahlgren 1960; Rowe and Scotter 1973). For example, aspen (*Populus tremuloides* Michx.) and jack pine (*Pinus banksiana* Lamb) require fire to regenerate in the boreal landscape (Runesson 2011). A fire creates a favourable condition for jack pine to germinate in the burned area because the species have cones that require heat of fire to release their seeds and re-establish themselves following the fire. Owning to its semi-serotinous cones, black spruce (*Picea mariana* [Mill.] B.S.P) may also become established following a fire, but this species grows slower than other species (e.g., jack pine); it may become established in the years following a fire. Aspen (*Populus tremuloides* Michx.) and birch (*Betula papyrifera* Marshall) are also able to re-establish quickly by sprouting from stumps and roots of burned trees or my producing abundant seeds that can be blown by wind over long distances. However, species such as balsam fir (*Abies balsamea* [L.] Mill.), white spruce (*Picea glauca* (Moench) Voss), and white cedar (*Thuja occidentalis*) are disadvantages during extensive wildfires. These species re-establish in the burned areas only when their seeds are blown into the burned area either by wind or brought by animals. Consequently, the species take longer to establish themselves in the burned areas; in some cases it takes more than 150 years for these species to reappear in the burned landscape (OMNR 2009). The differences in species adaptation to fire and stages of forest succession increase the biological diversity in ecosystems.

Wildfires affect wildlife population either directly by the heat and smoke of fires or subsequently weakened from habitat loss. The habitat loss caused by fire affects wildlife much more profoundly than the fire itself because food sources are scarce during fire seasons, and leads to losses within wildlife populations (Huff and Smith 2000). Although most wildlife population are directly affected by a fire, the degree of impact depends on various factors, including fire uniformity, severity, size, duration, and burn season as well as wildlife mobility. For example, animals with limited mobility (e.g., insects, older and weaker individuals) are more vulnerable than some large mammals (e.g., deer and moose) (Druhjell 2004). The rate of wildlife mortality also depends on the burn season. For example, if a fire occurs when animals (e.g., birds) are nesting or having young animals with limited mobility; the mortality rate is higher. Another factor that can lead to a loss (or reduction) of wildlife populations is the loss of habitat and the associated food sources following a fire. However, the change in species composition may provide alternate or even superior food sources for some animals (ESA 2002). For example, high severity fires that result in wildlife mortality can benefit other fauna, such as bears, coyotes, eagles, and common ravens (Druhjell 2004). These animals may find an increased availability of food sources as the reduced forest cover makes prey (and dead animals) more visible. Additionally, fire removes the lichen from the ground and it can severely affect some animals (e.g., caribou) but favours moose that feed on the advance growth (new saplings) that emerge after a fire (Runesson 2011).

Besides, wildfire can affect the physical and chemical properties of soil and hydrological processes, including the loss or reduction of structure, soil organic matter, and reduced permeability (Rowe and Scotter 1973). The changes in soil properties can result in increased hydrophobicity (water repellency) which results in decreased infiltration and increased ruff-off which eventually results in increased soil erosion (Bonan and Shugart 1989; Whelan 1995).

1.2.5. Forest management in the boreal forests

The physical dimension of the boreal forest resources may lead to extensive forestry activities and exaggerates the estimates of the feasible potential timber production, but not all forest resources in the boreal region are subject to industrial forestry harvesting (Kuusela 2011). Forest resources in remote areas or in extremely harsh climate are beyond the economic limit of harvesting operations. In Canada, 12% of the boreal forest area is not exploitable (protected by legislation), while less than 1% of Canada's forests are harvested annually. For example, in 2009 0.6 million ha, which is slightly larger than the size of Prince Edward Island (i.e., 0.56 million ha) were harvested across the country. In Ontario, forestry activities occur on 49% of the province's boreal forests and this takes place on what is referred to as the Area of Undertaking (AOU) (Figure 1.8), which refers to areas where forest management activities are permitted. For decades, the boreal forests of Ontario have been regenerated after clear-cutting because by law all forest harvested on Canada's public land must be successfully regenerated (NRC 2014).

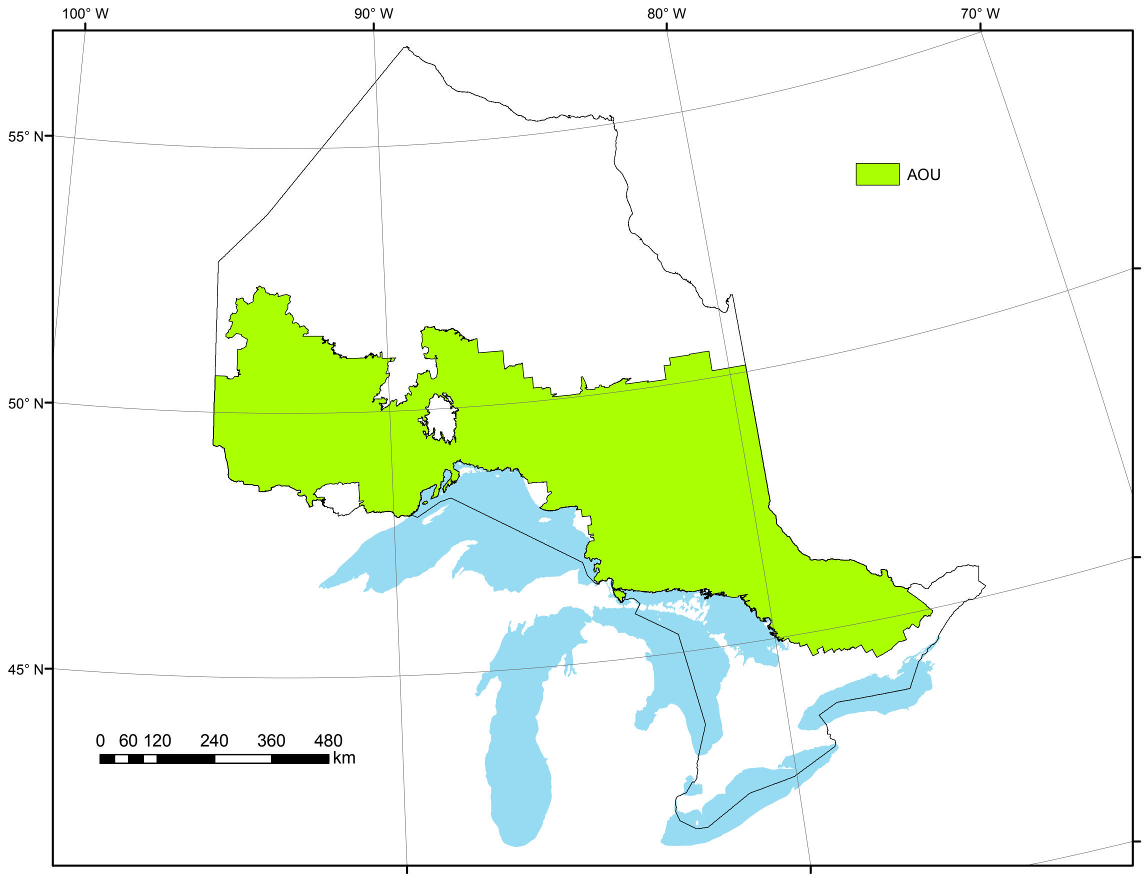


Figure 1.8. Area of the Undertaking (AOU) – the southern portion of Ontario's boreal forest where commercial forestry is currently permitted.

In order to implement sustainable forest management practices, the managed forest should resemble the forests formed by disturbances (Perera and Buse 2004; North and Keeton 2008). The more managed forests resemble the forests formed by fire disturbances, the greater the probability that the forest diversity is maintained (Smith and Hendry 1998). This requires that both the resources and biological diversity of the forests as well as the economic benefits (i.e., forest industry) of the forest be maintained (Delong and Tanner 1996; North and Keeton 2008). In order to sustain the resources and biological diversity, there has been an increasing interest in forest management approaches based on the Emulation of Natural Disturbance regimes (END) (Bergeron et al. 2007; Kramkowski 2012). END represents to

> "The management strategies and practices, at appropriate spatial and temporal scales, with the goal of producing forest ecosystems as structurally and functionally similar as possible to the ecosystems that would result from natural disturbances, and that incorporate the spatial, temporal, and random variability intrinsic to natural systems" (Perera and Buse 2004).

END is based on the premises that forests, particularly the boreal forests, are shaped by disturbances, specifically wildfire (Kramkowski 2012). It is suggested to maintain forest compositions and structures similar to those existing under natural disturbance regimes; this allows forest managers to reduce the negative impacts of harvesting on biodiversity (Klenk et al. 2008; North and Keeton 2008). In Canada, END as a harvesting and forest management approach is integrated in policies and practices for sustainable ecosystem management (Klenk et al. 2008). Specifically in Ontario, criteria for emulating natural disturbance patterns during harvesting were provided in the Forest Management Guide for Conserving Biodiversity at the Stand and Site Scales; the stand and site guide uses a combination of coarse and fine filter approaches to biodiversity conservation (OMNR 2010). Coarse filters create a diversity of ecosystem conditions, based on emulating natural patterns and processes, to provide habitat for the majority of native species of plants and animals while fine filters are applied when the requirements of particular species may not be adequately addressed by coarse filters alone (OMNR 2010). The objective of the forest management guide is "to direct forest management activities to maintain or enhance natural landscape structure, composition and patterns that provide for the long term health of forest ecosystems in an efficient and effective manner" (OMNR 2014).

Specifically, the coarse filter approach, which is based on emulating natural patterns, is aimed at providing directions for forest practitioners in the implementation of forest management practices that closely resembles the natural landscape created by fire in relation to the location, size of disturbance, residual structure, and species composition (OMNR 2001). The management guide in Ontario provides directions and guidelines related to landscape harvest patterns (i.e., percent of planned clearcuts, and spatial and temporal separation for planned clearcuts) and structural legacies (i.e., type and amount of structural elements that have to be retained during harvesting). Similarly, Alberta provides criteria for forestry companies for a sustainable forest management approach (i.e., for multiple environmental, economic, and social values of boreal forests) (Kramkowski 2012). However, Alberta does not have guidelines that specify and direct END to the same extent that Ontario does (Kramkowski 2012). In Quebec, the regulation respecting standards of forest management for forests in the domains of the state requires the presence of forest residuals within harvested areas, wooded edges, buffer strips between roads or water courses and harvested areas (Dragotescu and Kneeshaw 2012).

The management guidelines in Ontario require the retention of wildlife trees, standing individual trees or stems or small clumps of trees or stems (< 0.1 ha) within areas of operations for improving forest management strategies (OMNR 2010). The retention requires a thorough examination and understanding of the composition, spatial arrangements, and distribution of residual patches. For example, the structural elements that should be retained and the geographic locations that are most suitable for this retention need to be determined. In this instance, END works to imitate natural disturbances either at a stand level (e.g., retaining the number of snags or individual trees per unit area) or at the landscape level (e.g., size, shape, and distribution of unburned areas) (Delong and Tanner 1996; Kramkowski 2012).

1.3. Post-fire landscape structure

1.3.1. Landscape structure

Landscape ecology deals with the study of landscapes; specifically the composition and configuration of a landscape (McGarigal et al. 2002). The term landscape has been defined differently depending on the phenomenon under consideration, but the definitions invariably include an area of land containing a mosaic of patches (McGarigal et al. 2001). Forman and Godron (1986), for example, defined it as a spatially heterogeneous land area composed of clusters of interacting ecosystems that is repeated in similar form throughout. Turner et al. (2005) defined a landscape as an area that is spatially heterogeneous in at least one factor of interest. A landscape has also been considered as an area of land containing an interesting pattern that affects and is affected by an ecological process of interest (e.g., wildfire). A landscape is composed of three generalized elements – spatial components that make up a landscape: patches, corridors, and matrix; the extent and configuration of each of these elements define the patterns of the landscape (Forman and Godron 1986). The combination of pattern of patch-corridor-matrix has also been used to describe landscape structure and infer the underlying agents of pattern formation (Duning and Ziuzhen 1999).

Similarly, fire disturbances in a forested landscape can be explained using the patch-corridor-matrix model of a landscape where the forested landscape and undisturbed vegetation patches within a disturbance can respectively be described as the forest matrix and patch respectively. Forest fire disturbances often generate patterns of biological and ecosystem diversity at different geographical scales (Bergeron et al. 2007). Within a perimeter of a single fire event, for example, there is a large variation in the patterns and types of legacies left after the fire incidence (Bergeron et al. 2007). Since the patterns of post-fire landscape structure are useful to understanding of fire disturbances, forested landscapes, and support sustainable forest management practices, studies have been undertaken to examine their patterns, characteristics, and variabilities (Delong and Tanner 1996; Cuesta et al. 2009; Perera et al. 2009; Dragotescu and Kneeshaw 2012).

1.3.2. Spatial language

Since one of the necessities of pattern research is simplifying spatial concepts into meaningful spatial units (O'Neill 1988), the terms and concepts that pertain to post-fire landscape structure are defined. Terms such as wildfire, fire event, fire perimeter, and unburned remnants are intuitively spatial concepts, but their usage as universal language is less obvious (Andison 2012). A brief description of some of the spatial concepts that are used frequently throughout this dissertation along with a hypothetical (pictorial) representation is shown below; the descriptions are in relation the boreal forest landscapes. The spatial language as described in Figure 1.9 is a conceptual representation of a wildfire and its impact on forested landscapes.

Figure 1.9. Spatial language: summary of the spatial terms and features used in this study.

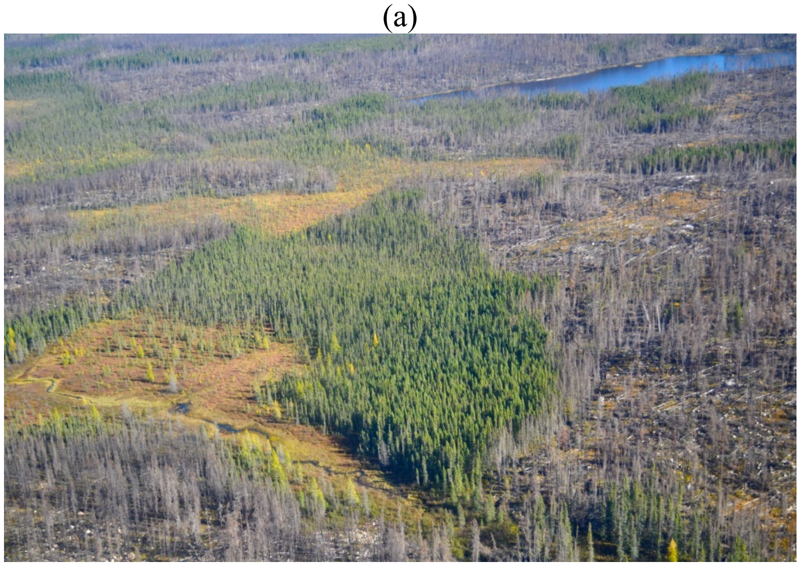
A fire event is a discrete forest disturbance event in time and space that disrupts the physical and biological structure of an ecosystem and the availability of resources. In this study, a fire event is used to describe a fire-disturbed area(s) that occurs during a specific period and is usually caused by a single ignition. A single fire event may include one or more disturbance patches that originated by spot fires (i.e., wind driven embers) and occurred beyond the existing fire boundaries.

Fire footprint – this refers to the spatial boundary that encompasses boreal wildfire processes; the area within the most probable locations of the outer fire boundary. It is defined based on: 1) a binary conceptualization of a classified map where 1 = all burned pixels and 0 = unburned pixels, and 2) a focal window analysis where a 3×3 focal window was passed over the binary layers 0 and 1; a focal sum was computed as a measure of a pixel's membership in the fire, resulting in focal sum values ranging from 0 (when al pixels are unburned) to 9 (when all pixels are burned). Any pixel with a focal sum value ≥ 1 indicates probability of membership of a fire and is coded as 1 to present a footprint. The footprint was then shrunk inward by 1 pixel to avoid outward bias by the 3×3 focal function

Figure 1.10. A complex mosaic of post-fire landscape structure: a mixture of burned (dark), partially burned (grey), and unburned (green) areas.

Post-fire residual patch – the wildfires do not burn the entire landscape; they rather create a complex mosaic of post-fire landscape structure, with different degree of burn severity (a mosaic of burned, partially burned, and unburned areas (Figure 1.10). The post-fire residuals are broadly defined as remnants of the pre-fire forest ecosystems that have retained their structure and were not entirely reduced to ash or charcoal during the fire. The composition of living and dead vegetation patches and their spatial arrangements differ depending on the fire behaviour and other pre-fire forest characteristics. Besides, following large wildfires, the most visible component of the fire footprint are the burned areas where all living structures are dead. Also, the resulting spatial mosaic encompasses: 1) partially burned areas where the fire passed through but did not kill the entire vegetation and 2) unburned areas that escape burning and retained within the fire footprint (Figure 1.11). The focus of this study was on unburned areas (hereafter described as post-fire residual patch) that entirely escape fire and left completely unaltered within the fire perimeter.

(a)



(b)

Figure 1.11. A mosaic of burned and unburned areas. (a) Completely burned areas within the fire footprint with high abundance of dead trees that are reduced to charcoal. (b) Presence of unburned areas (green) within a wildfire, including residual patches that escape fire. In this case,

the large island around the wetland escapes burning while the surrounding areas burned during the fire.

A post-fire residual patch is conceptually defined as a mix of live (and dead) vegetation that form a spatial continuum, ranging from undisturbed patches of live trees to a single tree stem (Nikora et al. 1999; Swystun et al. 2001; Perera et al. 2009a). Residual patches can be classified into different categories based on specific study goals or the scale of observations. Based on the spatial scale of their occurrence, for example, residual patch can be classified as 1) live tree patches (patch-level residuals), 2) standing live trees, and 3) snags (standing dead trees) (Perera et al. 2007; Routledge 2007).

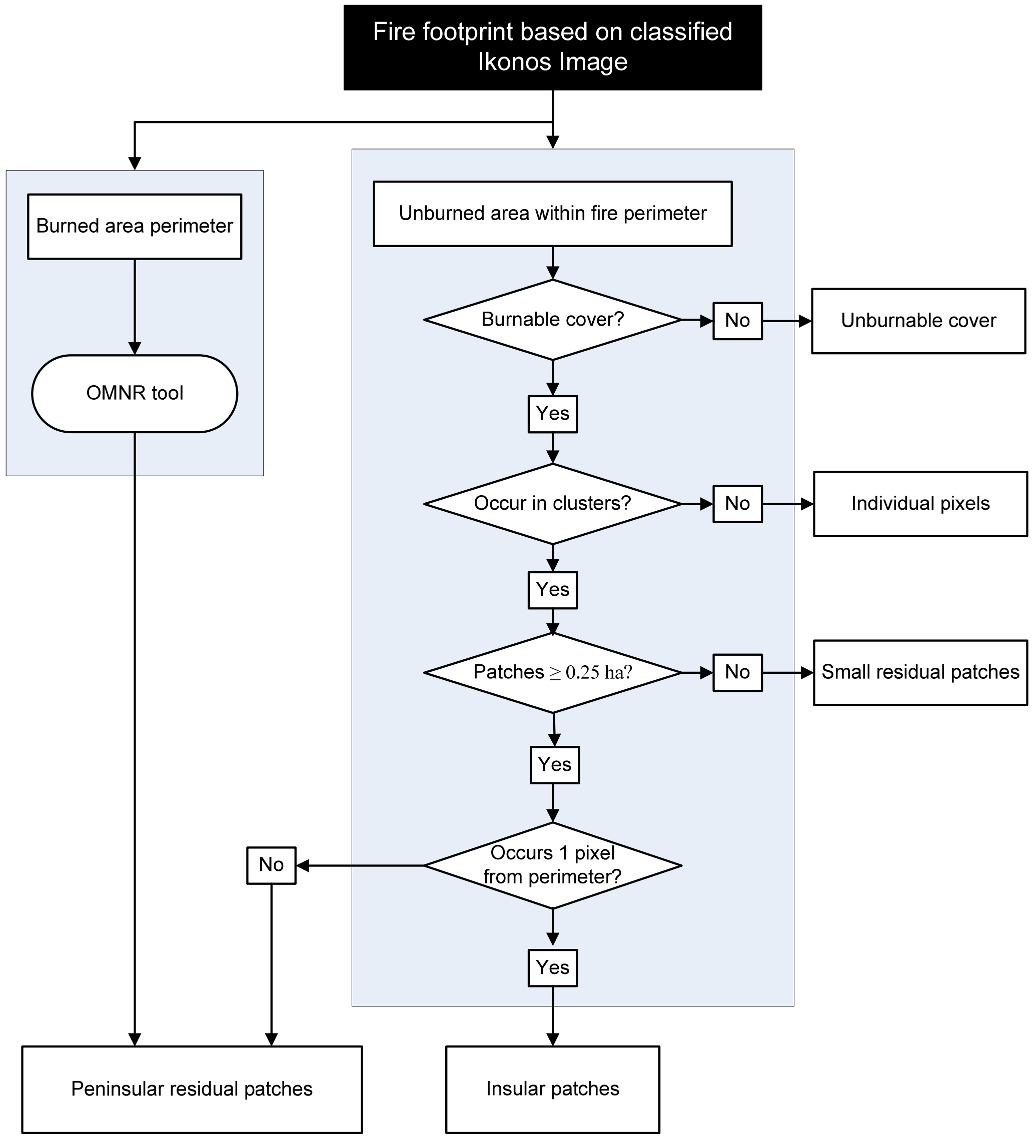


Figure 1.12. Stepwise criteria used to determine insular and peninsular residual patches for fire footprints. Both insular and peninsular patches are extracted from classified Ikonos images based on their size and location in relation to the fire footprint (Source: Perera et al. 2009a).

The conceptual definition of residual patch has been evolved to ensure consistent interpretation and usage of the term. In the Forest Management Guide for Natural Disturbance Pattern and Emulation (NDEP), for example, a residual patch has been used to describe both insular and peninsular patch types (Figure 1.9) (OMNR 2001). The concepts of insular and peninsular patches have also been described respectively as residual islands and residual matrix respectively (Andison 2004; Dragotescu and Kneeshaw 2012). As shown in Figure 1.9 (above), insular patch refers to undisturbed vegetation patches that are entirely contained within a fire event, at least 1 pixel dimension inward from the fire perimeter (at the grain of representation) while a peninsular patch is the undisturbed forested patch contained within a fire event but physically connected to the surrounding forested matrix. Specifically, an insular patch has been used to refer to as undisturbed treed or vegetated land cover class greater than or equal to 0.25 ha while a peninsular patch is defined an area that extends into the disturbance and has a base of less than 400 m (for fires ≤ 260 ha) or 1,000 m (for fires > 260 ha, and generally is longer than its base width (OMNR 2001). The method used to define and extract insular and peninsular patches based on their size and location in relation to the disturbance area perimeter is presented in (Figure 1.12).

The use of the term residual has been changed from the one described in NDEP (OMNR 2001) to reflect the patterns considerations, the variation in residual patch size, and suitability of site-specific habitats (OMNR 2010). The live and dead trees that remain standing after a disturbance in fire-origin forests haven generally referred to as residual trees or snags, or residual structure. In the recent management guide (OMNR 2010), for example, wildlife trees are used in place of 'residuals' or 'residual trees', which sometimes led to confusion, as mappable stands of trees are usually referred to as 'stand level residuals' (OMNR 2010). Wildlife trees are standing individuals trees or stems, or small clumps of trees or stems, within areas of operations; a clump of wildlife tree is < 0.1 ha in size (OMNR 2010). In this study, residual patches are used to describe any (live) undisturbed forest patch that are entirely contained within a fire event but physically not connected to the perimeter of the fire footprint. This is regardless of age and type of forest species that form the patches.

1.4. Scaling for understanding spatial pattern

1.4.1. The meaning of scale

Owing to the scale multiplicity in pattern and processes, scale holds the key to understanding the pattern-process relationships. Both scale and scaling are inevitably related to landscape ecology (Wu 1999; Wu and Qi 2000) and have been prominent in characterizing spatial patterns over multiple scales. The notion of scale refers to size in space and time; size is a matter of measurement (Allen and Hoekstra 1992). Scale is also often understood as expressing dimensions of time and space (Linke et al. 2007); consequently it has been used to describe both spatial and temporal scales. Spatial scale is usually considered as the product of grain and extent (Wiens 1989), which in remote sensing, relate to the spatial resolution (length of a pixel's edge in one dimension) and area coverage, respectively (Gustafson 1998; Allen and

Hoekstra 1992). The effect of scale on diverse spatial phenomena (e.g., patterns of post-fire landscape structure) can be studied using these two components of a scale (Wu and Qi 2000).

1.4.2. Scaling and scale effects

Apart from the concept of scale, attention has been given to the concepts of scaling and scale effects. Scaling focuses on what happens to the patterns and characteristics of an object when its scale (size or dimensionality) is changed; therefore, it is defined as is the process of information transformation or extrapolation over multiple scales (Marceau and Hay 1999; Wu 1999; Wu and Li 2006). However, scaling is a challenge in both theory and practice (He and Mladenoff 1999; Wu et al. 2000) because of the non-linearity relationship between processes and variables, and landscape heterogeneity that determines the process (Wiens 1989; De'ath and Fabricius 2000). The first step in designing a scale-dependent experiment is to identify the factors operational at a given scale of observation (i.e., the spatial scale of the focal question) (Marceau and Hay 1999), which depends on the processes, organism, or responses of interest (Wiens 1989). In order to understand the scaling theory, three levels of analysis can be formulated: 1) the focal level in question (L_0), 2) the level below that (L_{-1}), and 3) the level above that (L_{+1}) (Allen and Hoekstra 1992). Defining the focal level of a hierarchy is the most important factor in the theory because focal level determines the resolution of the observations (O'Neill et al. 1991). The scaling theory suggests that when one studies a phenomenon at a particular hierarchical level (L_0), the mechanistic understanding comes from L_{-1} whereas the significance or context of that phenomenon can be revealed at L_{+1} (O'Neill et al. 1991; Allen and Hoekstra 1992; Wu 1999). The three levels can be described as micro (L_{-1}), focal (L_0) and macro (L_{+1}) scales respectively.

The process of information transformation or assessing the scale-dependency experiment can be accomplished by changing grain, extent, or both (Wu 1999). While working with scaling, one must distinguish between two forms of scaling: up-scaling and down-scaling. Up-scaling refers to a process that transfers information from local scale to derive processes at macro scale (Wu et al. 2000). Up-scaling can be achieved using is a resampling techniques, which are designed to transform an image data set acquired at finer spatial resolution to a coarser spatial resolution representation of the same image. Conversely, down-scaling is a method of transforming information from macro scale to local scale; decomposing information at one scale into its constituents at smaller scales (Marceau and Hay 1999). In general, up-scaling and down-scaling can also be described as aggregation and disaggregation methods respectively (Figure 1.13).

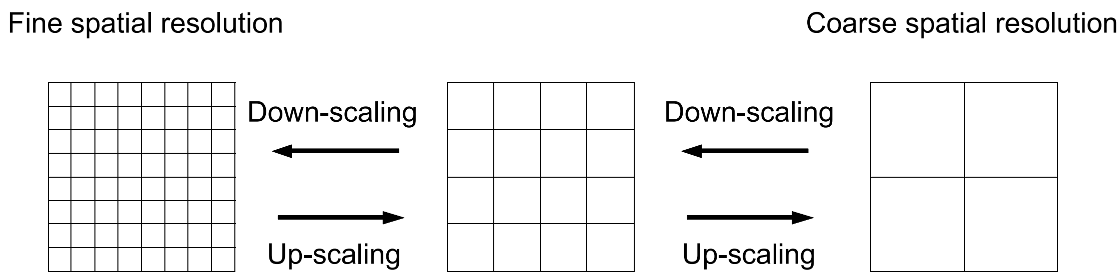


Figure 1.13. Scaling techniques: aggregation and disaggregation methods for implementing multi-scale analysis.

Remote sensing provides the desired data for up-scaling and down-scaling and the possibility of undertaking studies to understand the behaviour of variables when changing scale and derive appropriate rules for scaling (Marceau and Hay 1999). This study was provided with multiple spatial resolution data: 4, 8, 16, 32, and 64 m spatial resolutions, hereafter referred to by R_4, R_8, R_{16}, R_{32}, and R_{64} (Remmel and Perera 2009); hence a multi-scale analysis approach for characterizing spatial patterns across a gradient of scales would be performed.

While examining the issue of scale and scale effect in various aspects of spatial analysis, it is important to mention the concept of Modifiable Areal Unit Problem (MAUP). MAUP is a problem that occurs in spatial analysis of aggregated data in which the same basic data generates different results when aggregated in different ways (Wong 2009). For example, if the sizes of the pixels are changed or shift in location of the grid relative to the real scene on ground, it can then lead to a numerous datasets which will provide different results. An object (e.g., a residual patch) might have also different shape and size when derived from different images at different spatial resolutions; this problem is referred to as the MAUP (Openshaw 1984). There are two issues of concern related to the MAUP: scale and zonation; the MAUP involves both the effects of altered pixel size and the way of its alternation in a spatial context (Openshaw 1984). For example, in order to understand the spatial patterns of residual patches at landscape level, aggregation of fine resolution data (R_4) to coarser resolution data (R_{64}) is performed. This leads to a problem in spatial analysis where areal units are aggregated to different sizes; this is known as the aggregation effect of MAUP. The process in which the number of pixels kept constant or unchanged, but their arrangement changes is a zonal process which gives rise to various zonation or zoning effect; this involves a change in zones or grouping scheme (Wong 2009).

1.4.3. Scaling for characterizing spatial patterns

Understanding the effect of scale on detecting the patterns of spatial objects is an important step in landscape ecology or in assessing the relationship between patterns and processes, but there is no single all-encompassing scale at which all measurements can be made (Wiens 1989). It has been argued that geographic phenomena tend to have characteristic spatial and temporal scales or spatiotemporal domains (Allen and Hoesktsra 1992). Moreover, the amount of information available, variables that can be measured and the scale at which the process operate would not be the same across multiple scales. Thus, scaling theory becomes an important approach for characterizing the patterns over multiple scales; as variables and processes important at one scale may not be useful at another scale. If one changes the scale of reference, the phenomena of interest change, and information is often lost as observational scale changes (Riitters et al. 1995). For example, at the scale of a sub-event scale, it might be reasonable to ignore coarser-scale variability in temperature. Conversely, if the extent of our observational scale increases, the variability in temperature may also become important and should be accounted.

Moreover, spatial patterns and processes often occur over multiple scales (He and Mladenoff 1999) and there are hierarchical linkages among the scales; so information transformation among the scales is an essential component of landscape ecology (Wu and Li 2006). A central theme of landscape ecology is that particular phenomena should be addressed at their characteristic scales (Turner 1989), and hence a scaling rule should be established to understand the patterns across gradient of scales. However, a successful scaling strategy must address the complex aspects of ecological systems: scale-dependence process and spatial nonlinearities (Wu 1999). Observations made on a single scale can capture only those patterns and process pertinent to that scale of observation, but the complexity arises when an analysis involves multiple scales (Wu 1999). Scaling is also important when a prediction is desired to capture patterns at a certain scale (e.g., coarse scale or spatial resolution), based on information obtained at another scale (e.g., finer scale) (Wiens 1989; Wu and Li 2006).

1.5. Characterizing post-fire vegetation residual patches

1.5.1. Motivation

Wildfire is one of the natural factors affecting forest age structure, species composition, and forming heterogeneous landscape patterns following the disturbances (Chu and Guo 2013). Owing to the variation in fire intensity, disturbance size, vegetation cover, topography, fuel properties, and local weather, the patterns and characteristics of post-fire forests vary spatially

(Huang et al. 2006). Within a burned landscape, there may be some areas that escape fire, some areas that experience low-intensity fire, and some that experience high-intensity fire. Using the geoinformatics tools, researchers have been working to address the effects of wildfire, including pre-fire land cover mapping, assessing active fire characteristics, and charactering post-fire forest ecosystem responses. While a number of rigorous approaches have been developed to reflect the first two effects (Chu and Guo 2014); there is a lack of reliable and broadly replicable measurement approach for characterizing the spatial patterns and characteristics of post-fire forest conditions. Emphasis has been placed on understanding fire patterns and behaviour (e.g., fire spread, fire intensity, and fire severity); rather than the patterns of post-fire landscape structure.

Measuring forest structure following fire disturbances and characterizing their variability across different landscapes (and spatial resolutions) lays a foundation for assessing natural process (e.g., wildfire) because the heterogeneity in landscape elements (e.g., patches) influence the natural processes (Turner 1989; Turner et al. 1997; Blaschke et al. 2002). The processes, in turn, determine the formation of the spatial mosaic of landscapes (Kerby and Fuhlendorg 2007). Additionally, characterizing and measuring the spatial patterns of residual patches and their variabilities in composition and configuration: 1) provides baseline data for wildlife studies as they serve as habitat for different wildlife population over multiple scales; 2) helps to examine fire behaviour (fire intensity, severity, and spread) because a change in landscape patterns affects the subsequent patterns and behaviour of a wildfire (van Wagtendonk 2004); and 3) is useful for implementing disturbance-based forest management practices (Cuesta et al. 2009; Evans and Cushman 2009; Perera et al. 2007). For developing a framework for real world applications (i.e., forestry operations) that emulate natural disturbances, adequate understanding of forest disturbances and the characteristics of the subsequent landscape patterns are desired (Johnson et al. 1998).

Furthermore, knowledge about post-fire forest conditions across different spatial resolutions requires accurate, timely, and spatially explicit information on the patterns and characteristics of the residual patches. This relies on 1) how well we measure and understand the spatial patterns and characteristics of post-fire landscape structure (Andison 2013), and 2) how robust and consistent are the measurement frameworks for characterizing the patterns. The patterns and characteristics of residual patches could vary spatially across different landscapes or fire events as a function of topography, abundance of natural firebreak features, variations in fuel availability, and local weather. The measurement and quantification of the patterns of residual patches can also vary across different spatial resolutions as the composition and configuration of landscape structure are often sensitive to scale changes. However, there is a lack of a consistent, repeatable, and robust measurement framework for characterizing residual patches and assessing the scale effects on spatial pattering. There is a need for consistent methodologies and assessment tools that helps us mapping and identifying post-fire forest conditions, measuring the patterns and variabilities of residual patches across different landscapes and spatial resolutions, and identifying areas where residual patches are likely to occur. In response, this study uses different geospatial tools for the design and implementation of a repeatable, robust measurement framework for characterizing residual patches. The study also develops a systematic (modelling) approach to measure the learning rules that dictate areas where residual patches are likely to occur within fire disturbed landscapes and test the validity of the model on a large and independent fire event.

1.5.2. Scale for examining residual vegetation patches

Boreal forest fires involve factors and processes operating at different scales (King and Perera 2006) and thus the resulting patterns and variabilities can be studied on a wide range of scales (Pickett et al. 1999). The choice of an appropriate scale, or spatial resolution and extent, depends on the phenomenon under investigation (information desired about the surface properties), analysis method used to extract information, and research goals (Woodcock and Strahler 1987).

The effects of fire disturbances can also be studied at different geographic scales. Accordingly, a framework for depicting the effects of fire at different scales of space and time has been suggested (Moritz et al. 2011). This considers three geographic scales at which forest fire disturbance can be analysed: fire regime, wildfire, and flame (Moritz et al. 2011). A framework based on a different level of analysis – fire regime, fire event, sub-event, tree, and leaf level – can also be considered for assessing the effects of fire disturbances. For theoretical and practical reasons, analyses are often undertaken using large landscape units (e.g., fire regimes) (Cifaldi et al. 2004). Yet, at the fire event level, spatial patterns are common phenomenon and scale multiplicity is inherent in spatial heterogeneity (Wu et al. 2000). Therefore, the patterns of post-fire residual patches were examined at the fire event level, which is similar to the wildfire scale of Moritz's framework, but over multiple scales (i.e., different grain sizes or spatial resolutions), including R_4, R_8, R_{16}, R_{32}, and R_{64}. In this study, the term scale (multi-scale) is used to describe the 5 spatial resolutions considered.

1.5.3. Research objectives

A landscape ecological perspective provides information about how a change in fire behaviour influences the patterns of landscape structure (Haire and McGarigal 2009, 2010). An improved understanding of the characteristics of post-fire residual structure formed under natural conditions is also useful to develop rules for effective implementation of natural disturbance emulation strategies. This requires 1) a thorough examination of the characteristics of residual vegetation patches formed by wildfires; 2) understanding the factors that explain their occurrence, characteristics, and distribution; and 3) accurate and spatially explicit information that determine the occurrence and characteristics of residual patches. However, there is a lack of reliable or repeatable methods to measure and examine the characteristics of residual patches and predict the likely occurrence of residual patches within a fire disturbed landscape, given various environmental gradients. Therefore, this dissertation develops a replicable approach to study wildfire residual patterns in relation to the following research goals: 1) characterize the spatial patterns of post-fire residual patches, 2) assess the factors affecting residual patch occurrence, 3) develop a spatially explicit predictive model that generates probability maps for the existence of residual patches within a burned landscape, and 4) implement and validate the approaches with an independent dataset.

Moreover, the amount of information available and the variables that can be measured would not also be the same across multiple scales (He and Mladenoff 1999). For example, observations made at one scale may lose important information when operated on another scale; information often changes as observational scale changes (Turner 1989; Riitters et al. 1995; Kok and Veldkamp 2001; Perveen and James 2010). Additionally, when the scale of analysis is changed, different processes, responsible for the observation of patterns become increasingly evident (Benson and MacKenizie 1995). Therefore, it is also my goal to characterize the patterns and variabilities of residual vegetation patches across a spectrum of 5 spatial resolutions.

1.5.4. Research questions

Post-fire vegetation residual occurrence and pattern is the result of complex interactions of several factors (Bonan and Shugart 1989) that include climate (Foster et al. 1998; Turner et al. 1997; Swystun et al. 2001; Perera et al. 2007), time of burning, spatial extent and heterogeneity of the fire (Turner et al. 1997), variability in fire weather, elevation, and fuel conditions (Epting and Verbyla 2005), edaphic factors such as soil moisture and texture (Schroeder and Perera 2002), fire event geometry and behaviour (Perera et al. 2007), pre-fire characteristics, spatial variability, proximity to water surface (Turner et al. 1997; Perera et al. 2007; Cuesta et al. 2009) and topography (Haire and McGarigal 2009). Most of these inferences are based on post hoc observations rather than on testing a priori hypotheses; but observations made following the fire would be constrained in space and time and may not provide a comprehensive picture of the post-fire forest characteristics. This body of knowledge prompts one to quantify the spatial patterns, assess the geo-environmental factors that are associated with their occurrence, and test hypotheses on the combined effects of the geo-environmental factors on residual patch occurrence and distribution. To investigate these, different research questions (and sub-questions) are formulated:

1) What are the patterns and variabilities of post-fire residual patches?

   Sub-questions:
   - What are the spatial patterns and characteristics of residual vegetation patches?
   - Which measure(s) of spatial patterns are sensitive to scale change?
   - Is it possible to identify a scaling rule that determines the patterns across a gradient of scales?
   - Are certain land cover types less likely to burn than others?
   - What is the spatial association of residual patch occurrence in relation to natural firebreak features?

2) What are the predictor variables that explain the occurrence of residual patches within a disturbed landscape?

   Sub-questions:
   - What are the most important predictor variables that govern the occurrence of residual patches within burned landscapes?
   - What are the marginal effects of the most important predictor variables that explain the residual patches?

3) What is the combined effect of the geo-environmental factors that shape residual patch occurrence?

   Sub-questions:
   - Can a predictive model be developed to predict residual patches within a fire event?
   - What is the predictive performance of the model for determining residual patch occurrence?
   - Can the predictive model be inverted to build maps of likely residual stand locations?

1.5.5. The study area

The study is based on various fire events that occurred within the boreal forest ecozone of Ontario. Ecozones are areas of the Earth’s surface representing large and generalized ecological units characterized by abiotic and biotic factors (Wiken 1996). The following description of an ecozone follows William et al. (2009). An ecozone is a large area of land and water characterized by a distinctive bedrock domain that differs in origin and chemistry from the bedrock domain immediately adjacent to it. The characteristic bedrock domain has a major influence on the ecosystem processes and biota occurring, and hence on the patterns of disturbances in the region. Ecozones are ecosystem classification systems that are defined based on key abiotic processes functioning at national and continental scales. The Canadian system of ecosystem classification divides the country into twenty major units, 15 Terrestrial Ecozones and 5 Marine Ecozones (Wiken 1996). Of 15 Terrestrial Ecozones in Canada, the Hudson Bay Lowlands, the Boreal Shield, and the Mixed Plains Ecozones occur in Ontario (Figure 1.14). The Boreal Shield Ecozone refers to an area where the Canadian Shield and the boreal forest overlap, and is the largest ecozone in Canada, stretching 3,800 km from Newfoundland to Alberta; covering more than 1.8 million square kilometres (20% of Canada’s land area) (Wiken 1996). The Boreal Shield Ecozone within Ontario is also known as the Ontario Shield Ecozone, the largest ecozone in Ontario. The fire events considered in this study are contained within the Boreal Shield Econzone, specifically within the Ontario Shield Ecozone (Figure 1.14). The boreal forest region in Ontario in general and in the Ontario Shield ecozone in particular is further divided into ecoregions (e.g. 0E, 1E, 2W, 2E, 3E, and 3S) based on geoclimatic patterns (Hills 1961) (Figure 1.14). An ecoregion is a unique area of land and water nested within an ecozone that is defined based on different climatic variables, including temperature, precipitation, and humidity (William et al. 2009; McKenney et al. 2010).

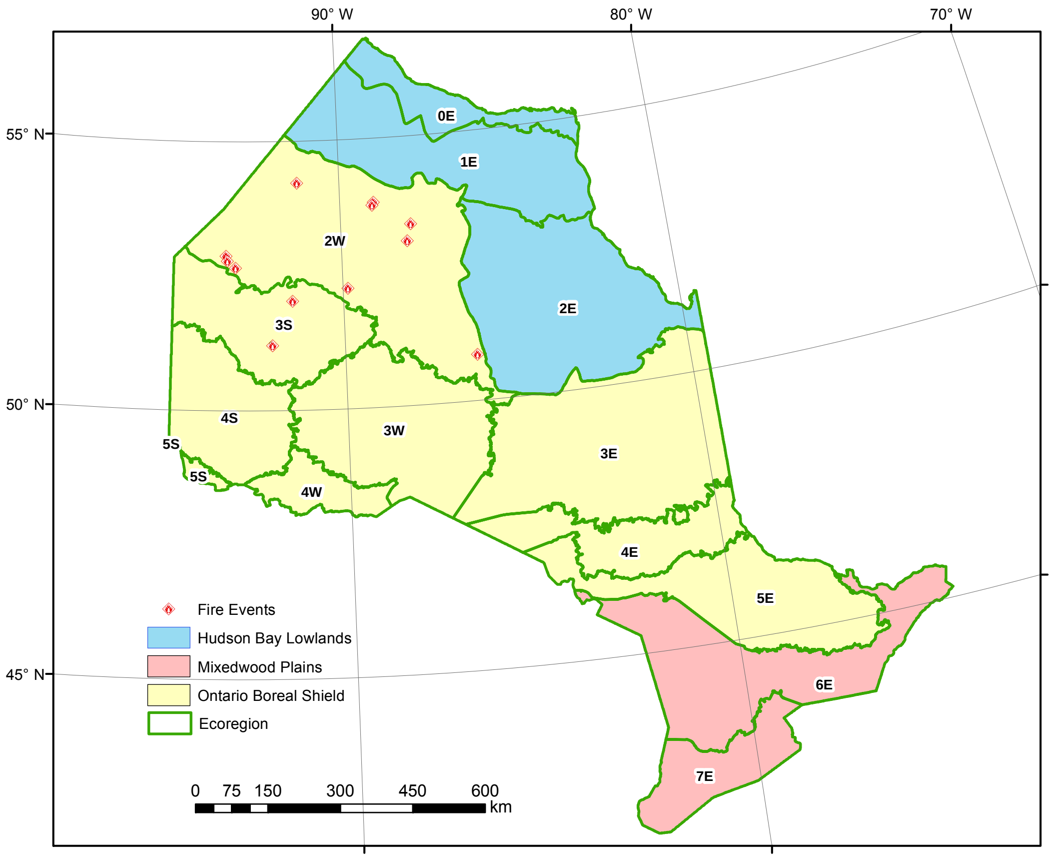


Figure 1.14. Locations of the 12 fire events studied in dissertation in relation to Ontario's ecozones and ecoregions.

The Ontario Shield is generally characterized by relatively cold and moist (and long) winters and short, warm summers (Thompson 2000). Owing to its large geographic extent, however, there is a wide range of temperature, precipitation, and humidity patterns in the ecozone; with annual precipitation ranges from 500 mm in the west to 850 mm in the east and daily temperature that ranges from -15˚C in January to 17˚C in July (William et al. 2009). The conditions in the southern part of the ecozone are more moderate. Similarly, the topography varied depending on both local bedrock and surficial deposits. The region is heavily forested, with open water (lakes and rivers), wetlands (peatlands), and shrubs. The northern part of the ecozone is dominated by coniferous species such as black spruce (*Picea mariana* [Mill.] B.S.P), white spruce (*Picea glauca* (Moench) Voss), balsam fir (*Abies balsamea* [L.] Mill.), jack pine, (*Pinus banksiana* Lamb) tamarack (*Larix laricina* [Du Roi] K. Koch.), and intolerant hardwoods including white birch (*Betula papyrifera* Marshall) and trembling aspen (*Populus tremuloides* Michx.). The southern portion of the ecozone is predominantly occupied by conifer species (e.g., pine) and mixed and deciduous forests of tolerant hardwoods, including oak (*Quercus* L), sugar maple (*Acer saccharum* [Marshal]) and American beech (*Fagus grandifolia* [Ehrh]). The ecozones are further subdivided into ecoregions, which are characterized according to broad but spatially explicit environmental features such as climate, geology, terrain, vegetation, soils, and water, as well as regional human activity (Wiken 1996).

This study focuses on the northwestern part of the ecozone (and the province) where the conifer-dominated boreal forests are naturally and frequently affected by fire disturbances. Wildfire is the dominant force of natural change in the region, but the frequency, intensity, and size of burns vary depending on climate, forest type, and local landscape features (Thompson 2000). The high risk of fire severity and a longer fire season in the region change the natural disturbance regime in many forest types (Flannigan and Wotton 2001). Therefore, emulating natural disturbances has been suggested as an important management paradigm for achieving sustainable ecosystem management. On specific way to emulate natural disturbances is to have better understanding of the patterns of fire disturbances and post-fire landscape structure. This study examines the patterns of residual patches and importance of various factors at the fire event level as part of an effort to provide a consistent and repeatable method for understanding the characteristics residual patch occurrence, probabilities and variability of residual patches, mechanism and causal factors of residual structure.

2. Characterizing spatial patterns of post-fire residual patches in boreal wildfires

Abstract

Wildfires typically contain a considerable number of unburned residual patches of various size, shape, and composition. These residual vegetation patches can occupy substantial areas of fire footprints; thus understanding the patterns of residual patches provides insights for emulating forest disturbances in harvesting operations. In this study, eleven boreal wildfire events within Ontario; each one ignited by lighting and never suppressed are studied. The spatial patterns of post-fire residual patches are assessed based on selected spatial metrics (related to composition, configuration, and fragmentation). Characterizing the occurrence of post-fire residuals, their spatial patterns, and variability at multiple spatial resolutions is also imperative to examine the effects of analytical scales because spatial patterns are scale dependent. One way of understanding these relationships is to examine how the patterns of residual patches change with scale. The effects of analytical scale (i.e., spatial resolution) on characterizing the spatial patterns are assessed; the patterns were examined at five spatial resolutions R_4, R_8, R_{16}, R_{32}, and R_{64}. The results show that the responses of the landscape metrics can be grouped into three categories: monotonic and predictable response, monotonic change with no simple scaling relationship, and non-monotonic change with erratic responses. The study also finds that certain land cover types that are less abundant on the landscapes (e.g., treed wetland and sparse conifer) dominate the residual patches in some fire events (e.g., F01 and F04).

Keywords: residual patches, spatial patterns, analytical scale, scale effect, spatial metrics, landscape metric scalograms, land cover, fire footprint

2.1. Introduction

Landscapes are complex and heterogeneous land areas containing patterns formed by different forms of disturbances (Forman and Godron 1986; Linke et al. 2007). Specifically, a forested landscape often changes in response to different elements, including fire disturbance, insect infestation, global changes in climate, and human activity (Perera and Euler 2000). The most apparent process and change in a forested landscape is disturbances from wildfire. This has been responsible for the formation of heterogeneous elements within a landscape. Spatial heterogeneity in a landscape has a close relation with stability and biodiversity where high heterogeneous landscape encourages interactions (Duning and Xiezhen 1999).

Wildfire is a major natural disturbance and an important factor that shapes the landscape structure in the boreal forests. Fires in boreal forests are often intense and frequent (Johnson 1995; Cui et al. 2009) and consume substantial forest cover (Perera et al. 2009b), but do not burn the entire landscape (Whelan 1995; Johnson et al. 1998; Leduc et al. 2007). Owing to the variations in weather and site conditions (e.g., vegetation, topography, and natural firebreaks) (Rochadi et al. 1999; Perera et al. 2007), forest fire shapes the patterns of forest structure (Agee 1998; Linke et al. 2007; Hely et al. 2010) and creates a complex and heterogeneous landscape mosaic comprising patches of different size, age, shape, and tree species compositions (Turner

1989; Diaz-Delgado et al. 2004; van Wagtendonk 2004; Mermoz et al. 2005; Madoui et al. 2010; Vinatier et al. 2010). The spatial patterns of post-fire landscape structure (e.g., forest land cover) are useful to understand various ecological processes such as species dynamics and fire disturbances. The patterns also have direct implications for various aspects of forested landscapes: economic values (i.e., selection of sites for harvesting), social concerns (i.e., conservation of wilderness) (Thompson 2000) and ecological values (i.e., habitat for various organisms). Understanding the patterns of post-fire residual structure helps forest managers to determine the structural elements that should be retained to emulate fire disturbances and preserve the biological diversity of the ecosystems.

One particular way of understanding fire disturbances and their effects is assessing the patterns of landscape structure following a fire. Wildfires affect the physical landscape structure, age class distributions, ecotones, and positions of forest boundaries (Weber and Flannigan 1997). This study focused only on forest landscape structure, referring to the pattern of a landscape that is determined by its type of use and its structure (i.e., size, shape, arrangement, and distribution of landscape elements (patches, corridors, and matrix) (Walz 2011). In this study, the term landscape structure refers to the patterns of post-fire residual patches, specifically the composition, arrangement, and the resulting spatial relationships among individual patches.

2.1.1. Post-fire residual patches

Wildfire is one of the main natural disturbances consuming substantial forest cover, influencing and reshaping the landscape mosaic of boreal forests (Madoui et al. 2010). One of the characteristic features of wildfires is the existence of unburned areas within a fire-disturbed landscape, which are referred to as post-fire residuals. The presence of a residual patch is due to different geo-environmental factors that interactively affect fire behaviour and the resulting patterns of post-fire landscapes. The term residual patch is broadly defined as remnants of the pre-fire forest ecosystems that have retained their structure and were not entirely reduced to ash or charcoal during the fire. In this study, the term residual patch is used to describe remnants of the pre-fire forest ecosystems (i.e., live undisturbed vegetation patches) that are not physical connected to the footprint perimeter; this is regardless of size, age, and species composition of the patches. For detailed description on the types and meanings of different patches and fire footprint, please refer to (§1.3).

2.1.2. Landscape pattern metrics (LPM) for spatial pattern analysis

Understanding the patterns of residual patches plays an important role in inferring ecological processes such as fire disturbances and species dynamics (Griffith 2004; Mermoz et al. 2005; Vinatier et al. 2010); this has been central to the study of landscape ecology (Diaz-Delgado et al. 2004). In dealing with landscape ecology, the basic characteristics of a landscape (i.e., structure, function, and change) should be understood (Forman and Godron 1986; Turner 1989). Landscape structure has been used extensively in the landscape ecological literature, primarily to describe both landscape composition and configuration (Gustafson 1998; Linke et al. 2007). A landscape's composition is described by the number of categories and amount of different spatial elements within a landscape but without being spatially explicit (McGarigal et al. 2002; Remmel and Csillag 2003; Linke et al. 2007). Landscape configuration refers to the physical distribution of patches within the landscape (McGarigal and Marks 1995; Remmel and Csillag 2003; Griffith 2004; Cifaldi et al. 2004; Lin et al. 2010). In order to understand the interaction between spatial patterns and process, the spatial heterogeneity of a landscape must be identified and quantified in meaningful ways (Turner 1989; Wu et al. 2000; Blaschke et al. 2002).

One of the characteristic features of a wildfire is the tendency to generate important biological diversity, which is used to describe the degree of heterogeneity in ecosystem structure and composition (Burton et al. 2008). One particular way of addressing the spatial heterogeneity in a landscape is by computing series of landscape pattern metrics (LPM) (Turner 1989; Corry and Lafortezza 2007); hence, an emphasis has been placed on developing methods to quantify landscape structure. LPM refer to indices obtained from categorical maps, and are focused on the characterization of the geometric and spatial properties of landscape patterns (McGarigal et al. 2002). The metrics have been widely used to characterize spatial heterogeneity, infer ecological processes (e.g., forest disturbances and species dynamics) (Riitters et al. 1995; Forman and Godron 1986; Griffith 2004; Lin et al. 2010).

Boreal forest fires involve factors and processes operating at different scales (King and Perera 2006) and thus the resulting patterns and variabilities can be studied on a wide range of scales (Pickett et al. 1999). However, there is considerable uncertainty regarding the appropriate scale at which measurements and analyses are undertaken (Griffith 2004; Cifaldi et al. 2007; Linke et al. 2007). LPM used to measure landscape structure relies on digital spatial data; yet the characteristics of the data are constantly changing depending on how scale is defined (Turner 1989; Corry and Lafortezza 2007) and grain sizes are aggregated (He et al. 2002). Additionally, spatial patterns manifest as processes operate over multiple spatial scales (Turner 1989; Ostapowicz et al. 2008; Wu et al. 2000); hence interpretation based on data from one scale may not apply to another (Perveen and James 2010). Because of this multiplicity, scale holds the key to understanding pattern-process interactions; this has led to the hierarchical perspective in landscape ecology. The choice of an appropriate scale depends on the analysis method used to extract information about the phenomena and specific research objectives investigated (Woodcock and Strahler 1987). The relationship between patterns and scale has also been an integral component landscape ecology (Wu and Li 2006), and as a result the definition of scale has to be well established.

2.1.3. Scale and its importance for pattern analyses

Various researchers have approached the issue of scale and scaling from related but different perspectives. In landscape ecology, for example, the scaling of patterns and processes is often addressed by considering multiple scales at which spatial pattern analyses are undertaken (e.g., Benson and MacKenzie 1995; Moody and Woodcock 1995; Wu et al. 2002; Zhu et al. 2006). Scale in landscape ecology involves both grain and extent, which are related to the spatial resolution of a given study area and area of coverage respectively. To understand the scale effect, the spatial patterns over multiple scales should be studied and hierarchical linkages among them should be established using scaling approaches. There are two approaches to multi-scale analyses: 1) the direct method that uses inherently multiple scale approaches, and 2) the indirect multi-scale method that uses single-scale methods repeatedly at different scales (Wu et al. 2000). The direct methods contain multiple-scale components in their mathematical formulation or procedures, and thus are either hierarchical or multiple-scaled (Wu et al. 2000). Some of the direct methods used in landscape ecology include wavelet analysis, lacunarity analysis, and spectral analysis. The indirect approach on the other hand can use methods that are designed for single-scale analysis, such as the wide variety of landscape metrics (e.g., shape and area related metrics) as well as statistical measures (mean, variance, correlation, and regression coefficient). The most common approach to study the scale effect issue of scale, and implemented in this study, is the indirect approach, (Wiens 1989). The indirect methods was applied because it allows one to compute the various aspects of spatial patterns (composition, configuration, and fragmentation), and compare the LPM over multiple scales.

2.1.4. Research framework

Several studies have described the spatial patterns of natural fires (Diaz-Delgado et al. 2004; van Wagtendonk 2004; Mermoz et al. 2005; Collins et al. 2007; Meddens et al. 2008; Hely et al. 2010; Dragotescu and Kneeshaw 2012). Despite their importance for understanding fire disturbances and species dynamics, there are relatively few studies undertaken to characterize the spatial patterns of residual patches and their spatial distribution within a disturbed landscape (Schmiegelow et al. 2006; Madoui et al. 2010). Moreover, previous studies examined the effect of scale change on measures of spatial structure (e.g., Turner 1989; Wiens 1989; Benson and MacKenzie 1995; Gustafson 1998; Nikora et al. 1999; Kok and Veldkamp 2001; Wu et al. 2002; Zhu et al. 2006; Haire and McGarigal 2009). The studies revealed that the LPM are sensitive to the changes in grain size, but the response of LPM to changing grain size varies depending on the dataset and aggregation techniques applied. In spite of this, our understanding of the interactions between scale and landscape pattern is limited; hence landscape pattern quantification remains an important issue for investigation and more analyses are needed to characterize patterns over multiple scales (Leduc et al. 2007). Moreover, little is known about the effects of scale on parameters that characterize the spatial structure for fine spatial resolution data (i.e., less than 30 m); yet this spatial resolution remains useful for detecting spatial structure (Corry and Lafortezza 2007).

The objectives of this chapter are to: 1) characterize the spatial patterns of post-fire residual vegetation patches, 2) examine the sensitivity of the metrics to changing grain size, 3) identify some general rules for comparing LPM obtained at different scales (i.e., establishing scaling rules), 4) assess the impact of land cover types on residual patch occurrence (i.e., are particular land cover types more likely to generate residual patches?), and 5) evaluate the spatial association of natural firebreaks (surface water) and fire perimeter with the occurrence of residual patches. To achieve these objectives, eleven spatial metrics (derived at landscape and class level) were analysed, across five spatial resolutions R_4, R_8, R_{16}, R_{32}, and R_{64}.

In this chapter, I hypothesized that 1) the metrics that characterize the patterns of post-fire landscape structure would not be consistent across the fire events as the fire size and fire intensity would vary across the fire events; 2) the pattern and characteristics of residual patches, with respect to size, shape, and composition is sensitive to scale change; 3) the occurrence of residual patches is explained with respect to certain non-burnable areas, and 4) a tendency exists for residual patches to be concentrated near surface water. By addressing these, the study serves as a proxy for understanding the range of variability for fire in forest ecosystems (Collins et al. 2007). This study also provides useful information concerning the status and dynamics of boreal forests, and elements for long-term benchmark monitoring and conservation related to disturbances.

2.2. Methods

2.2.1. Study area

Wildfires in boreal forests are sometimes stand-destructive and large that may burn tens of thousands of hectares, but small fires that burn areas of less 100 ha are the most frequent (Thompson 2000). In this study, the patterns and characteristics of post-fire residual patches were examined based on 11 fire events occurred in northwestern Ontario, having footprint areas ranging from approximately 58 to 4225 ha (Table 2.1). The focus is placed on 11 fire events of different sizes but it is has to be noted that it is the largest fires that are primarily responsible for the change in landscape structure. Yet, frequent wildfires are very common in boreal forest; such recurring fire behaviour is also responsible for the occurrence of different landscape structure, specifically residual patches.

Table 2.1. The study fire footprints, corresponding fire areas and burn year, and the proportion of residual patches. The fire footprint codes (fire ID) are those used in Remmel and Perera (2009).

Fire Footprint ID *	Burn year	Fire Footprint Extent (ha)	Total unburned area (ha)	% of unburned area: fire perimeter
F01	2002	4525.3	1466.7	32.4
F02	2002	80.6	20.7	25.7
F03	2002	80.5	17.0	21.1
F04	2002	1574.9	341.9	21.7
F05	2002	2286.1	720.8	31.5
F06	2002	3741.8	1634.0	43.7
F07	2002	940.5	287.0	30.5
F08	2003	3072.2	1212.5	39.5
F09	2002	57.7	11.2	19.4
F10	2003	3276.9	945.7	28.9
F11	2003	719.3	105.3	14.6

The 11 fire events are located within one of the largest ecoregions in Ontario (i.e., Big Trout Lake ecoregion – 2W), which sits within Ontario Boreal Shield Ecozone (Figure 2.1). The three ecozones within Ontario are described in (§1.5.5). The boreal forest region in Ontario in general and the Ontario Shield ecozone in particularly are divided into ecoregions (e.g., 2W, 3E, 3S, 3W, 4E, 4S, 4W, 5E, and 5S) based on geoclimatic patterns (Hills 1961). An ecoregion is a unique area of land and water nested within an ecozone that is defined based on different climatic variables, including temperature, precipitation, and humidity (William et al. 2009; McKenney et al. 2010). The climate within an ecoregion influences the vegetation types, soil formation, and other ecosystem processes (e.g., forest disturbances), and associated species. This in return affects wildfire disturbances and the subsequent post-fire landscape structure. The fire events considered in this study are within the 2W ecoregion that is within the Ontario Ministry of Natural Resources (OMNR) extensive fire management zone, where fires are monitored and recorded, but not actively suppressed or harvested (Perera et al. 2009a). The spatial extent of each of the eleven fire events is shown in (Figure 2.2).

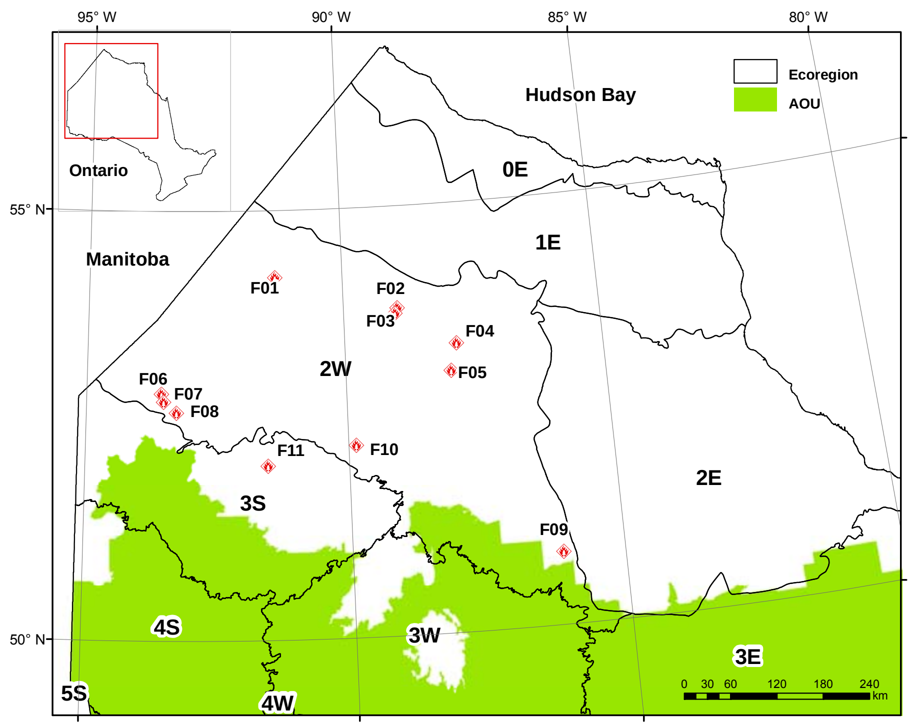


Figure 2.1. Location of the 11 fire events in relation to Ontario's ecoregion and the Area of Undertaking (AOU). The AOU is the area where forest harvesting operations are permitted.

This ecoregion (2W) is located in a cold and dry part of the province. The climate is characterized by long, cold, and dry winters, and short, warm and moist summers (Runesson 2011); with a mean annual temperature ranging from -4.1 to -0.1 °C, a mean growing season lengthy of 147 to 170 days, mean annual precipitation of 550 to 786 mm, and mean summer rainfall between 222 and 297 mm (Williams et al. 2009). The relief is characterized by flat plains, undulating upland areas and dissected uplands with ridges and escarpments (Baldwin et al. 2000). The landscape is also characterized by extensive peatlands in low-lying areas (William et al. 2009). The region is primarily dominated by coniferous species (approximately 41% of the region is covered by this cover type) while more than 30% of the ecoregion is covered by various types of wetlands and surface water. The region is specifically characterized by mix of coniferous and deciduous species, typified by jack pine (*Pinus banksiana* Lamb), white spruce (*Picea glauca*

(Moench) Voss), black spruce (*Picea mariana* [Mill.] B.S.P), paper birch (*Betula papyrifera* sp.), balsam fir (*Abies balsamea* [L.] Mill.), and trembling aspen (*Populus tremuloides* Michx.). This ecoregion is susceptible to fire, but they are generally smaller than those in the southerly ecosystems in northwestern Ontario.

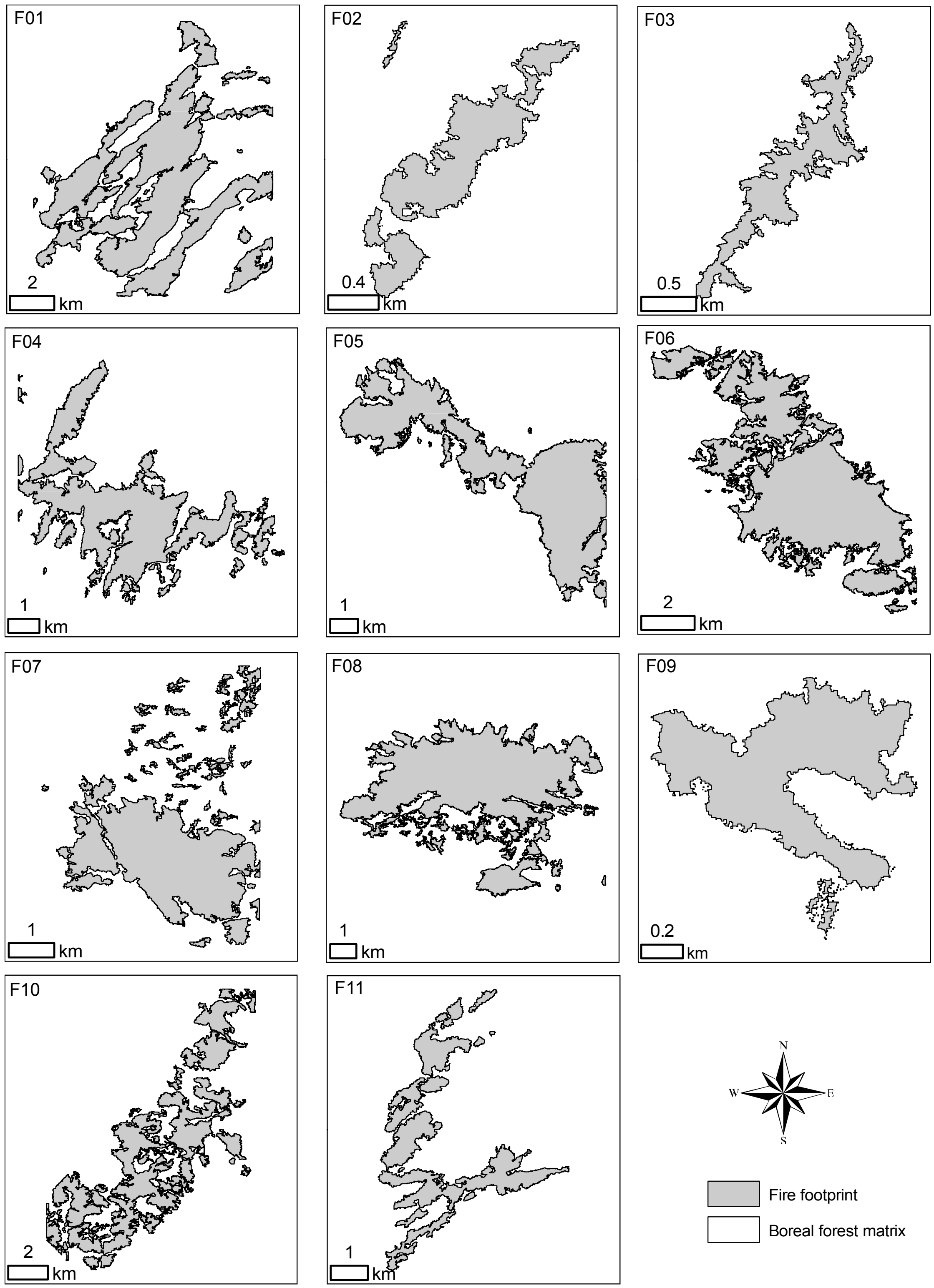


Figure 2.2. The spatial extent of the fire footprints derived from the classified Ikonos images for the eleven fire events used in the study.

2.2.2. Landscape data

The landscape data used in this study are based on post-fire vegetation residual maps delineated and extracted from classified Ikonos imagery captured between 1 June and 7 August 2005, with 14 land cover categories (Table 2.2) (Spectranalysis 2005). A standard classification process, without any specific emphasis on the fires and residuals, was implemented to classify the Ikonos imagery (Spectranalysis 2005). The base data for this study, maps of post-fire vegetation maps, were provided at multiple spatial resolutions: R_4, R_8, R_{16}, R_{32}, and R_{64}, which were obtained from a previous work by (Remmel and Perera 2009). In order to generate the multi-resolution data, they used a majority rule-based spatial aggregation method, such that the block size represented the desired spatial resolution (R_4, R_8, R_{16}, R_{32}, and R_{64}) within which the thematic majority class was assessed (Remmel and Perera 2009). The details about the approaches used to generate maps of residual patches for the eleven fire events, over the five spatial resolutions are documented in Remmel and Perera (2009).

Generally, there are two approaches to generate categorical spatial patterns with multiple spatial resolutions for a given landscape (Saura 2004). The most common and simpler method is spatial aggregation, which can be implemented using nearest neighbour techniques, majority rule-based, or random-rule based aggregation. Using nearest neighbour techniques often maintains the global proportion of each category in the original map but can lead to disaggregation. Spatial aggregation based on the majority rule have been used in landscape ecological studies based on the premises that 1) dominant classes increase in abundance while minor classes decrease in abundance or even disappear through aggregation processes; and 2) spatial patterns change with aggregations (He et al. 2002).

The second approach is directly classifying simultaneously gathered satellite images covering the same area but with different sensor spatial resolutions (Saura 2004). One of the problems with the latter approach is that satellite images from different sensors covering the same area and in specific time are not easily available. The approach, based on different sensors, has been less commonly used in landscape ecological studies (Benson and MacKenzie 1995). Therefore, the majority rule-based aggregation method has been used throughout this study to mimic the multiple spatial resolution data: R_4, R_8, R_{16}, R_{32}, and R_{64} (Remmel and Perera 2009). This is using an independent aggregation, scheme where the aggregation at each successive grain size always starts with the base data, as opposed 'iterative' aggregation scheme in which the aggregation at the next grain size is based on the already aggregated data of the initial grain size (Wu et al. 2002).

The aggregations were sequential, from 4 to 64 m spatial resolutions, with double increment. The five spatial resolutions were considered because 1) the aggregated spatial resolutions at R_4, R_8, R_{16}, R_{32}, and R_{64} ensured no re-sampling of partial pixels; no pixel was divided by the aggregation procedures, 2) the selected spatial resolutions potentially reflect the minimum patch size of the landscape guide (OMNR 2010); the minimum mapping unit or the smallest possible feature that can be mapped at R_{32} is 0.1024 ha, which is equivalent to the minimum patch size (0.10 ha) of the forest management guide. The spatial resolutions perfectly straddle this threshold, with R_4, R_8, and R_{16} being finer and R64 being greater than this threshold; 3) the selected spatial resolutions are similar to how some remote sensing device would view the landscape at different spatial resolutions. For example, 4 m spatial resolution is similar to datasets derived from Ikonos; 8 m is the closest to the 1:20,000 cartographic scale typical for aerial photography-based mapping associated with northern Ontario's forest management practices (Perera et al. 2009a); and 32 m is the closest to the datasets obtained using Landsat image.

Table 2.2. Categories of land cover obtained from IKONOS image classification. These categories follow the classes of the 2000 Ontario Provincial Land Cover Database (OMNR 2005).

Land cover category	**LCID**	**Description**
Complete burn	CB	Vegetated areas burned over their full extent, showing little or no evidence of vegetation
Partial burn	PB	Vegetated areas burned over part of their extent, showing evidence of sparse or scattered vegetation
Old burn*	OB	Old burns where charring is still evident but regeneration appears
Dense conifer	DC	Dense, predominately coniferous forest that may include some minor component of deciduous species
Sparse conifer	SC	Sparse, predominately coniferous forest which may include some component of deciduous species
Deciduous	DE	Dense, predominately deciduous forest which may include some minor component of coniferous species
Alder shrub woodland	AS	Alder shrubs with some large trees occurring almost exclusively along watercourses
Low shrub	LS	Low shrub areas that may include grasses but do not support trees, found in proximity to lakes, on the deltas of watercourses, and on old burns
Treed wetlands	TW	Bogs and fens with tree cover
Open wetlands	OP	Bogs and fens without tree cover
Water	WA	Surface water; includes some extensive string bogs
Marsh	MA	Inundated areas with emergent vegetation adjacent to surface water
Bedrock and non-vegetated	BV	Areas with little or no vegetation, primarily bedrock outcrop
Cloud and shadow	CS	Image areas containing no usable data because of cloud and shadow effects

*Old burn is obtained based on a pre-fire land cover map.

2.2.3. Landscape pattern analysis

Landscape pattern metrics should be carefully selected based on their minimal correlations among the indices used, their simplicity, and their sensitivity to landscape variations (Wu 2004). In this study, eleven measures (shown in Table 2.3) are selected based on previous work involving landscape metrics (Krummel et al. 1987; O'Neill et al. 1988; Turner 1989; Riitters et al. 1995; Cain et al. 1997; Meddens et al. 2008; Cuesta et al. 2009). The metrics are believed to explain the effects of scale on pattern analysis and the impact of land cover on residual patch occurrence.

Table 2.3. List of landscape pattern metrics used in the study: symbols, measurement units, and description. The level of analysis includes (L – landscape level and C – class level). Naming and scaling conventions are those of McGarigal and Marks (1995).

Metric name	**Symbol and Range**	**Description**	**Scope**	
Class area	CA – CA < 0	How much of the landscape is comprised of residual vegetation patches (ha)	C/ L	CA close to 0 when patches become rare; CA = total area as the landscape is dominated by a single patch.
% of landscape	%LAND – 0 < %LAND≤100	Computes the percentage of landscape occupied by residual patches (%)	C	%LAND is close to 0 when the patch type is increasingly rare
Largest patch index	LPI – 0 < LPI≤100	The ratio of the area of the largest residual patches to the total area of the landscape (%)	C/L	LPI approaches 0 when the largest patch becomes increasingly smaller; LPI = 100 when the largest patch comprises 100% of the landscape
Number of patches	NP – NP ≥ 0	The total number of residual patches in the landscape (fire event). CL and LL	C/ L	NP = 1 when the landscape contains 1 patch
Patch density	PD – PD > 0	The number of residual patches per unit area (per ha). CL and LL	C/ L	Higher PD higher spatial heterogeneity
Mean patch size	MPS MPS > 0	The average area of all patches in the landscape (ha)	L	Smaller MPS indicates heterogeneous landscape
Patch size standard deviation	PSSD PSSD ≥ 0	The standard deviation of patch size in the entire landscape (ha). It is a measure of absolute variation; it is a function of the mean patch size and the difference in size among patches	L	PSSD = 0 when all patches in the class are the same size or when there is one patch
Patch size coefficient of variation	PSCV PSCV ≥ 0	Measures of relative variability about the mean (i.e., variability as a percentage of the mean) (%). It is misleading in the absence of NP and PD	L	PSCV approaches 0 when the variability in patch size is small
Landscape shape index	LSI LSI ≥ 1	Measures the complexity of patch shape compared to a standard shape. In raster version, patch shape is evaluated based on a square as s standard shape	L	LSI = 1 when the landscape consists of a single patch of the corresponding type is circular (vector) or square (raster)
Mean shape index	MSI MSI ≥ 1	A patch level shape index average over all patches in the landscape	L	MSI = 1 when all patches of the corresponding patch type are circular or square; MSI increases as the patch shapes become more irregular
Mean fractal dimension	FRAC 1≤FRAC≤2	The summation of fractal dimension for all patches divided by the total number of patches in the landscape	L	FRAC approaches 1 for shapes with very simple perimeters (circles or squares)
Mean nearest neighbour distance	MNN MNN > 0	The average of the shortest distances between patches of the same type within the landscape	L	MNN close to 0 indicates patches of the corresponding patch type are close to each other; MNN is none if there is only one patch

The metrics were also selected to represent and measure different components of landscape structure: 1) indices that explain the overall landscape composition (CA, LPI, and %LAND), 2) spatial metrics that describe residual patch configuration (LSI, MSI, FRAC, and MNN), and 3) indices that explain the fragmentation (NP, PD, MPS, PSSD, and PSCV). The parameters that characterize the spatial characteristics of landscape structure were studied and quantified at the landscape and class levels of landscape structure. At the landscape level, aggregate characteristics of residual patches are studied, irrespective of land cover classes while information about patches of a given land cover type (i.e., vegetation residual patches) are obtained at the class level. The parameters concerning these levels of spatial heterogeneity provide different kinds of information on spatial patterns, including the level of fragmentation within the landscape, the type and composition, and the spatial arrangements of residual patches within the landscape. The landscape parameters are computed using FRAGSTATS 3.3 (McGarigal et al. 2002) using the binary raster post-fire residual patch maps.

2.2.4. Spatial patterns of residual patches and scale effect

Scale multiplicity is inherent in spatial heterogeneity (Wu et al. 2000); the effects of scale are thus inevitable for investigating landscape structures. Multi-scale approaches have also been suggested to quantify the effects of scale on landscape characteristics. In this study, the indirect approach to multi-scale analysis, which is related to the spatial aggregation problem, is used to characterize the patterns of residual patches. The approach uses a single-scale method by repeatedly measuring parameters at multiple scales. Understanding the pattern and structure along a hierarchical scaling ladder can be implemented by changing grain size, extent, or both across successive domains of scale (Wu 1999). The scale multiplicity is performed by changing the grain sizes into the desired spatial resolutions while the extent and thematic resolutions were kept constant, and repeatedly computing the metrics using the spatially aggregated data (R_4, R_8, R_{16}, R_{32}, and R_{64}). The spatial structure of the residual patches, at the landscape and class levels, is quantified using the parameters listed in Table 2.3. Information obtained at the landscape level was used to assess the sensitivity of the indices to scale changes.

2.2.5. Spatial distribution of residual patches: effects of land cover

The occurrence of residual patches is attributed to various factors (e.g., weather variables, ignition sources, fuel, vegetation, and topography) that interactively affect fire behaviour and the subsequent post-fire landscape structure (Collins et al. 2007; Hely et al. 2010). The impact of climatic conditions are usually manifested at a broader geographic scale (e.g., regional or landscape level) overlong period of time while other variables (e.g., fuel, topography, natural firebreaks, or vegetation types) can be realized at finer scales (e.g., fire event). The topography of Ontario includes flat plains, undulating uplands and dissecting uplands with ridges and escarpments. These kinds of topographic features can have a considerable impact on the post-fire landscape structure. Based on the 1 km DEM of Ontario, elevation of the 1 km grid ranges from a minimum elevation of sea level, around Hudson Bay, to a maximum elevation of about 610 m, west of Thundery Bay (Mackey et al. 1994). However, the topography within each of the fire events does not vary substantially (Figure 2.3); this is based on DEM data that vary in resolution from 0.75 arc seconds to 3 arc seconds, and vertical resolution within 5 m.

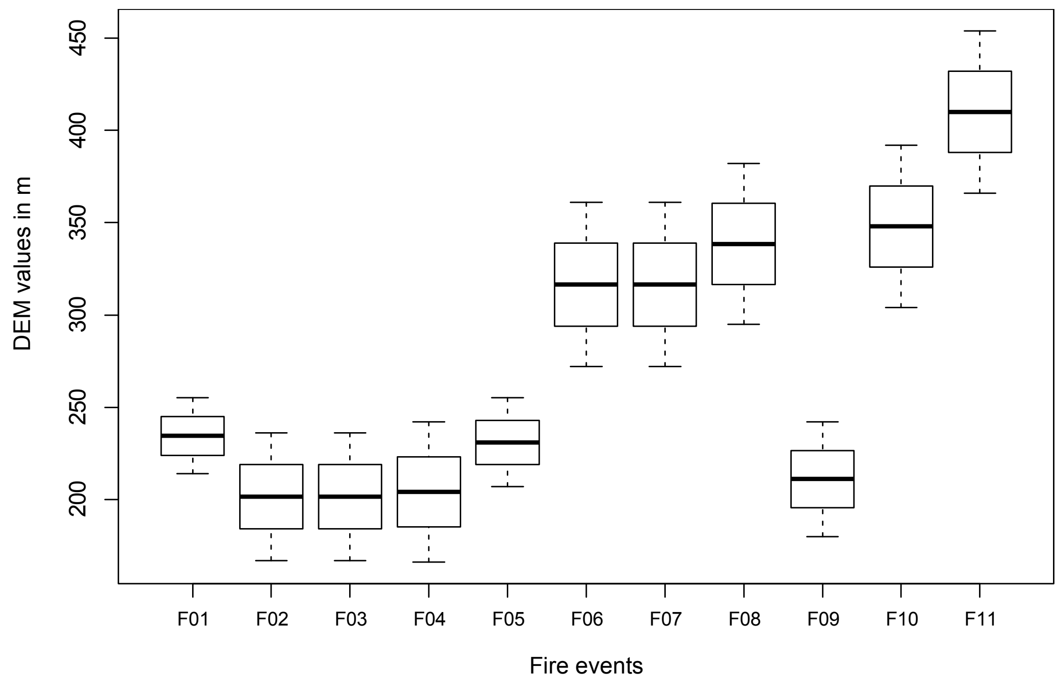


Figure 2.3. A box plot that shows the variability of DEM values across the 11 fire events in ecoregion 2W; each box in the plot is based on the DEM values obtained at five different spatial resolutions.

Furthermore, vegetation cover types, depending on the fuel load, play an important role in determining the patterns of fire spread, and the subsequent post-fire landscape structure (Mermoz et al. 2005). This is attributed to the fuel load or moisture content of fuel particles of the cover types (Nelson 2001). For example, vegetation cover with high moisture content can slow the rate of burning; even prevent fire spread and eventually affect the patterns of post-fire landscape structure. Owing to the variation in fuel type (i.e., whether they are alive or dead) and amount of fuel available and its spatial distribution, the impact of vegetation cover on the patterns of residual patches can vary (Viegas 1993; ESA 2002; van Wagtendonk 2004). Therefore, knowledge of the types of vegetation cover and their spatial distributions is necessary for understanding the wildfire processes and managing natural resources. The tendency for certain cover types to escape burn (and create residual patches within a burned landscape) was thus examined based on information extracted at the class level of landscape analysis.

2.2.6. Spatial association with fire breaks and fire perimeter

In order to determine the proximity or spatial association of residual patches to surface water, the proportion of residual patches within increasing distance bands from the water were computed. Initially, a series of external 100 m wide buffers rings that extended outward from the water body polygons were generated. The buffer rings were created to the extent where all the residual patches would be covered. The number of buffer rings generated varied with the distribution of residual patches within the fire perimeter. The residual patch area was then computed for each buffer ring, and the spatial association with proximity to surface water would be assessed. The spatial association of residual patches with increasing distance inward from the fire perimeter was also investigated. A series of 100 m inward buffer rings were created from the edge of the fire perimeter; this was repeated until the area of the bands was 0 m^2. The number of inward buffers was also different for each fire event studied, with 2 buffer rings for the smaller events (F02, F03, and F09) and 17 buffer rings for one of the largest events (F06). Similarly, the area of the residual patches within each buffer was computed to assess the spatial relationship of residual patch occurrence relative to distance from the fire perimeter. The process of creating inward and outward buffer rings and computing residual patch areas within each ring was conducted for each fire event at the five spatial resolutions.

2.3. Results

The results are presented in four sections: patterns and characteristics of residual patches, effects of changing grain size, impact of land cover types on residual occurrence, and the spatial association of residual patches with natural firebreaks (surface water) and fire perimeter. In each section, the results of pattern analyses are given for each fire event observed.

2.3.1. Patterns of unburned residual patches

Within the eleven fire events studied in this dissertation, at 4 m spatial resolution, there were 1629 residual patches, with a mean patch area of 1.61 ha. The minimum and maximum patch areas were 0.25 ha (157 pixels) and 151.97 ha (94982 pixels), respectively. The total number of residual patches with the events was large (Table 2.4), but most of the vegetation patches (i.e., more than 75%) were less 1 ha. Those small residual patches (≤ 1 ha) were responsible for 21% of the total unburned area (Table 2.4). This indicates a degree of fragmentation of residual patches within a disturbed landscape.

Table 2.4. Unburned vegetation residual patches: the total number (and area) of residual patches, and proportion of residual patches < 1 ha across the 11 fire events.

	Total residual patches					**Residual patches < 1 ha**			
Event	Total no.	Total area	Min. area	Max. area	Mean area	Total no.	% of total no.	Patch area	% of patch area
F01	469	641.53	0.25	52.43	1.37	353	75.27	170.48	26.57
F02	9	3.56	0.25	3.56	0.40	9	100.00	3.56	100.00
F03	6	3.96	0.29	1.03	0.66	5	83.33	2.93	74.07
F04	102	103.70	0.25	10.83	0.45	81	79.41	36.24	34.95
F05	170	344.90	0.25	97.13	2.03	119	70.00	56.85	16.48
F06	327	787.97	0.25	151.97	2.41	233	71.25	102.61	13.02
F07	59	90.90	0.25	43.39	5.57	45	76.27	20.48	22.53
F08	136	361.03	0.25	107.00	2.23	124	91.18	54.76	15.17
F09	3	1.86	0.32	1.21	0.62	2	66.67	0.65	35.06
F10	308	379.13	0.25	16.92	1.23	223	72.40	107.79	28.43
F11	40	26.28	0.27	3.03	0.66	34	85.00	14.56	55.42
Total	**1629**	**2744.83**			**1.61**	**1228**	**75.38**	**570.92**	**20.80**

The characteristics of residual patterns at a given spatial resolution was explored to assess spatial heterogeneity and determine the range of variation in the amount, size, shape and spatial arrangement of residual patches within the fire perimeter; this has been accomplished by computing landscape metrics at 4 m spatial resolution (Table 2.5).

Table 2.5. Patterns of selected landscape metrics computed for the 11 fire events at 4 m spatial resolution, at the landscape level of analysis where aggregate characteristics of residual patches are examined.

Event	**CA**	**NP**	**PD**	**LPI**	**LSI**	**MPS**	**MSI**	**FRAC**
F01	641.53	385	60.01	8.30	130.04	1.67	5.91	1.37
F02	3.56	8	224.62	35.76	11.29	0.45	3.88	1.33
F03	3.96	6	151.45	25.93	8.53	0.66	3.52	1.29
F04	103.7	99	95.47	10.44	46.98	1.05	4.38	1.32
F05	344.9	159	46.10	29.78	54.39	2.17	4.30	1.30
F06	787.97	282	35.79	22.30	112.36	2.79	5.49	1.34
F07	90.9	57	62.70	47.73	31.50	1.59	4.08	1.30
F08	361.03	151	41.82	30.04	52.83	2.39	4.31	1.31
F09	1.86	3	161.50	64.94	10.59	0.62	5.72	1.39
F10	379.13	261	68.84	5.16	101.71	1.45	5.73	1.37
F11	26.28	38	144.61	11.53	22.62	0.69	3.55	1.28

2.3.2. Spatial pattern analyses: effect of changing grain size

The results of the selected spatial metrics are summarized in the form of landscape metric scalograms (Wu 2004), in which pattern indices are plotted against grain sizes; this is intended to evaluate the scale effect on characterizing spatial patterns. For each fire event, a subset of the pattern metrics was computed and a scalogram was generated for each metric. Since the computed values of the spatial metrics have different measurement units (and hence are not directly comparable to each other), all the values of the metrics were normalized by dividing each metric by its maximum value. In general, changing grain size for fine resolution data had substantial effects on the LPM values. While there was some consistency among some of the metrics' values, the magnitude and pattern of the response curves varied for most of the metrics across the fire event landscapes. The effects of grain size on spatial patterning can thus be grouped into three general categories: 1) monotonic and predictable response, 2) monotonic change with no simple scaling relationship, and 3) non-monotonic change with erratic responses.

2.3.2.1. Monotonic decreasing and predictable response

This includes the LPM that decreases with increasing grain size in a remarkably consistent power law relationship, with a coefficient of determination (R^2) > 96%. Three of the LPM examined exhibit such a predictable response across all the events considered (Figure 2.4). The three metrics are all related to shape, including LSI, MSI, and FRAC. These metrics changed predictably with increasing grain size, exhibiting scaling relationships that were consistent across the landscapes. This showed that shapes of the residual patches become less complex and irregular as data become increasingly spatially aggregated (or resampled). Figure 2.5 shows the variability of the LPM exhibiting monotonic decreasing and predictable response, across the eleven events with a power law best-fit model. The three shape-related metrics show a similar trend across all the events despite the differences in the spatial extent of the fire events, and the spatial composition, and configuration of the subsequent post-fire landscape structure.

2.3.2.2. Monotonic change with no simple scaling relationship

This encompasses spatial metrics that increase or decrease with increasing grain sizes, but do not exhibit a predictable response across all the events. One may expect that some measures of landscape structure such as CA, MPS, LPI, and MNN would increase monotonically with increasing grain size simply because of the progressive aggregation of smaller patches into large ones. For the same reason, other measures of landscape such as NP and PD would tend to decrease monotonically with increasing grain size. These have been evident for some of the metrics computed for some of the fire events; yet the increase or decrease in the parameters is not readily predictable across all the events. Figure 2.6 shows the LPM that decrease or increase with increasing grain size but did not show consistent responses among different fire event landscapes. For example, in F01, some LPM (e.g., NP and PD) exhibit monotonic change but the same LPM (NP and PD) show erratic behaviour for F06. Despite the monotonic change of the LPM, there is an unexpected deviation from a predictable pattern. However, a statistically significant regression coefficient between LPM and grain sizes was attained for some of the fire events, yet different mathematical models are generated for each LPM across different events (Table 2.6). Hence, it is not easy to fit a robust scaling law to predict patterns across all the fire events. This is reflected in Figure 2.7 and Figure 2.8 where the variability of the LPI values exhibiting a monotonic change with no simple scaling rule is presented. As a diagnostic, the results show that the metrics characterizing the pattern of residual patches vary across the events.

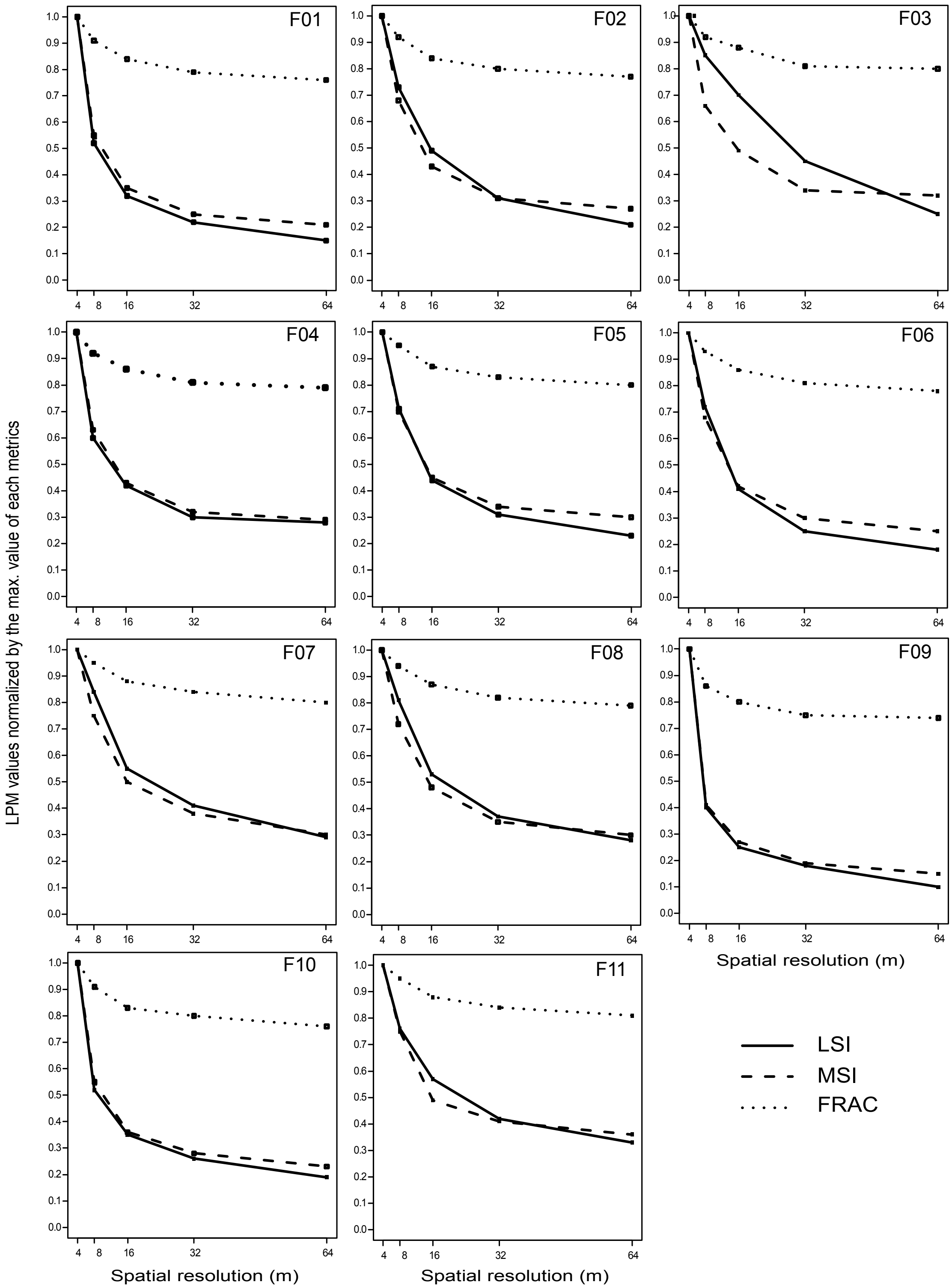


Figure 2.4. Scalograms showing effects of changing grain size on landscape: landscape metrics that exhibit monotonic decreasing function and predictable response with increasing grain sizes. The y-axis shows the LPM values normalized by max. value of each metrics.

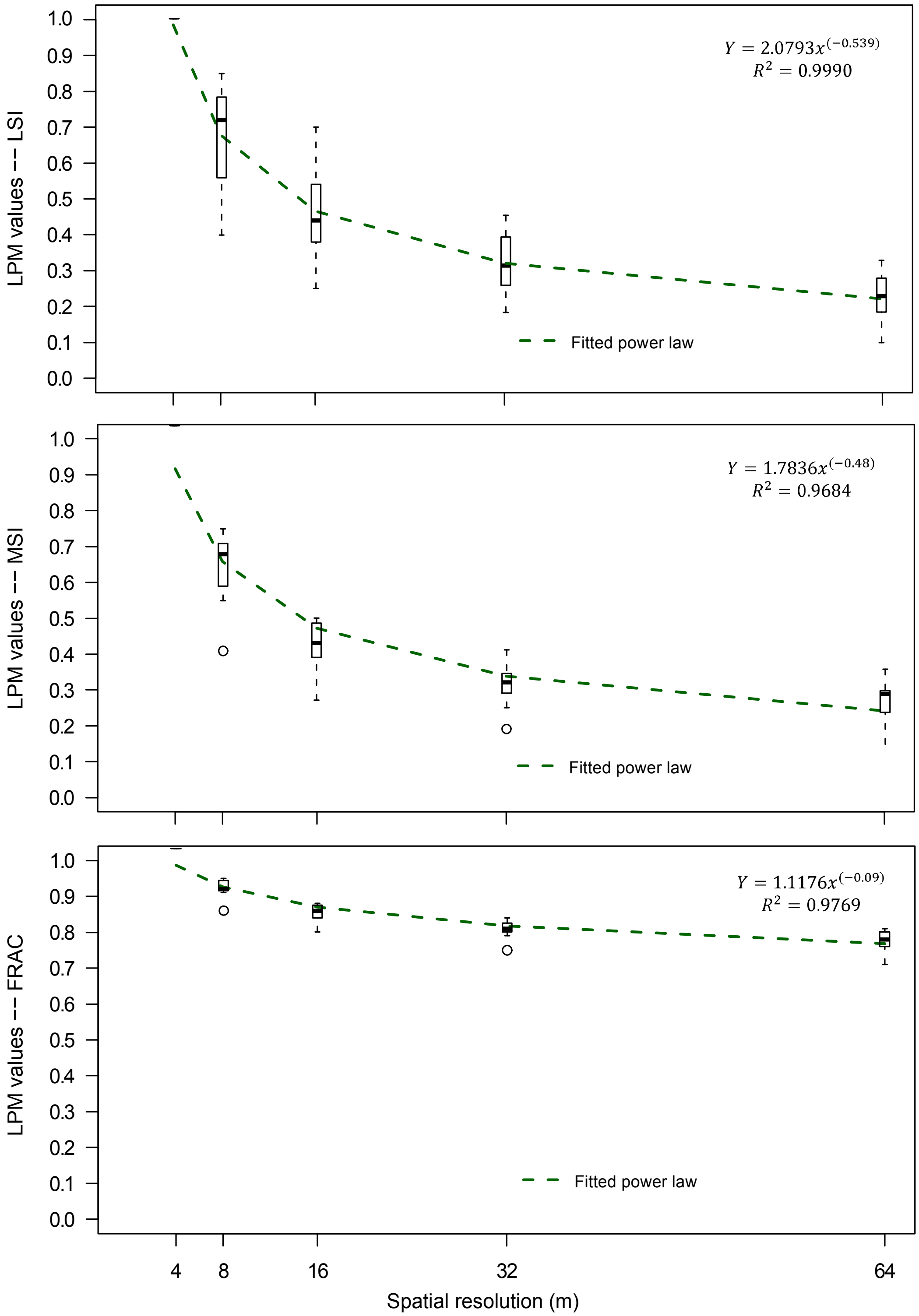


Figure 2.5. Scalograms showing effects of changing grain size: the variability of LPM values exhibiting a monotonic and predictable relationship across all the fire events. The box plots are based on the LPM values obtained from the 11 fire events.

Table 2.6. R^2 values showing the best fit models of the measured landscape metrics in relation to the 5 grain sizes used in the study. The best fit generates different regression types: Power law (black), Linear (red); Exponential (green), Logarithmic (blue), and shaded (LPM with erratic response curves).

	R^2 for LPI with practicable response curves across all fire events			R^2 for LPI with monotonic changes (decreasing or increasing) but no simple scaling relationship across different fire events							
Event	LSI	MSI	FRAC	CA	LPI	NP	PD	MPS	PSSD	PSCV	MNN
F01	0.9868	0.9587	0.9726			0.9886	0.9778	0.9831	0.8139		0.9886
F02	0.9965	0.9706	0.9695								
F03	0.9289	0.9586	0.9603								
F04	0.9444	0.9532	0.9687	0.9656			0.9840	0.9952	0.9917	0.9800	0.9659
F05	0.9940	0.9659	0.9787	0.9573				0.9902		0.8659	0.9987
F06	0.9916	0.9781	0.9845								0.9899
F07	0.9887	0.9917	0.9923		0.9205		0.9675			0.9550	0.9899
F08	0.9911	0.9804	0.9872				0.9497	0.9513	0.9679		
F09	0.9683	0.9323	0.9016								
F10	0.9694	0.9495	0.9650				0.9923	0.9844	0.9793	0.9268	0.9735
F11	0.9989	0.9577	0.9847	0.9741			0.9424	0.9535			

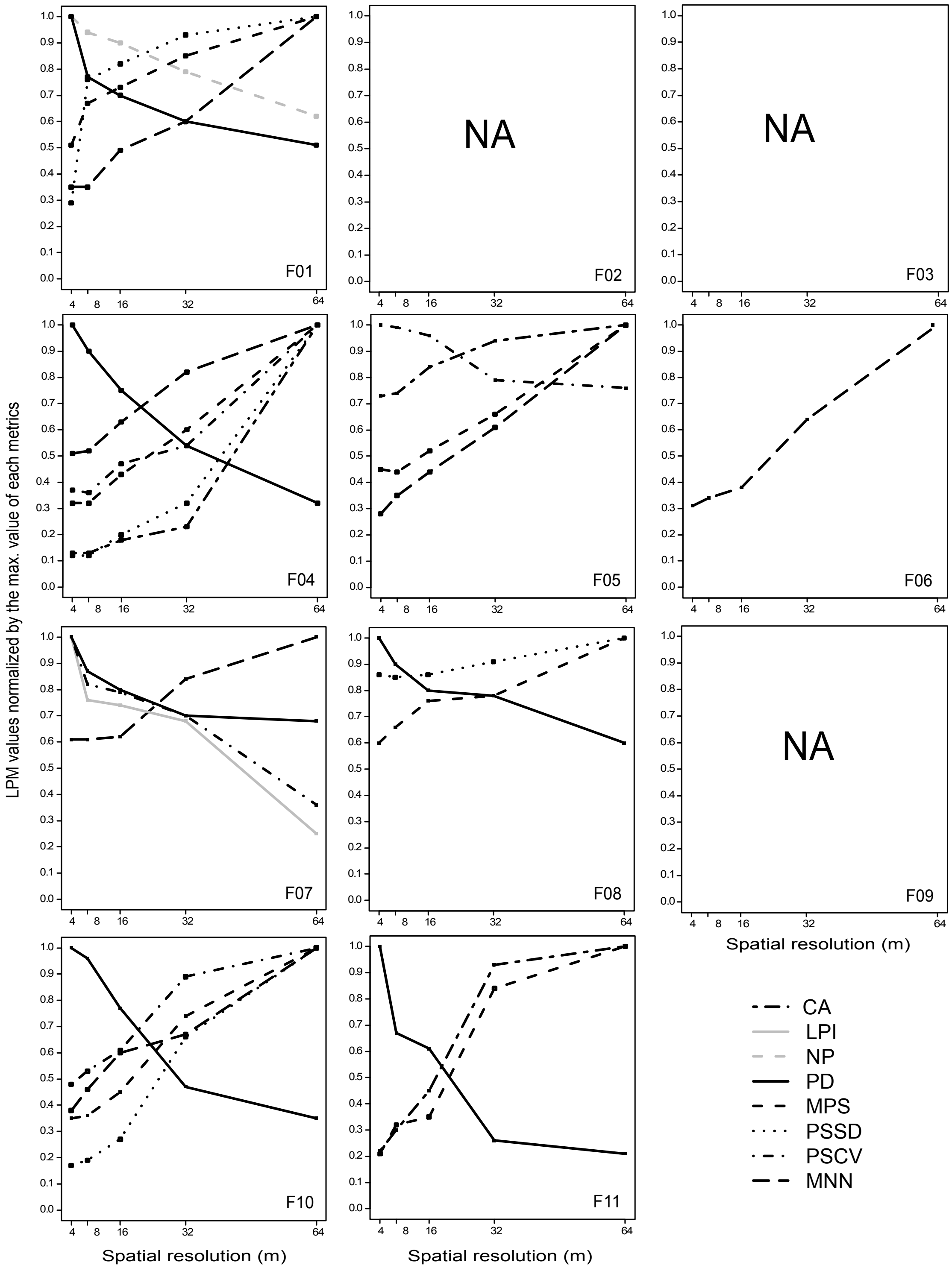


Figure 2.6. Scalograms showing effects of changing grain size on landscape: landscape metrics showing monotonic change (decreasing or increasing) with no simple scaling relationship across the 11 fire events (NA indicates none of the metrics in the event has this response curve). The y-axis shows the LPM values normalized by max. value of each metrics.

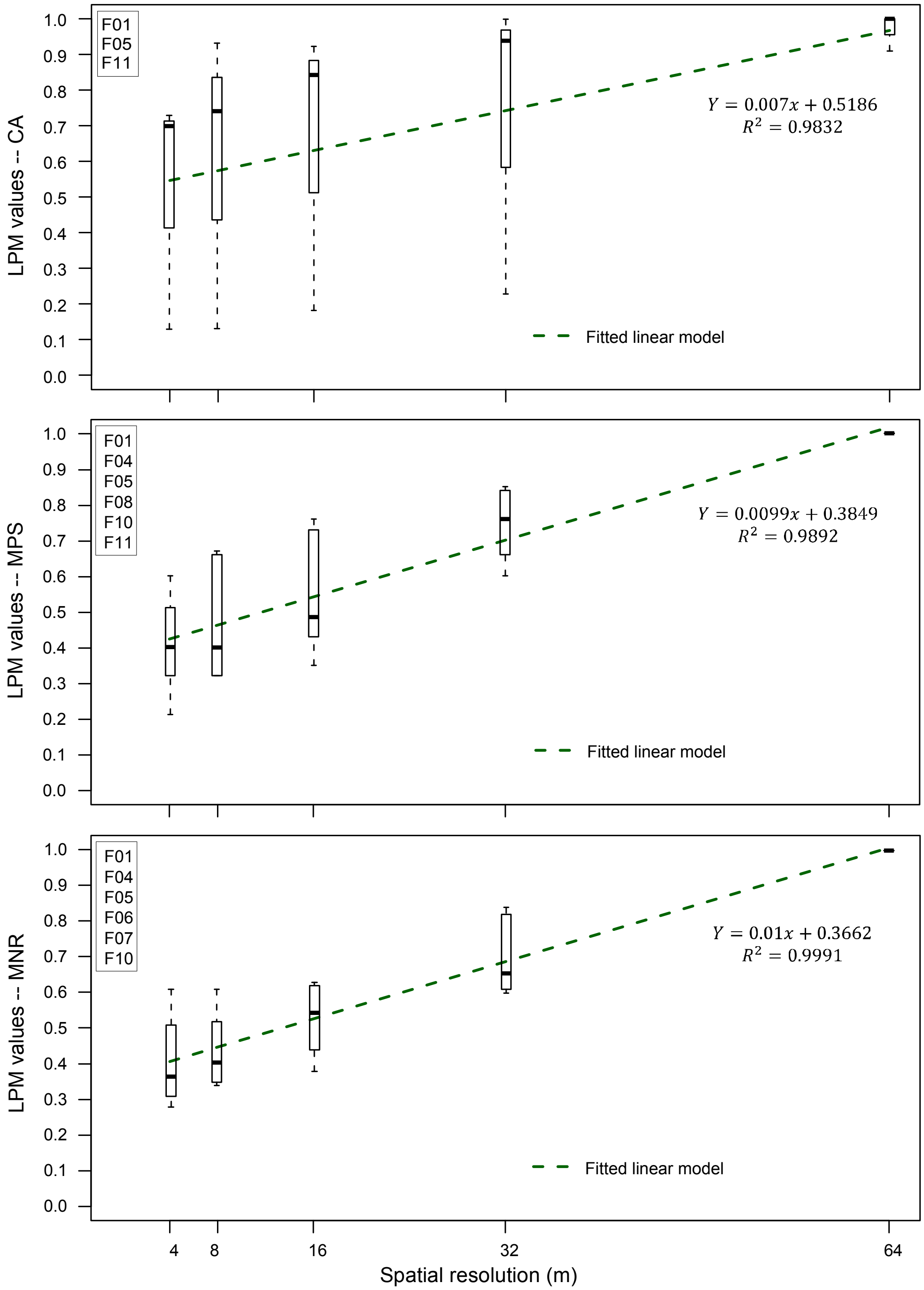


Figure 2.7. Scalograms showing effects of changing grain size: the variability of LPM values exhibiting a monotonic change across certain fire events but with no robust scaling rule across the 11 fire events. The box plots are based on the LPM values obtained from the fire events listed in the top-left corner of each scalogram.

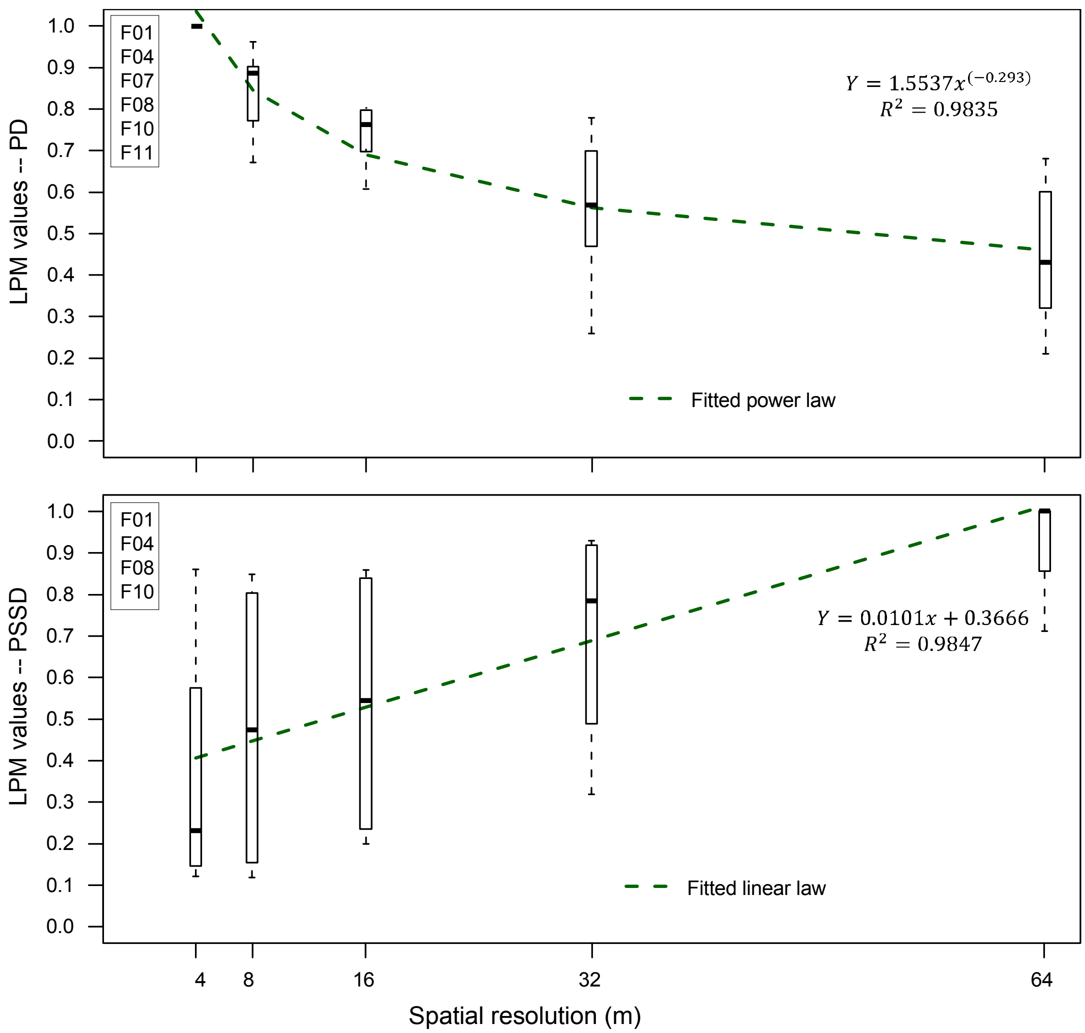


Figure 2.8. Scalograms showing effects of changing grain size: the variability of LPM values exhibiting a monotonic change across certain fire events but with no robust scaling rule across the 11 fire events. The box plots are based on the LPM values obtained from the fire events listed in the top-left corner of each scalogram.

2.3.2.3. Non-monotonic and erratic response

Figure 2.9 shows the LPM that exhibit non-monotonic change and erratic responses among different events, particularly for the small sized events (F02, F03, and F09). For these three events, all the LPM (except the three shape related metrics: LSI, MSI, and FRAC) exhibited erratic responses to increasing grain size. On the other hand LPM such as CA, LPI and NP tended to be erratic across most of the landscapes but LPM that show erratic behaviour were not similar among different landscapes, as shown in Figure 2.9. The non-monotonic changes with increasing grain sizes demonstrate the unpredictable patterns of the LPI across the landscapes. This indicates that it is not easy to develop a simple scaling law to predict patterns for these specific metrics at different scales.

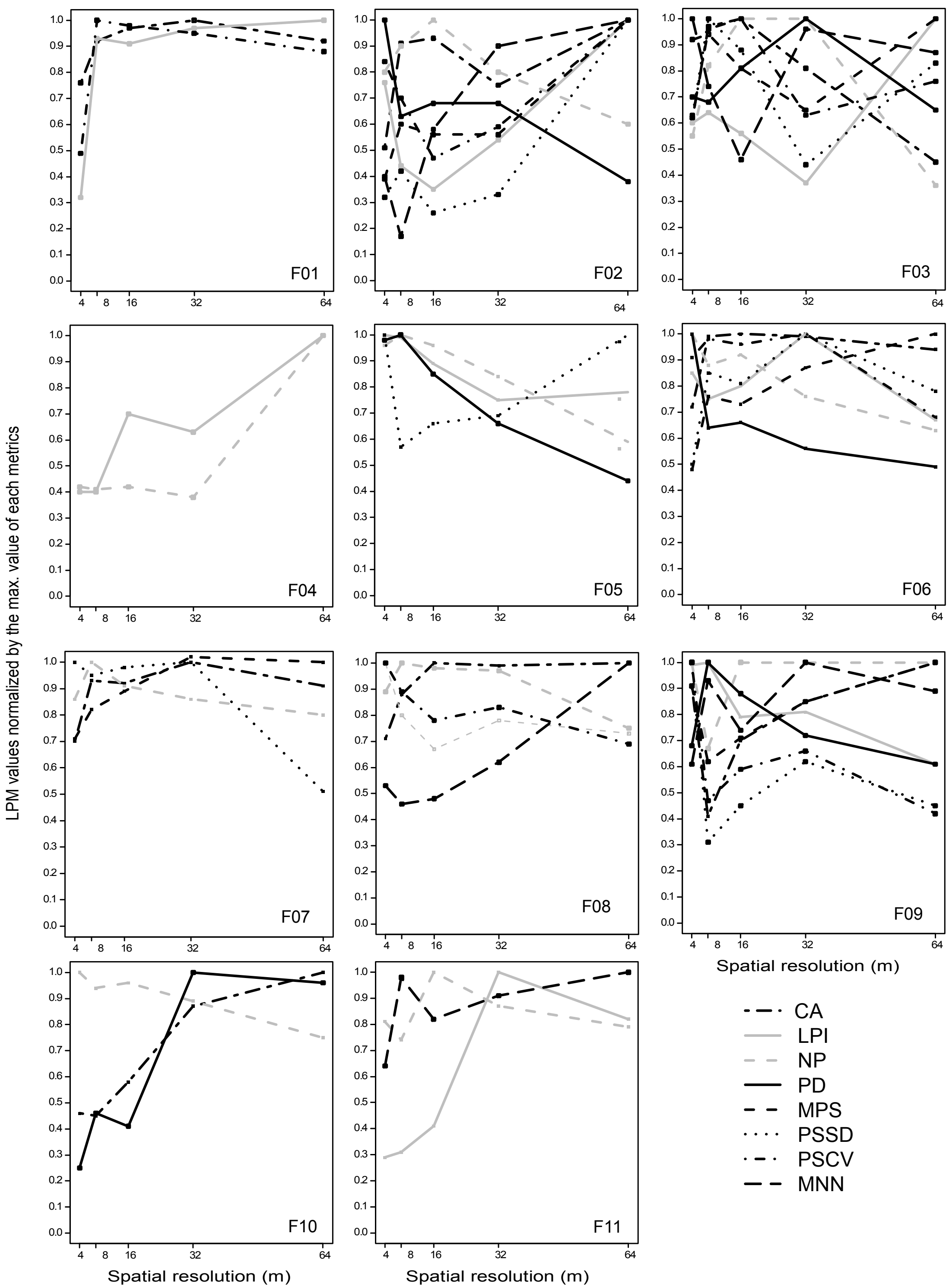


Figure 2.9. Scalograms showing effects of changing grain size on landscape: landscape metrics showing erratic response curves and non-predictable pattern across the certain fire events (NA indicates none of the metrics in the event has this response curve). The y-axis shows the LPM values normalized by max value of each metrics.

2.3.3. Residual patch distribution: impact of land cover type

The post-burn land cover composition within the unburned patches was examined for the 11 boreal wildfires using selected class-level metrics (e.g., CA and %LAND). This was to address whether particular land cover types were more likely to generate residual patches than other types. These two metrics explain the size of residual patches occupied by specific cover types. The results show that the proportion of residual patches occupied by different cover types tend to vary across the events (Figure 2.10), depending on the dominance of different land cover types in the landscape and resistance of certain cover types to fire. However, sparse conifer was over-represented within residual patches across all fire events except for F01, F03 and F10. In these three events, treed wetland (in F01), dense conifer (F03), and low shrub (F10) classes were more likely to escape burning. Specifically in F01, more than 50% of the residual patches were occupied by treed wetland while deciduous trees occupy more than 40% of the residual patches in F03. However, majority of the residual patches were covered with sparse confer; there was a situation in which approximately 50% of the residual patches comprised of sparse conifer (e.g., F04, F05, and F11).

The fragmentation level of residual patches in relation to land cover types was also examined using selected metrics (NP and LI), which can explain the landscape heterogeneity. Figure 2.11 shows the variability in NP across 5 spatial resolutions by land cover types; the result shows that the number of residual patches occupied by different cover types is substantial and highly variable across the five spatial resolutions. Similarly, the variability in the largest patch index (LPI) across the five spatial resolutions by land cover is estimated and shown in (Figure 2.12). Despite the variability in LPI values, it is investigated that the LPI is likely to be associated with the land cover types that dominate the residual patches (higher %LAND values). In F01, for example, the largest patch is occupied by treed wetland, which was predominant land cover in residual patches (i.e., more than 50%).

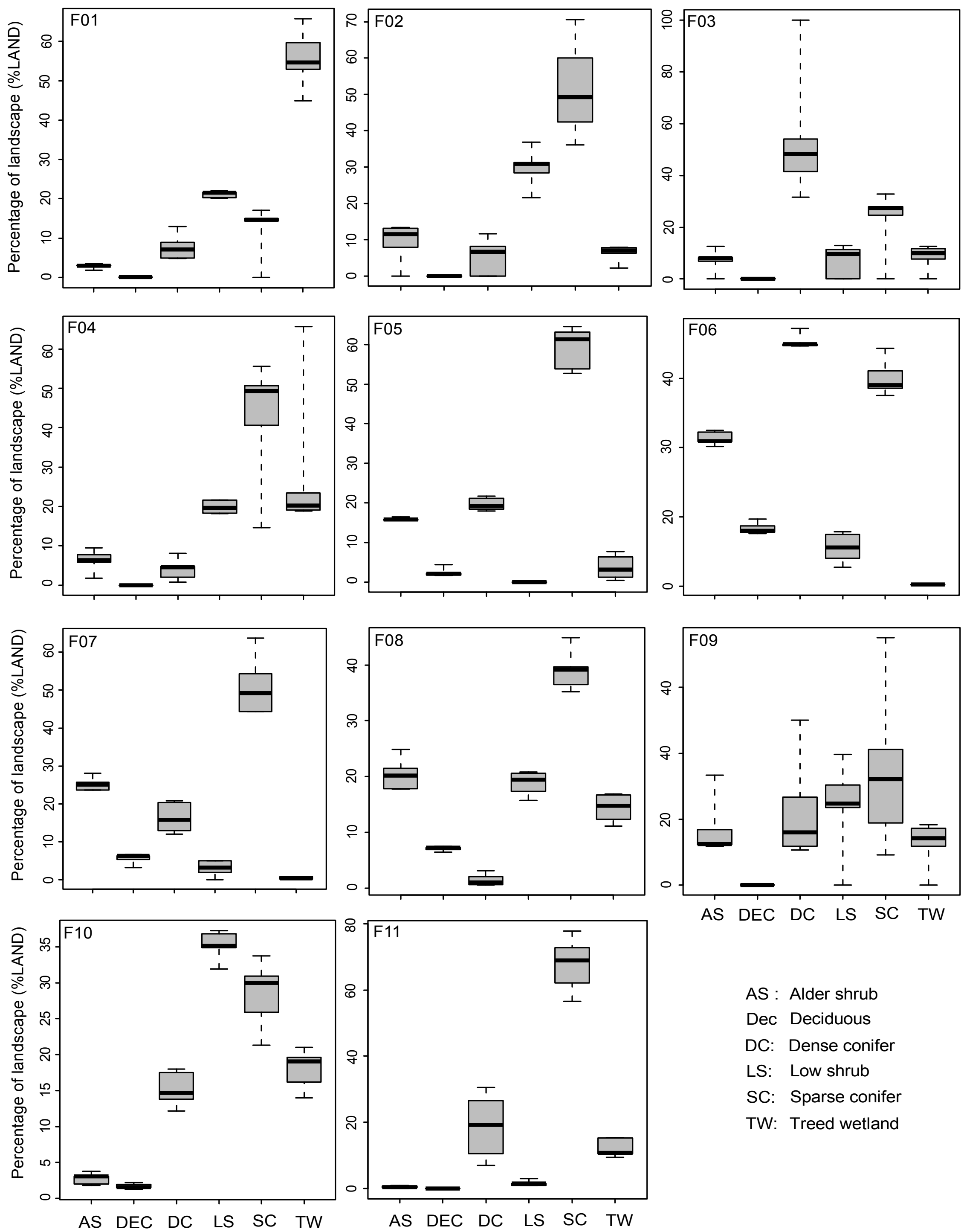


Figure 2.10. Land cover composition of residual patches: the variability in the proportion of fire footprint occupied by existing residual patches; the variability in %LAND across the five spatial resolutions by land cover types. Each box in the plot is based on the residual patch area computed across the five spatial resolutions by land cover types.

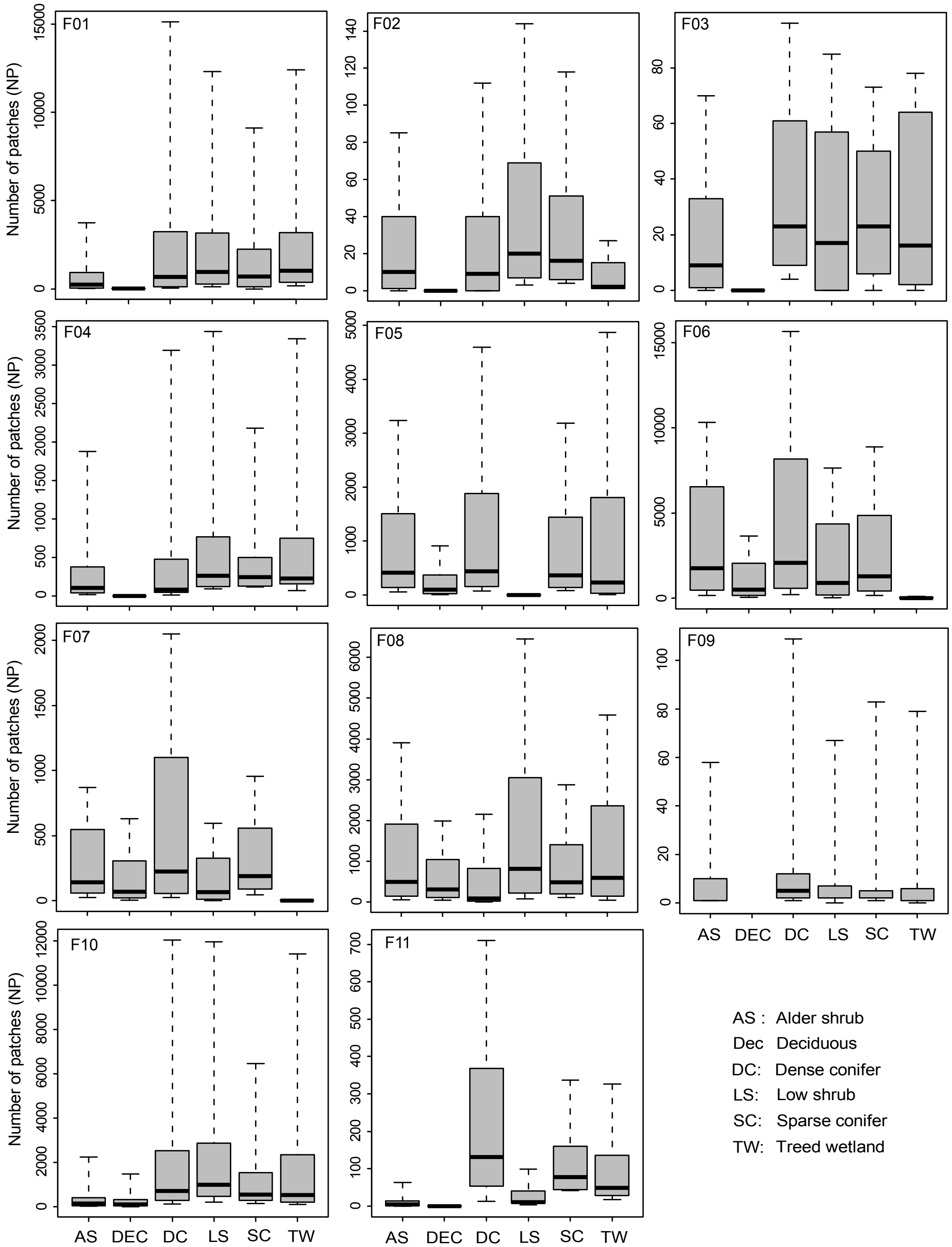


Figure 2.11. Residual patch fragmentation: the variability in the number of residual patches occupied by different cover types; each box in the plot is based on the metric values obtained at five spatial resolutions.

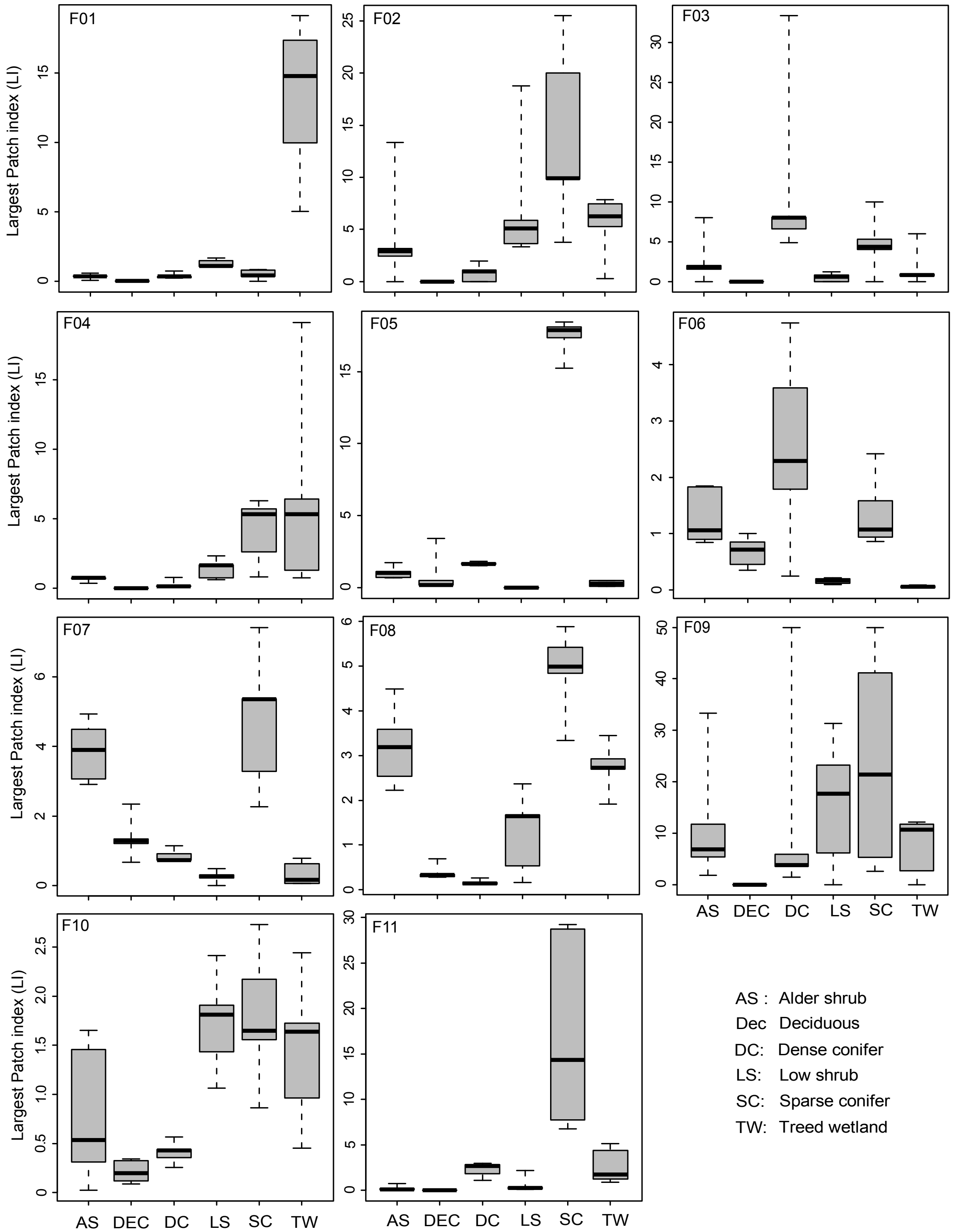


Figure 2.12. Residual patch composition: the variability in LPI across the five spatial resolutions by land cover types; each box in the plot is based on the metric values obtained at five spatial resolutions.

2.3.4. Spatial association with surface water and fire footprints

The adjacency analysis of residual patches with surface water and edge of fire perimeter was investigated at the five spatial resolutions. Figure 2.13 and Figure 2.14 show the variability of residual patch area, and the distance from surface water and edge of fire perimeter among the fire footprints observed at R_4 and R_{32} spatial resolutions. The results of the spatial associations at R_4 and R_{32} m spatial resolutions are presented because these resolutions are similar (or closest) to how a respective remote sensing device would view the landscape with Ikonos and Landsat images respectively. Despite the importance of surface water for the occurrence of residual patches, the variability of residual patch area (shown in Figure 2.13) suggested that it is difficult to generalize any kind of trends with increasing distance from natural firebreaks, across all the five spatial resolutions. However, the fitted models tended to indicate that the residual patch area decreases with increasing distance from surface water and fire perimeter, with a second-order polynomial model form.

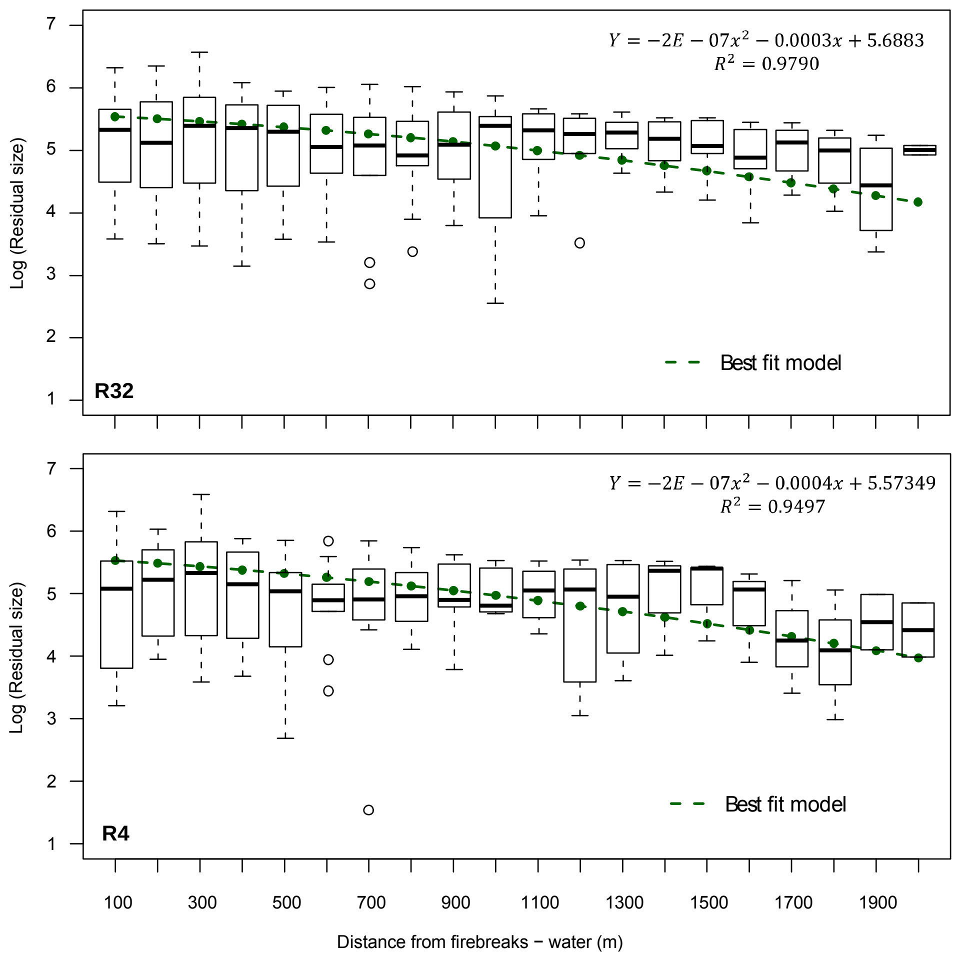


Figure 2.13. Variability in the proportionate extent of residual patches in external 100 m wide buffer rings with increasing distance from natural firebreak features (i.e., water); each box in the plot is based on the residual patch area computed across 11 fire events (at R_4 and R_{32}). The y-axis shows logarithm of residual patch area.

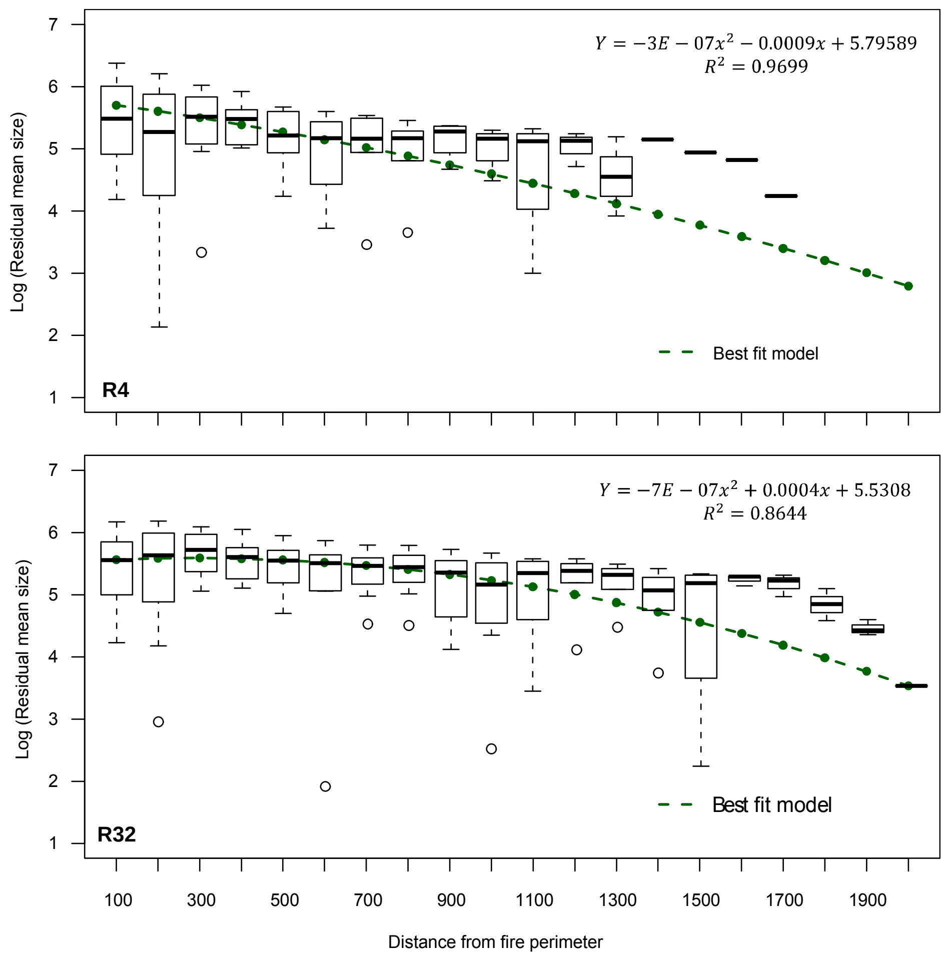


Figure 2.14. Variability in the proportionate extent of residual patches in internal 100 m wide buffer rings with increasing distance from footprint perimeters; each box in the plot is based on the residual patch area computed across 11 fire events (at R_4 and R_{32}). The y-axis shows logarithm of residual patch area.

2.4. Discussion

2.4.1. Spatial pattern analysis of residual patches

One of the most important questions asked by forest managers is how much residual patch must be preserved during harvesting operations. This requires knowledge about the proportion of unburned areas within a fire-disturbed landscape (Andison 2004). The proportion of the unburned areas in this study represent between 15% and 44% of the total fire area for fire events F11 and F06, respectively, with an average 28% of the total fire area; this includes not only residual patches but all unburned areas (e.g., surface water, and bedrock and non-vegetated areas). The residual patches account between 3% for F09 and 21% for F06, with an average 10% of the total area across all the fire events; this suggests the severity of the fire varies across the fire events. This is similar with the report made by Madoui et al. (2010), who found that the average proportion of residual patches in the Western Quebec was only 10% of the total area. The study was based on 33 fire events, obtained from classified Landsat imagery and having fire areas ranging from 3114 to 51882 ha; located in two different physiographic zones. The results of my study also suggested that the composition of residual patches in northwestern Ontario is comparable with the study undertaken in northern BC (Delong and Tanner 1996) and foothills of Alberta (Andison 2004). The studies indicated that residual patches accounted for an average of 3-9% in sub-boreal BC and 10-11% in the Alberta's foothills respectively.

The proportion of land area (CA) escaped from fire varies considerably across the fire events, where there was a situation in which more 21% of land area evades the fire. The proportion of the fire footprint occupied by residual patches was relatively substantial for the large sized fire events (e.g., F01, F06, and F08) while it was comparatively small for the small sized fire events (F02, F03, and F09), with a significant correlation coefficient ($r = 0.938$, $p < 0.001$) between the proportional area of residual patches and disturbance size. This is likely to happen because large fires often encounter natural firebreaks such as wetlands, surface water, and barren lands (Eberhart and Woodard 1987); hence result in a substantial area and number of residual patches. In small fire events, depending on fire intensity and natural firebreaks, there is a tendency for a fire to burn everything. Such a positive correlation was also obtained in a study undertaken by Delong and Tanner (1996) in the sub-boreal of BC. The study on the foothills of Alberta however suggested the independence of proportional residual area from the disturbance size.

The study on the patterns of residual patches in BC and in the Albert's foothills, and the results of my study indicated that the composition of residual patches varied across the fire events from different studies. The variation can be explained in relation to fire intensity and severity, but the differences in the patterns can also be attributed to differences in the spatial resolution of the source data used to study the patterns and residual patch definition. The latter is related to the way in which the 'residual patches' are defined with respect to survival levels of species; the incorporation of partially disturbed islands in the definition. It also refers to the differences in the minimum resolution (i.e., the minimum island polygon used to define a residual patch). This study was based on 0.25 ha as the minimum island area to define a residual patch while other studies based their definition using 0.20 ha (Delong and Tanner 1996) and 0.02 ha (Andison 2004) as a threshold value. This suggests that understanding the natural patterns of residual patches should require a clear definition of residual patch existence and resolution at which the minimum residual island is delineated; yet this is a goal-specific assessment.

The other question that might be asked by forest managers is related to the heterogeneity level within a disturbed landscape as this reflect the abundance of natural firebreak features and variation in fuel distribution. The fire events were characterized by a large number (>75%) of small residuals (< 1 ha); indicating a high degree of fragmentation of residual patches. The existence of large numbers of small residual patches also reflects the abundance of high concentration of natural firebreaks within the landscapes or diversity within the fire event. Moreover, the heterogeneity of a fire-disturbed landscape can be explained by various metrics. LPM such as NP, PD, LPI, and MPS, for example, can serve as general indices of spatial heterogeneity (fragmentation) of a landscape, but they are not spatially explicit measures (McGarigal et al. 2002). Specifically, NP and PD of a particular habitat type may affect a variety of natural processes; for example it may alter the stability of species interactions or opportunities for coexisting of different species in a landscape and affect the propagation of disturbances (e.g., fire) across a landscape (McGarigal and Marks 1995). In this study, the NP was relatively high for the large sized fire events (e.g., F01 and F06) and it was lower for small sized fire events (F09); indicating a significant correlation between NP and fire footprint area (r = 0.973, p < 0.001) and higher degree of fragmentation for large sized events. This could be attributed to the diversity of land cover types (fuel distribution) and abundance of natural firebreaks within the fire events. LPI is also another indicator of spatial heterogeneity, where the least heterogeneous landscape has higher LPI values. The LPI values shown in Table 2.4 indicate that two of the large sized fire events (F01 and F10) were the most heterogeneous landscapes with LPI values of 8.30 and 5.16 respectively; yet LPI values is not significantly correlated with disturbance size (r = -0.555, p = 0.0762). During fire recurrence, the highly fragmented residual patches may be more resistant to the propagation of fire disturbances; hence more likely to persist in a landscape than patches that are contiguous (McGarigal and Marks 1995).

PD is believed to facilitate comparison among landscapes of various sizes; where a landscape with a greater PD would have more spatial heterogeneity. PD was negatively correlated with the extent of fire disturbance (r = -0.80, p = 0.0080); suggesting that small sized fire events tend to exhibit greater heterogeneity. Although it is scale dependent, a smaller mean patch size indicates a spatially heterogeneous landscape. It is discovered that the MPS was positively correlated with the area of the fire event (r = 0.78, p = 0.0041), indicating that small sized events might be more fragmented than larger ones. The relationship between different metrics (LPI, PD, and MPS) and fire disturbance size provided different explanation on the fragmentation levels and fire disturbance size. These prompt one to suggest that the spatial heterogeneity of the fire events might not easily be explained with a combination of LPM parameters.

Furthermore, residual patches maintain different natural shapes or variations that eventually affect the calculation of shape level metrics (Kachmar and Sanchez-Azofeifa 2003). Residual shapes can be ‘circular’, isolated, and identifiable homogenous areas surrounded by burned areas or homogenous areas that follow the perimeter of a linear natural (e.g., river stream) or anthropogenic features (e.g., roads or transmission lines). With respect to shape metrics, a relatively circular form of residual patches (MSI = 1.5 – 1.6) was observed in a study undertaken by (Drogatescu and Kneeshaw 2012). In a study by Andison (2004), an MSI of 1.3 and 2.9 was also obtained depending on the size of the fire. In this study, an MSI value of greater than 3.55 was computed across all the fire events; indicating that the patch shapes are non-circular. Similarly, FRAC values greater than 1.28 was observed for all fire events, confirming the departure of the residual patches from a Euclidean geometry. The non-linear shape of the residual patches provides more suitable interior habitat conditions (Bergeron et al. 2007). The shape of the residual patches were irregular and complex, and might have been caused by irregular shapes of the water courses, wetlands, and other natural variations within the fire perimeter.

2.4.2. Sensitivity to scale change

The effects of grain size on characterizing the patterns of landscape structure have been reported in different studies (Turner 1989; Wiens 1989; Benson and MacKenzie 1995; Wu et al. 2002; Zhu et al. 2006; Haire and McGarigal 2009), but the effectiveness of the methods remains questionable (Wu et al. 2000). In this chapter, a single-scale method was used to assess the impacts of grain sizes on characterizing post-fire landscape structure. Changing grain size can have considerable (quantitative and qualitative) effects on how LPM changes over multiple scales (Turner1989; Zhu et al. 2006; Corry and Lafortezza 2007). The results of the multiscale analyses supported the hypothesis that the LPM used in this study would be sensitive to changing grain sizes, across the events. Although the sensitivity to scale change varied greatly among the indices and spatially across the sites, the effects of grain size on measures of LPM were grouped into three categories: monotonic change and predictable patterns, monotonic change (decreasing or increasing) with no simple scaling relationship, and non-monotonic change with erratic responses.

Previous studies indicated that landscape shape index decreases following a power law relation as grain size increases (Wu et al. 2000; Corry and Lafortezza 2007). This was evident in this study where both LSI and MSI exhibited consistent and predictable patterns, with a power law relation throughout the fire events. Hence, simple mathematical regression equations can be formulated to predict patterns at coarser spatial resolutions. With respect to the fractal dimension metrics, Wu et al. (2002) found that the behaviour of landscape fractal dimension in response to changing grain size did not have a predictable trend; this was based on remote sensing data of a boreal forest landscape (near Thompson, MB, Canada). In a similar study by Wu et al. (2000), the fractal dimension of a landscape remains constant over a range of grain sizes and begins to fluctuate after grain sizes exceeds 50 m. However, my study showed that the fractal dimension tended to behave the same way as the other shape related metrics (i.e., a decreasing power relation) (Figure 2.4). These indicate that new and predictable landscape features would emerge over coarser grain sizes. The consistent patterns of the shape related metrics is likely because there are limited values for shape related parameters (Griffith 2004).

Other measures of landscape structure such as CA, MPS, LI, MNN, NP, and PD are expected to increase or decrease monotonically with increasing grain size simply because of the progressive aggregation of smaller patches. In a study conducted by Wu et al. (2000), the two measures of spatial heterogeneity: NP and PD showed a remarkably consistent power-law relationship, suggesting that these indices can be predicted over a wide range of grain sizes. Corry and Lafortezza (2007) also found that CA had a predictable pattern while NP and MPS showed a stepped-function (i.e., staircase-like response). In this study, the magnitude and pattern of the response curves of these measures of LPM vary among the events, with CA, MPS, MNN, NP, and PD exhibiting monotonic decreasing behaviour with no simple scaling law for five of the fire events. The response curves of these parameters for the rest of the events were erratic. Similarly, PSSD increases almost linearly with increasing grain size while PSCV decreases in a power-law fashion (Wu et al. 2000). The PSSD and PSCV in this study showed a non-linear increase and decline with increasing grain size, respectively, for some of the fire events while they showed an erratic response for the certain events. The differences in the magnitude of LPM response curve across the events do not allow one to develop scaling laws or predict patterns at coarser spatial resolutions. The difference in magnitude of response curves also suggests that large number of samples (i.e., a wide range of grain sizes) would generate better scaling laws for transferring information more effectively.

In order to understand the role of scale in spatial heterogeneity, it is useful to identify simple relationships between metrics measured at different scales and examine the scaling laws exist for different patterns (Wu et al. 2000). However, the exact relationship among metrics measured at multiple scales varies across landscapes and may not allow information extrapolation (Turner 1989). The result of this study supported this view that the response curves from most of the LPM, except LSI, MSI and FRAC, did not allow one to derive robust scaling laws to predict patterns at coarser spatial resolutions across different events. This suggests that multi-scale analysis of residual patches at limited grain sizes (R_4, R_8, R_{16}, R_{32}, and R_{64}) would remain a challenge in landscape ecology. Furthermore, the discrepancies of the estimated parameters can also be attributed to the way in which grain sizes are aggregated. Studies have shown that different aggregation methods may have considerable effect on landscape patterns (Wichham et al. 1995; He et al. 2002). For example, for landscapes with greater local heterogeneity, the results of aggregation might be less adequate (Benson and MacKenzie 1995). On the other hand, important properties of landscapes are not preserved under some aggregation schemes (Benson and MacKenzie 1995). He et al. (2002) also indicated that majority and random-rule based methods of spatial aggregation led to different results in cover type proportions and altered spatial patterns. Therefore, it would be an interesting task to assess the impact of spatial aggregation methods that affect the characteristics of landscape metrics (Wu et al. 2002). While the single-scale method was useful to characterize and understand the patterns of post-fire landscape structure, having large number of resampled data (i.e., grain sizes) would provide a better approach to study multi-scale structural problems in landscape pattern over multiple scales.

2.4.3. Residual patch occurrence and land cover composition

Although there is directional effects related to wind and fire growth (Burton et al. 2008), it is assumed that the area within the fire perimeter had an equal chance of burning. Yet, there are certain areas that escape burning and the occurrence of those unburned areas is attributed to various factors, including abundance of natural firebreaks, less fuel availability, or the type of land cover or species type within the fire perimeter. The most obvious question is that why there is a tendency for certain land cover types burn preferentially than others. The abundance of residual patches is often associated with the dominant cover types in the landscape. However, there are some land cover types that dominate residual patch occurrence despite their low abundance in the landscape (Madoui et al. 2010). For instance, in a study by Kafka et al. (2001), deciduous forest cover was more likely to dominate the existing residual patches regardless of its low dominance in the landscape. Burton et al. (2008) likewise concluded that compared with the coniferous forests, deciduous forest is less likely to burn in the boreal mixed wood region and dominate the post-fire residual patches.

The findings of my study support this view to a certain degree that some land cover types are more likely to evade burning than other land cover types, despite their low abundance. In F01, for example, more than 71% the landscape is dominated by sparse conifer and dense conifer, but treed wetland and low shrubs are predominant in residual patches (Table 2.7).

Because of the wetness of the land, treed wetlands usually serve as a natural firebreak; hence it was expected that residual patches would be dominated by treed wetland. In different fire events, (e.g., F07 and F09) sparse conifer dominated the residual patches despite its low abundance (e.g., dense conifer is the dominant cover type). This indicated that sparse conifer has a low tree density and may not be prone to high-intensity fires because of a lack of fuel (Madoui et al. 2009). On the other hand, there was a situation in which land cover types with high abundance was positively associated with residual patch occurrence. This provides an overview of the land cover composition of the existing residual patches; the relative importance of land cover type as a variable to explain residual patches is presented in Chapter three.

Table 2.7. A table summarizes the proportion of the dominant land cover types in the fire footprints and the land cover types that dominate the existing residual patches in each fire event.

Fire event	**The most dominant land cover type (% of the landscape)**	**The second most dominant land cover type (% of the landscape)**	**Land cover type that dominate residual patches ***	**Land cover type associated with the LPI ****
F01	Sparse conifer (41%)	Dense conifer (30%)	Treed wetland	Treed wetland
F02	Sparse conifer (66%)	Treed wetland (31%)	Sparse conifer	Sparse conifer
F03	Treed wetland (57%)	Sparse conifer (42%)	Dense conifer	Dense conifer
F04	Sparse conifer (41%)	Dense conifer (21%)	Sparse conifer	Sparse conifer
F05	Water (38%)	Dense conifer (20%) and Old burn (17%)	Sparse conifer	Sparse conifer
F06	Dense conifer (86%)	Sparse conifer (8%)	Dense conifer	Dense conifer
F07	Dense conifer (76%)	Sparse conifer (12 %)	Sparse conifer	Sparse conifer
F08	Sparse conifer (41%)	Dense conifer (34.5%)	Sparse conifer	Sparse conifer
F09	Dense conifer (67%)	Sparse conifer (30%)	Sparse conifer	Sparse conifer
F10	Old burn (66%)	Sparse conifer (13%)	Low shrub	Low shrub
F11	Sparse conifer (44%)	Dense conifer (37%)	Sparse conifer	Sparse conifer

* Summary of Figure 2.10 and ** summary of Figure 2.12.

The spatial heterogeneity of residual patch occurrence in relation to land cover types using two measures of spatial heterogeneity: NP and LI was also investigated. The number of residual patches occupied by the six different cover types (alder shrub, deciduous, dense conifer, low shrub, sparse conifer and treed wetland) was substantial and highly variable across the five spatial resolutions considered. The large number of residual patches throughout the land cover types: 1) signifies that the residual patches are spatially distributed in the landscape due to high concentration of natural firebreaks in the landscape, and 2) indicates that the land cover types that constitute the residual patches are highly fragmented across space; meaning that a single residual patch is likely to be occupied by more than one cover type. Moreover, the variability in LPI shown in Figure 2.12 indicates that the proportion of the landscape area occupied by the largest residual patch is associated with the dominant cover type in the landscape (Figure 2.10, Figure 2.12, and Table 2.7). Based on this, it is sensible to argue that the dominant land cover types are associated with low level of spatial heterogeneity. Finally, the variability in NP and LPI across five spatial resolutions (Figure 2.11 and Figure 2.12) revealed that grain sizes had a substantial effect on computing the LPM over multiple scales.

2.4.4. Spatial association with natural firebreak features and fire perimeter

I discovered that certain land cover types are more likely to evade fire and hence form residual patches, but the concentration and spatial distribution of the residual patches are also associated with other factors (e.g., weather conditions, topography and fire breaks). It was reported that the spatial distribution of residual patches within the fire perimeter would be random under the influence of weather conditions (Madoui et al. 2010). The influence of weather conditions during a fire event could be substantial but the conditions are less likely to vary within a fire event scale. It was anticipated that there is a tendency for residual patches to be associated with proximity to natural firebreaks, mainly to surface water. The study showed that the residual patches are associated with the proximity to surface water to a lesser degree, specifically with increasing proximity to water. Despite the variation in residual patch area (Figure 2.13), the abundance of residual patches showed a trend with increasing distance from surface water, across all the five spatial resolutions. As a diagnostic, the relative abundance and distribution of residual patches within the fire perimeter, in relation to distance to surface water would mostly be uneven. Similarly, the spatial distribution (and variability) of residual patch area, across the five spatial resolutions showed some trend with closer proximity to fire perimeter edge; for example, large residuals tend to concentrate with closer proximity to the fire edge. The overall variability of residual patch area and their spatial association with distance from water or fire edge suggested that 1) the distribution (and impact) of surface water across the events was uneven and 2) the spatial distribution of residual patches within fires would be attributed to other geo-environmental factors. Therefore, the next section (Chapter 3) investigates the parameters that govern the occurrence of residual patches for the same landscape (11 fire events) considered in the study.

2.5. Summary and conclusions

Wildfire in boreal forests is usually intense and consumes a substantial amount of forest cover, but does not burn the entire landscape; it retains certain undisturbed vegetation patches within a fire-disturbed landscapes. Exploring the patterns, characteristics, and distribution of the heterogeneous landscape mosaic created by wildfire has been fundamental in landscape ecology. Also, improved understanding of the characteristics of post-fire residual structure in natural environment allows forest managers to implement effective forest management practices. This requires a comprehensive assessment of the characteristics of post-fire residual patches. To this end, my study sought to implement a reliable and consistent method for measuring and assessing the spatial patterns and characteristics of residual vegetation patches. In order to examine the characteristics of the residual patches, a number of related but distinctive research objectives were formulated.

The main objective of this chapter was to characterize the spatial patterns of post-fire residual patches using different metrics and examine the sensitivity of the metrics to changing grain sizes. The study involved different metrics that explain the composition, configuration, and the fragmentation of residual patches within the fire footprint. Burton et al. (2008) stated that the severity characteristics of wildfires, particularly large fires, vary across boreal North America due to the variation in climatic, vegetation, and fuel availability (Burton et al. 2008). This view was reflected in my study where the proportion of residual patches (CA and %LAND), which may indicate the severity level of a fire, varied across the fire events owing to the variation in fire behaviour (size, intensity, and severity). The uneven burn severity within the fire events can also be explained by the local effects of the sites (topography, natural firebreak features, and other non-vegetated features) and stand attributes of severity. This further intensities the diversity in the distribution and configuration of post-fire landscape structure.

The metrics related to the composition of residual patches such as CA and %LAND are also useful to determine the proportion of post-harvest residual patches that should retained during harvesting. The results of my study showed that the proportion of the residual patches varied from 3% for F09 (the smallest fire event with 58 ha) to 21% for F06 (the largest fire events with 3741 ha); with average 10% across the 11 fire events where the average fire footprint area was 1850 ha. The metrics that describe the spatial configuration and fragmentation of the residual patches such as NP, PD, LPI, and MPS can also be used to determine the spatial arrangements of the post-harvest residuals. Although the spatial heterogeneity might not easily be explained with a single LPM, this study found that there was high degree of fragmentation due to the abundance of natural firebreak features; the degree of fragmentation was also associated with the extent of the fire perimeters.

I also assessed the sensitivity of the metrics to changing grain sizes and developed a scaling rule to determine patterns across multiple scales. The results of the multi-scale analyses prompted me to infer that the effects of scale on spatial patterning can be summarized into three categories: monotonic change and predictable pattern, monotonic change with no simple scaling relationship, and non-monotonic change with erratic responses. The pattern analysis revealed that shape metrics LSI, MSI and FRAC exhibited a predictable response, following a power law relation. This was consistent with studies undertaken by others, indicating that a robust scaling rule can be derived to determine patterns across multiple scales. Nevertheless, most of the LPM used in this study showed either a monotonic change with no simple scaling law or a non-monotonic change with erratic response across different sites. For some of the LPM with monotonic changes, however, it would be possible to draw a correlation (negative or positive) between the metrics and grain sizes. This hinders one to develop a robust scaling rule to transfer information (predict patterns) across different events at coarser spatial resolutions, and hence a site-specific or scale-specific pattern analysis is desired. Based on the findings on the five different spatial resolutions, I was prompted to conclude that multi-scale analyses of post-fire residuals at limited grain sizes (R_4, R_8, R_{16}, R_{32}, and R_{64}) would remain a challenge in landscape ecology. Although a single aggregation method (i.e., independent aggregation method as opposed to 'iterative' method and majority rule as opposed to random rule-based method) was used, I tended to infer that spatial aggregation methods can also have an impact on characterizing patterns over multiple scales. Therefore, the multi-scale analysis undertaken, based on a single-scale method, at the given grain dictates further theoretical and empirical studies.

Another objective of this chapter was to determine whether particular types of land cover are likely to escape burning and whether post-fire residual patches are associated with fire edge and surface water. Addressing this issue would allow forest managers to answer the question 'where to retain post-harvest residual patches'. The incidence of post-fire residuals is usually associated with the dominant land cover type, and this has been manifested in most of the fire events. However, the occurrence of residual patches in a disturbed landscape was not only related to the dominant land cover, but also to land cover types that are less prone to high fire severity such as treed and open wetlands. In some of the fire events, for example, certain land cover types (e.g., treed wetland and sparse conifer) were more represented in residual patches despite their low abundance in the landscape. This could be associated with low tree density, high moisture content of species, and wetness of the land surrounding the fire.

3. Estimating the variables that govern the existence of residual vegetation patches within a fire disturbed boreal landscape

Abstract

Wildfires are frequent boreal forest disturbances in Ontario and emulating them with forest harvesting has emerged as a common forest management goal. Wildfires typically contain a considerable number of unburned residual patches of various size, shape, and composition; understanding the characteristic features of these residuals provides insights for effectively emulating forest disturbances by harvesting operations. The occurrence of residual patches within eleven boreal wildfire events; each ignited by lighting and never suppressed, is studied. The importance of different geo-environmental factors that are believed to influence the existence of residual patches within a disturbed landscape is studied using Random Forest. The factors include distance from natural firebreaks (wetland, bedrock and non-vegetated areas, and water), land cover type, and topographic variables (elevation, slope and ruggedness index). The effects of analytical scale (i.e., spatial resolution) on determining the importance of each of the predictor variables were also assessed; the importance of the variables was examined at five spatial resolutions (R_4, R_8, R_{16}, R_{32}, and R_{64} m). The results show that natural firebreak features, specifically wetlands, are among the most important variables that explain the occurrence of the residual patches. Topographic variables are usually contributing factors, particularly in rugged terrain, but in this study, topographic variables of ruggedness index, slope and elevation are found to be less informative in explaining the presence of residual patches. Besides, there is some variability in the relative ranking of the importance values for predictors across the studied fire events. I also found that the importance of the predictor variables exhibit a slight variation along the gradient of scales (spatial resolutions) for a single event, indicating the effect of scale on variable importance.

Keywords: Boreal forest, fire disturbance, residuals, random forest, predictor variables, variable importance, spatial resolution, marginal effect

3.1. Introduction

Forest management practices in the boreal forest often alter the species composition and forest structure and reduce the biodiversity of a landscape (Long 2009). The managed forest should emulate the patterns occurred following natural disturbance so that the forest biodiversity within the landscape would be maintained (Dragotescu and Kneeshaw 2012), as managed forest lands have a controlling influence on natural processes (e.g., wildfire) (North and Keeton 2008). One of the approaches for maintaining biodiversity, while managing forests, is to retain certain structural elements (e.g., live trees) of forest habitat within harvest units. The patterns of such residuals are expected to mimic the incidence of unburned islands following a fire disturbance. This is the natural disturbance emulation approach that has been suggested as a model for sustainable forest management strategy in various boreal ecosystems (North and Keeton 2008). The natural emulation approach proposes using knowledge of natural patterns of post-disturbance remnants as guides for mimicking human related disturbance activities, such as harvesting (Andison 2004). Such a natural pattern approach is considered as an effective tool in the conservation of biodiversity and promotes ecosystem resilience (Long 2009), which is the capacity of ecosystems to absorb disturbances without undergoing fundamental change (Drever et al. 2006). It was also argued that traditional management practices tend to produce homogenous forest cover than those naturally disturbed; hence increase the likelihood of unexpected catastrophic change within an ecosystem (Drever et al. 2006). Therefore, natural-disturbance-based management allows forest managers to maintain the biodiversity, and the structural and compositional heterogeneity of forest ecosystems.

The natural approach to forest management practice requires understanding of the various aspects of wildfire disturbance regimes (e.g., disturbance frequency, disturbance intensity and attributes, types of post-fire remnants) (North and Keeton 2008). One way by which such a natural approach to forest management is implemented in the boreal forest is by understanding the patterns of the legacies following a wildfire and implement management practices that mimic the natural patterns. In the boreal forest, a considerable number of unburned patches of various sizes, shape, and composition occur within a burned landscape. This requires exploring the patterns and characteristic features of the residual patches and improving our understanding of the mechanisms and causal factors of residual structure. This provides insights for implementing natural-disturbance-based management practices. To this end, this study was determined to implement a reliable and broadly applicable data mining approach to examine the geo-environmental factors that are responsible for the occurrence and distribution of residual patches within a given burned landscape.

3.1.1. Spatial language: residual and null-residual patches

The study considered a binary response variable: the presence and absence of residual vegetation patches, hereafter described as residual and null-residual patches respectively. The residual patches are defined based on the work of Remmel and Perera (2009) in which the NDPE guide (OMNR 2001) was used to define and extract the residual patches (§1.3). The residual patches, as noted in Remmel and Perera (2009), are composed of treed and vegetated land cover types, contiguous pixels (patches) of greater or equal to 0.25 ha, and are contained within the fire footprint. The null-residual patches, which are supposed to be within the burned areas, are used to describe the absence of residual patches within the fire footprint; they can also be described as ‘pseudo’ or ‘burned’ residual patch within a fire footprint. The null-residual patches were simulated by random placement of patches within the burned landscape in which their shape, size, and orientation mimic the actual residual patches (Figure 3.1).

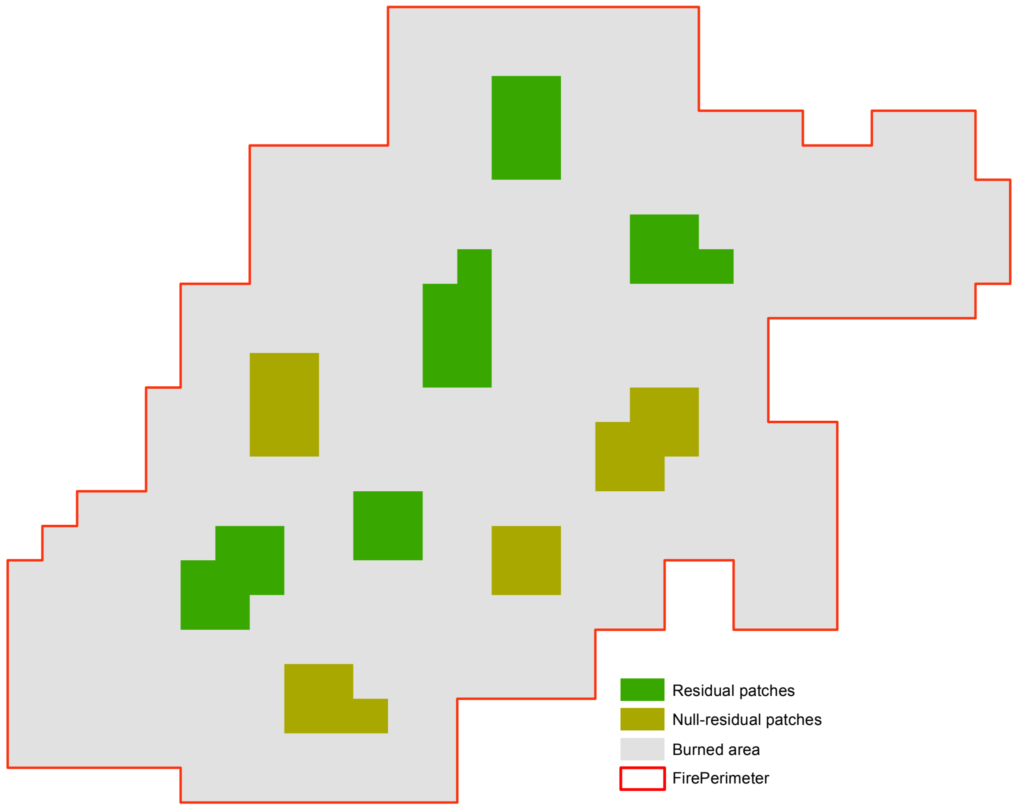


Figure 3.1. Hypothetical example of spatial language: existing residual (presence data) and null-residual patches (absence data) within a fire footprint. The number, size, shape, and orientation of null-residual patches mimic the existing residual patches.

3.1.2. Random Forest

Wildfire processes and the resulting landscape patterns involve complex interactions of various factors; the relationship among these factors may be strongly non-linear (De'ath and Fabricius 2000). The patterns of post-fire landscape structure are also unplanned, complex, and heterogeneous making accurate predictions difficult (Turner et al. 1997; Guisan and Zimmerman 2007). The need to handle such complex interactions among the variables led to the development of various statistical or machine learning approaches (such as regression tree analysis (RTA), support vector machines (SVM), and classification and regression trees (CART), in conjunction with GIS and remote sensing (Austin 2002).

Classification and regression trees (CART) is a machine learning method used for constructing prediction models by recursively partitioning sample data into smaller (and increasingly homogeneous) groups, using decisions based on conditions applied to independent variables. Classification trees are designed for grouping observations where the dependent variables are categories, while regression trees are used when the outcomes are continuous variables. CART-based analyses can be used for interactive exploration, modelling complex datasets, and prediction of patterns and processes (De'ath and Fabricius 2000; Schroff et al. 2008). A CART approach has the: 1) ability to handle nonlinear relationships, 2) flexibility to deal with a broad range of response types, 3) ability to handle missing values in the response and explanatory variables and, 4) ease of construction and interpretation (De'ath and Fabricius 2000; Breiman 2001; Strobl et al. 2008).

Recently, there has been a growing interest in ensemble-learning methods, which use multiple models to obtain better predictive performance. Unlike the standard CART, where a single classification tree is produced for classification, ensemble-learning methods generate many classified outcomes and aggregate their results (Liaw and Wiener 2002; Rodriguez et al. 2006). The methods have grown in prominence due their ability to handle large numbers of predictor variables (Cutler et al. 2009) and overcome the problem of over-fitting associated with CART (Breiman 2001; Evans and Cushman 2009). The most widely known ensemble methods are bagging, boosting, and random forest (RF). RF is a CART based on the bootstrap method that provides well-supported predictions with a large number of predictor variables (Cutler et al. 2007; Strobl et al. 2008), for both classification and regression problems. The algorithm has become popular and appears to be powerful in different applications, including gene expression data analysis (Archer and Kimes 2008), abundance of soft coral reefs (De'ath and Fabricius 2004), agricultural management practices (Watts and Lawrence 2008), and ecohydrological (Evans and Cushman 2009). It has also been applied in remote sensing mapping and vegetation prediction (Iverson et al. 2004; Pal 2005; Gislason et al. 2006; Schroff et al. 2008; Chehata et al. 2009) and prediction of species or vegetation type occurrence (Peters et al. 2007).

The principle of RF is to combine many binary decision trees built using several bootstrap samples coming from the learning sample and choose randomly a subset of explanatory variables at each node (Breiman 2001; Genuer et al. 2010). RF, similar to CART, explains the variation of a single response (dependent) variable by one or more explanatory (independent) variables. The response variable could be categorical (classification trees) or numerical (regression tree) while the explanatory variables can be categorical and/or numeric. RF is also similar to bagging in that bootstrap samples are drawn to construct multiple trees (Rodriguez et al. 2006; Prasad et al. 2006). In training, the algorithm creates multiple bootstrapped samples of the original training data, and then builds a number of unpruned classifications for each bootstrapped sample set (Chehata et al. 2009). In a typical bootstrap sample, two-thirds of the data are used for constructing any particular tree. A classification tree is fit to each bootstrapped sample, but at each node only a small number of randomly selected variables (e.g., square root of the number of variables) are available for the binary decision tree (partitioning) (Breiman 2001). Observations in the original dataset that do not occur in a bootstrap sample (i.e., one-third of the data that are not used in the construction of a tree) are called out-of-bag (OOB) observations. The OOB sample is used to estimate the prediction or classification error and evaluate variable importance (Genuer et al. 2010).

The ensemble learning method drew the attention of many researchers because it: 1) is a nonparametric and nonlinear classifier that does not require any assumption on data distribution (Breiman 2001; Strobl et al. 2008), 2) adds an additional layer of randomness into the training of the trees (Breiman 2001; Liaw and Wiener 2002); 3) runs on large datasets and can handle thousands of variables; they do not over-fit (Breiman 2001; Strobl et al. 2008; Watts and Lawrence 2008; Chahata et al. 2009); 4) has high predictive performance (Strobl et al. 2008) and is computationally efficient in both training and classification (Schroff et al. 2008); and 5) is faster than bagging and boosting, and provides additional pieces of information (e.g., importance of predictors, internal accuracy measure and proximity analysis) (Breiman 2001; Cutler et al. 2007). Yet, it is important to note that RF is more of a 'black box' approach because one cannot examine the individual trees separately (Prasad et al. 2006).

3.1.3. Research framework

Madoui et al. (2010) noted that the presence and distribution of residual patches can be attributed to one of the following hypotheses: 1) given wind patterns, certain areas of the forest matrix may be preferentially spared from burning, 2) owing to the variation in their fuel properties, some land cover types are less susceptible to fire, and 3) the presence of natural firebreaks (e.g., wetlands and water) trigger some areas of the forest matrix to be spared from burning. Moreover, a number of studies (e.g., Vera 2001; Ryan 2002; Perera et al. 2007; Cuesta et al. 2009) have undertaken to assess the factors that govern the occurrence of post-fire residual patches. The studies indicated that the existences of residual patches is attributed to complex interactions among various factors such as wind variation, topography, vegetation type, fire size, natural firebreak features, and pre-fire vegetation characteristics. Most of these inferences are based on post hoc observations rather than on testing a priori hypotheses; but observations made following the fire would be constrained in space and time, and may not provide a comprehensive view of the post-fire forest characteristics. Yet, some of the approaches used to evaluate the factors (e.g., standard regression analysis or classification trees) have been challenged to provide a reasonable analysis of the variables that govern the residual occurrence because of the linearity assumption of the techniques. Therefore, my study used an advanced machine-learning algorithm (RF) to examine the importance of different predictor variables that could explain the occurrence of residual patches. This includes distance from natural firebreaks (wetlands, bedrock and non-vegetated areas, and water), land cover type, and topographic variables (elevation, slope and a ruggedness index: RI). The idea of RI was introduced to express the amount of elevation difference between adjacent cells on a digital elevation grid (Riley et al. 1999).

The objectives of this chapter were to: 1) assess the variables that explain the existence of residual patches, 2) evaluate the variability of the measures of variable importance across five different spatial resolutions (R_4, R_8, R_{16}, R_{32} and R_{64}), and 3) investigate the marginal effect of the predictor variables on the occurrence of residual patches.

Furthermore, some research hypotheses were formulated to address the importance of the predictor variables to explain the residual patches. Some of the factors that explain residual patches (e.g., topographic variables and natural firebreaks – wetlands and surface water) are more important at a local scale (e.g., fire event level) while others (e.g., climate) are determinant at larger geographical extent (e.g., landscape or ecosystem level) (Swystun et al. 2001; Cuesta et al. 2009). Hence, I hypothesized that variables associated with natural firebreaks would be among the most important predictors to explain residual patches. Topographic variables are also considered among the most important variables that affect the patterns of post-fire residuals (Meddens et al. 2008; Madoui et al. 2010); residual patches are more prevalent in rugged terrain than in plain or flat lands. Thus, in this study I also hypothesized that topographic variable (e.g., slope and ruggedness index) would be informative for the occurrence of residual patches. Boreal wildfires involve factors operating at different scales (King and Perera 2006), and the configurations of patch characteristics are sensitive to a change in spatial resolution (Wiens 1989; Cain et al. 1997; Zhu et al. 2006). Thus, it is expected that that the measures of variable importance would be sensitive to scale change.

3.2. Methods

3.2.1. Study area

In this study, the relative importance of different predictor variables that explain the occurrence of residual patches was explored at the fire event level. This is part of an effort to develop and provide a reliable and repeatable method for understanding residual patches and investigating the agents that govern the patterns of post-fire landscape structure and fire behaviour. The study was based on 11 fire events that are located within one of Ontario's largest ecoregions: the Ontario Shield Ecozone (for details see §1.5.5). Fires in boreal forests are sometimes large and extensive that burn tens of thousands of hectares; however, small fires (< 100 ha) are the most frequent, yet it is the largest fires that primarily shape the landscape structure of the boreal forests (Thompson 2000). The 11 fire events studied in ecoregion 2W varied in size ranging from 58 to 4525 ha (the extent of each of the fire print is presented in Figure 2.2 of §1.5.5), and for the sake of ease of analysis the 11 fire events were categorized into three broad groups based on their size (Table 3.1).

Table 3.1. The three categories of fire events, categorized based on their size.

Fire event-class size	**Fire footprint extent (ha)**	**Fire footprint ID**
Large sized events	Fire footprint extent ≥ 3000	F01, F06, F08, and
Medium sized events	Extent > 100 and < 3000	F04, F05, F07, and
Small sized events	Fire footprint extent ≤100	F02, F03, and F09

3.2.2. Landscape data: presence and absence data

3.2.2.1. Presence data – residual vegetation patches

The use of RF for classification and prediction requires a response variable and explanatory variables. The response variable in my study was the presence and absence of a residual patch, which are described as residual and null-residual patches respectively. The presence (residual patches) data were extracted from classified Ikonos images (Spectranalysis 2005). The delineation and categorization of residual patches from classified Ikonos images follow the OMNR's definitions of residual patches (Remmel and Perera 2009). OMNR's definitions of residual patches are based on size and location of a patch in relation to the perimeter of a disturbed landscape (OMNR 2001). The residual patches (presence-data) considered in this study were clusters of unburned pixels, which are: 1) composed primarily of treed land cover classes (e.g., dense conifer, sparse conifer, deciduous, and alder and shrub woodland) and other vegetated land cover classes (low shrub, treed wetland, open wetland, and marsh), 2) ≥0.25 ha of the cover class, and 3) > 1 pixel inset from the fire perimeter (Remmel and Perera 2009). For detailed description on the types and meanings of different patches, refer to §1.3. Since the study was based on multi-scale analyses, the residual patches were obtained over the five spatial resolutions: R_4, R_8, R_{16}, R_{32}, and R_{64}; hence the relative importance of the predictors would be assessed across a gradient of scales (spatial resolutions) (§2.2.2).

3.2.2.2. Absence data – null-residual patches

The response variable often incorporates the presence and absence-data but majority of the data that are available today consist of presence-only data (Zaniewski et al. 2002). Yet presence-only data are the most difficult element to integrate into statistical modelling. This study was based on presence-absence data where residual patches extracted from classified Ikonos images were considered as presence-data. However, information pertained to absence-data is often unavailable or difficult to obtain; hence a computer simulation approach has been suggested to algorithmically generate 'pseudo' absence (Zaniewski et al. 2002). Therefore, the initial step in the analysis was to algorithmically extract the presence-data (null-residual patches) within the burned areas. An algorithm was designed (Figure 3.2) to randomly generate null-residual patches within burned areas in which the size, shape, and orientation of the null-residual patches mimic the residual patches.

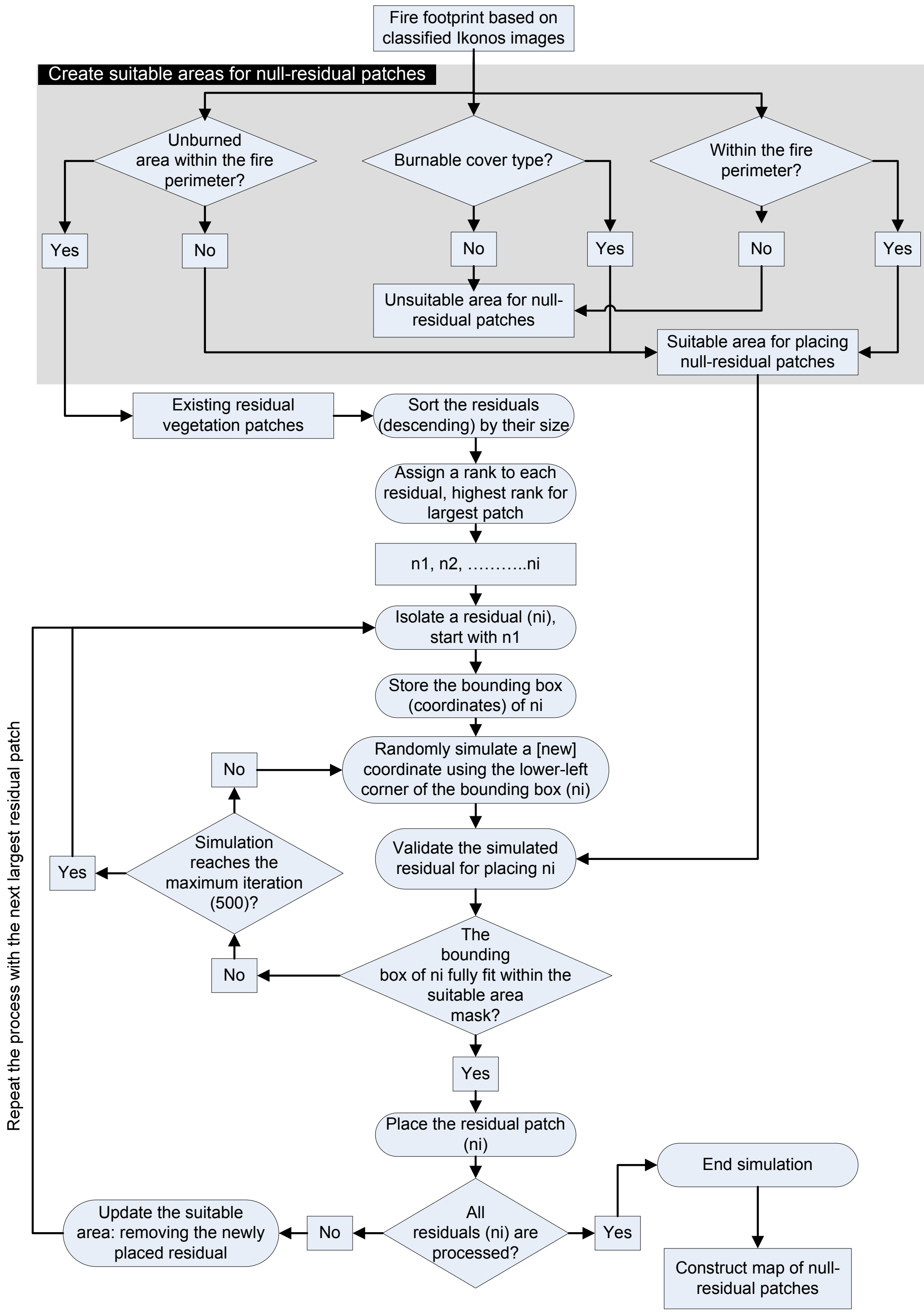


Figure 3.2. Conceptual simulation design producing map of null-residual patches (absence-data). The algorithm was designed to randomly generate absence-data in areas defined as suitable for placing null-residual patches. The initial step in the algorithm was to determine areas that are considered suitable for placing the null-residual patches.

The algorithm involves two main parts: 1) identifying suitable areas for extracting null-residual patches and 2) randomly placing the residual patches within valid and identified burned areas. The suitability layer for null-residual patch placement within the burned landscape was based on clearly defined rules, such that a null-residual patch: 1) should be contained within the perimeter of a fire footprint, 2) should not occupy or overlap an existing residual patch, and 3) must comprise burnable land cover types. The second step in the algorithm was to randomly place the residual patches (as null-residual templates) within the burned areas, one at a time, such that map of null-residual patches would be constructed. The random placement includes: 1) sorting all the residual patches by their size in descending order, with the largest patch to be simulated first; 2) assigning a rank for each residual, giving rank 1 for the largest residual, rank 2 for the second largest, and continues until the smallest residual was assigned with its respective rank; 3) isolating the largest residual patch (ranked first) and defining and storing the bounding box of the residual (i.e., lower-left and upper-right corner of the residual); 4) for each residual patch, simulating a random coordinate which becomes the lower left corner of the bounding box of the residual patch; 5) testing whether the actual patch shape fits within suitability matrix as identified by the bounding box; and 6) if the simulated coordinates met the suitability requirement, the algorithm places the residual patch; if it fails to meet the criteria, the algorithm simulates a new random coordinate. The process continues until the maximum number of permitted iterations (500) is reached. The area occupied by a previously placed null-residual patch is no longer available or suitable for generating the subsequent largest residual. Hence, the algorithm updates the suitable areas by removing the newly placed residual from the suitable areas. The overall process continues until the smallest residual patch is placed in the burned landscape. The algorithm was applied to each of the eleven fire events, across the five spatial resolutions.

3.2.3. Predictor variables

The existence of residual patches within a fire-disturbed landscape, across five different spatial resolutions, was studied in relation to different explanatory variables that are believed to explain the residual patches: topographic parameters, natural firebreaks, and land cover types. Table 3.2 lists the possible determinants of residual patch occurrence.

Table 3.2. List of explanatory variables (and their descriptions) used in analysing the patterns of post-fire residual occurrences.

Explanatory variables	**Description**
Wetland (WL)	Euclidean distance to nearest wetland from a residual patch
Water body (WA)	Euclidean distance to nearest water from a residual patch
Terrain ruggedness index (RI)	Elevation difference between adjacent cells
Slope (SL)	Rate of maximum change in elevation values from each cell
Elevation (EL)	Average elevation value of a residual patch
Land cover (LC)	Land cover type where a residual patch occurs
Bedrock and non-vegetated (BV)	Euclidean distance to non-vegetation from a residual patch

Topographic variables: A fire ignited at the bottom of a slope spreads rapidly uphill because flames dry matter ahead of the advancing front, while a fire ignited on the top of a slope tends to spread slowly downhill in the absence of the pre-warming and drying process (Viegas 1993; ESA 2002). Rugged and undulating terrain is less severely affected by fire than flat areas (Chafter et al. 2004), indicating the impact of topographic features on fire spread and the patterns of post-fire landscape structure (Epting and Verbyla 2005). In this study, three topographic variables: terrain ruggedness index (RI), slope and elevation, which are derived from DEM were considered. The RI value for each cell was computed using an algorithm (AML) developed by (Evans 2004). The idea of RI was introduced by Riley et al. (1999) to express the amount of elevation difference between adjacent cells on a digital elevation grid. The model considers elevation values from a center cell and the eight first-order neighbours (Equation 1). The other two variables: slope and elevation were also obtained from the same DEM. The spatial resolutions of the DEM used in the study vary in resolution from 0.75 arc seconds (~20 m) to 3 arc seconds (~90 m).

$$RI = Y\left[\sum\left(x_{ij} - x_{00}\right)^2\right]^{1/2} \quad (1)$$

x_{00} is the elevation of the center cell; x_{ij} *is the elevation of each neighbor cell to cell (0,0)*

Natural firebreaks: natural firebreak features are gaps in forest cover that may escape fire and act as a barrier to slow or stop the wildfire spread. A firebreak may occur naturally where there is a lack of vegetation or forest fuel (e.g., water, wetland, or river) or is impeded by anthropogenic features (e.g., road, transmission line or highway). The abundance and distribution of natural firebreaks within a fire event can contribute in the incidence and spatial distribution of residual patches. The fire events considered in this study are characterized by substantial natural firebreak features, specifically wetland and surface water, which would have served as a barrier for fire spread and played a role in the occurrence of residual patches. The existence of residual patches in relation to different types of natural firebreaks including water, wetland, bedrock and non-vegetated area were assessed; all these predictors were extracted from the existing pre-fire Ontario Land Cover Data Base (OMNR 2005). The Ontario Land Cover Data Base provides a classification of 27 land cover types across the province of Ontario. The land cover classification was generated using a digital image analysis (supervised image classification) of Landsat-7 satellite images recorded between 1999 and 2002, most from 2000 onward (Spectranalysis 2004). The classification was conducted using the original spatial resolution of the source data (30 m spatial resolution); however the classified maps were resampled into the desired spatial resolutions (R_4, R_8, R_{16}, R_{32}, and R_{64}).

Land cover: land cover types also play an important role in determining the patterns of fire spread and post-fire landscape structure (Mermoz et al. 2005). This is attributed to the fuel load and moisture content of fuel particles of the forest ecosystem, as high values of moisture content slow the rate of burning; even prevent fire spread (Nelson 2001). Forest fuels have different properties depending on species, whether they are alive or dead, and the amount of fuel available along with its spatial distribution (Viegas 1993; ESA 2002; van Wagtendonk 2004). Similar to the variables related to natural firebreak features, the land cover variable was extracted from the pre-fire Ontario Land Cover Data Base.

Weather variables: the occurrence, intensity, seasonality and spread of fire in boreal forests and patterns of post-fire landscape structure also depend on weather and climate (Dale et al. 2001; Johnson 1995). The primary weather variables are temperature, precipitation, wind and relative humidity. Of these factors, by far the most important factor is wind, defined by its velocity and direction (Rowe and Scotter 1973). Wind blows through the forest, dries the foliage and makes it more flammable; the fire is further intensified by the abundance of forest fuel as well as topography (i.e., ascending slopes). Despite their contribution to the ignition and spread of wildfire, weather variables such as wind, temperature, and precipitation were not incorporated into this study. The spatial resolution of the weather related data was too coarse for the scale of observations considered in the study. Also, the fire events are located within a single ecoregion where the patterns of weather variables, including temperature, precipitation, and humidity are similar.

Most of the predictor variables were obtained in raster data format. The values of the raster data (i.e., values of each predictor variable) within the residual (and null-residual) patches were computed using zonal statistics, considering each residual (and null-residual) patch as a 'zone'. However, the extent of some of the residual patches was too large (Figure 3.3a) to get a reasonable estimate of the mean values of the predictors (e.g., mean distance to nearest wetland from a patch). For example, the residual patch shown in Figure 3.3a has a range of values (distance to water) and computing a single mean zonal value (mean distance to water from a patch) for the patch would give a misleading estimation. Therefore, a series of concentric 100 m inward buffer rings were created from the edge of the fire perimeter to calculate the mean (zonal) values of the predictors for a residual more effectively (e.g., Figure 3.3b). A geometric intersection of the residual (and null-residual) patches and the buffer rings was then applied. Thus, each portion of a residual in a single buffer ring (i.e., each intersected feature) was considered as a 'zone'; the mean (zonal) raster values of each of the predictors were then computed for each portion of a residual (and null-residual) patch. The overall process of buffering, geometric intersection, and variable estimation was performed for all the residual and null-residual patches throughout all the fire events across the five spatial resolutions. Once all the required variables were computed for the residual and null-patches, the relationship among pairs of predictor variables was explored using correlation matrices. This was performed to assess whether there was a significant correlation among pairs of variables and examine their redundancy.

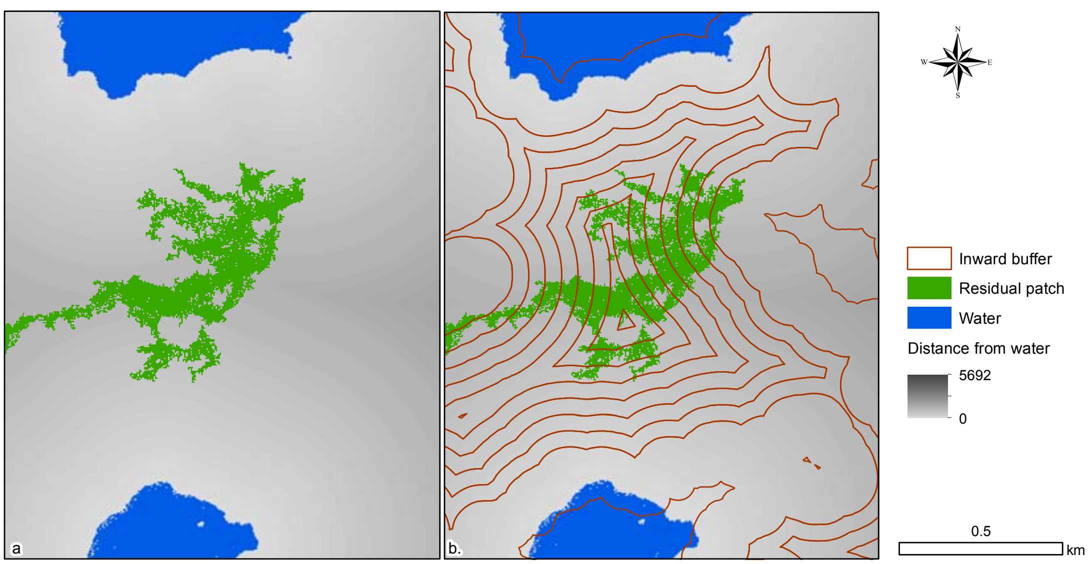


Figure 3.3. Estimating mean zonal values of predictor variables for each residual patch: a) the existing residual patch could be considered as a 'zone' for estimating the mean zonal values of each predictor, but the patch may be too large to be considered as a single patch as it has a range of values that would mislead the estimation, and b) a series of internal 100 m wide buffer rings were created from the edge of the fire perimeter and each portion of a residual in a single buffer ring as a 'zone'.

3.2.4. Random forest (RF) implementation

The existence of residual patches was examined using the RF algorithm (Breiman 2001) as implemented in R (Liaw and Wiener 2002). RF combines two sources of randomness that improve the prediction accuracy: bagging and random feature selection to construct each classification tree (Robert-Granie et al. 2009). The RF implementation in this study used classification trees because the response variable under investigation, presence (residual) or absence (null-residual) of patches is a binary (categorical) variable.

Two user-defined parameters, which affect the stability of the results, are required to execute RF (Strobl et al. 2009a). These parameters are usually optimized to minimize the generalization error (Iverson et al. 2004; Gislason et al. 2006; Peter et al. 2007). The first parameter is the number of trees to grow (n_{tree}); this is akin to the number of simulations or randomizations. There is no specific rule to define the number of trees to grow, but the general rule is that a small number of n_{tree} can result in poor classification performance while larger values of n_{tree} should provide a more stable classification (Liaw and Wiener 2002; Prasad et al. 2006). If auxiliary information such as variable importance is desired, a large number of trees are required. The second parameter is the number of randomly selected variables used to split the nodes (m_{try}). Compared with the standard classification tree, an additional random factor is included in the RF; at each node a random subset of m_{try} variables has to be set and the best splitting variable among those m_{try} is used to split the node. The m_{try} affects both the correlations between the trees and strength of the individual trees (Peters et al. 2007). This parameter requires some subjective judgment, but Breiman (2001) defined this parameter as $m_{try}= \sqrt{n}$ (where n is the number of predictor variables), with a minimum of m_{try} = 2. RF with different n_{tree} values (i.e., 100, 200, 300, 500, 600, 700, 800, 900, and 1000) at m_{try} = 2 and m_{try} = 3 was executed (for a single fire event) to initially assess the sensitivity of the model to n_{tree} and m_{try}. Overall, low OOB error values were observed for all n_{tree} constructed, with OOB error ranging from 9.56% to 9.89%o for m_{try} = 3 (Table 3.3). This suggests that the error estimates showed stabilization of the overall error across the n_{tree} values considered; adding more trees to the model did not change the OOB error substantially. Based on this, the RF was ultimately implemented in R with n_{tree}=100 and m_{try} = 3. Figure 3.4 shows a schematic view of the overall logic undertaken to execute RF as implemented in R.

Table 3.3. OOB estimate of error rate across different n_{tree} values (and m_{try} of 2 and 3) for F01. The OOB error is estimated internally using one-third of the cases left out of the bootstrap sample and not used in the construction of the tree.

ntree	**OOB error values (%)**	
	m_{try} = 2	m_{try} = 3
100	9.88	9.68
200	9.89	9.49
300	9.65	9.41
400	9.71	9.55
500	9.56	9.55
600	9.61	9.48
700	9.81	9.50
800	9.63	9.34
900	9.68	9.52
1000	9.68	9.45

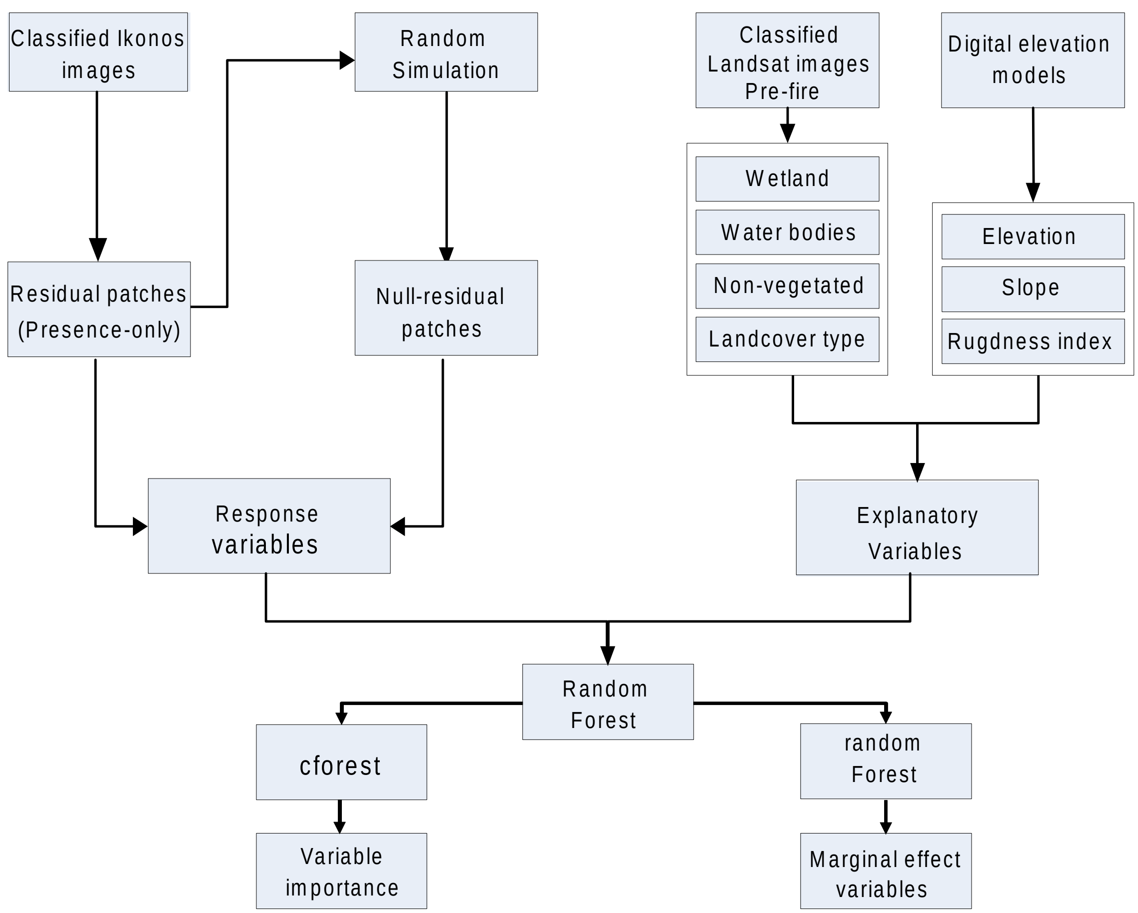


Figure 3.4. RF implementation for assessing the importance and marginal effects of the predictor variables. The two RF implementations (cforest and randomForest) were considered to identify the relative importance of the predictors and examine the partial dependency of the response variable on the predictors respectively.

3.2.4.1. Spatial variable importance

Random Forest performs several types of statistical analyses such as estimating local error, measuring missing values, and assessing the importance of predictor variables in high dimensional settings (Breiman 2001; Cutler et al. 2007; Strobl et al. 2008). RF estimates the importance of a variable at how much a prediction error increases when OOB data for that variable is permuted while all others are left unchanged (Liaw and Wiener 2002). A measure of variable importance based on permutation importance, the Mean Decrease in Accuracy (MDA) was computed. In a MDA, the OOB data are used to obtain estimates of variable importance by evaluating their contribution to the prediction accuracy. The measure of importance value in MDA considers the proportion of cases in the correct classes with permuted OOB data and the proportion of cases in the correct classes where OOB data are not permuted. The MDA thus averages the difference between these two accuracies over all trees in the forest and normalizes it by the standard error (Equation 3) (Robert-Granie et al. 2009; Cutler et al. 2007). The importance score of a variable is computed as follow:

$$VI_{ntree(xi)} = \frac{\sum_{ntree=1}^{ntree} Pred.n_{tree} - Pred.n_{tree(xi)}}{ntree} \qquad (2)$$

$$VI_{xi} = \frac{VI_{ntree(xi)}}{st} \qquad (3)$$

Where:

$VI_{xi}\ (or\ MDA)$ is the importance score of a variable

$Pred.n_{tree}$ is predicted class before permutation

$Pred.n_{tree(xi)}$ is predicted class after permutation

There are two random forest implementations in R: *randomForest* (Breiman 2001) and *cforest* (Hothorn et al. 2008). In the context of this study, the former is termed as the original permutation importance while random forest based on *cforest* is described as conditional variable importance. The variable importance, computed in *randomForest* and *cforest*, is based on a random permutation of the predictor variables, as described above. However, variable importance measures using *randomForest* are subject to biases in favour of variables with many categories, and continuous variables that affect variable selection in a single tree (Strobl et al. 2009b). The original permutation importance also overestimates the importance of correlated predictor variables (Strobl et al. 2009b). This means that correlated predictor variables tend to appear more important than uncorrelated variables (Strobl et al. 2009b). For example, if there are three variables (X_1, X_2, and X_3) in a dataset, and if two of the variables (X_2 and X_3) are correlated to each other; high importance values are assigned to the correlated variables X_2 and X_3; even if one of the variables (e.g., X_2) does not have any impact on the response variable Y. It is also described that variables with different scales of measurement are likely to be over/under estimated when the original permutation importance is applied.

An alternative permutation importance (i.e., conditional variable importance) has been suggested to overcome the above problems and to guarantee unbiased variable selection and variable importance for predictor variables. This approach considers the conditional effect of a variable on the response function. For example, if a variable's conditional effect is negligible (e.g., X_2); variable importance for that variable would be minimal; even if the variable has a significant correlation with the most influential variable in the dataset. Some of the variables considered in this study might correlate to each other (e.g., RI or slope) and the predictor variables are of different types (i.e., different categories and different scales of measurement). Therefore, conditional variable importance measure was performed to determine the set of variables that are deemed important to explain the residual patches for each fire event, across the specified five spatial resolutions.

3.2.4.2. Partial dependence plots (PDP)

After identifying the most relevant predictor variables, the next step was to get an idea of the dependence of the response variable on each of the predictors, using a partial dependence plot (PDP). PDP gives a graphical depiction of the marginal effect of a variable on the predictions of "blackbox" classification and regression (Cutler et al. 2007). It summarizes the effects of predictors on the probability of occurrence after accounting for the average effect of all other variables (Friedman 2009). It is designed to show the dependence of the response variable on X_1 as averaged over the distribution of values of the other predictor variables (X_2, X_3,……X_n). Yet, PDP may not show a comprehensive description of a variable's effect on the prediction, but it can show how the response variable changes as you change the predictor (Hastie et al. 2009). PDP visualizes not only the additive effects of each predictor on the response but also the interacting effects of predictors (Jun 2013). However, visualizing the effects of predictors is limited to low-dimensional views (Hastie et al. 2009); one can only display the effects of one or two variables.

PDP is used to graphically characterize the relationship between selected predictor variables and predicted probabilities of residual patch occurrence. According to Cutler et al. (2007), if classification problems, say K classes, there is a separate response function for each class. Letting $P_k(X)$ be the probability of membership in the k^{th} class given the predictors (details can be found in Cutler et al. 2007); $X = (X_1, X_2, X_3, \ldots X_n)$, the k^{th} response function is given by (Equation 4):

$$f_k(x) = log\ \ p_k(x) - \frac{1}{K}\sum_j \log(p_j(x) \tag{4}$$

Where *k is the class (*e.g., residual or null-residual) of *K* total classes, and p_j is the proportion of votes for class *j.*

For the case when *K* = 2 (presence = a, and absence = b), if *p* denotes the probability of "success" (i.e., presence of residual patches), then $p_a(x) = 1 - \ p_b(x)$. So, the above expression reduces to:

$$f_a(x) = logp_a(x) - \frac{1}{K}\sum logp_j(X) \tag{5}$$

$$f_a(x) = logp_a(x) - \frac{1}{2}[logp_a(x) + logp_b(x)] \tag{6}$$

$$f(x) = 0.5log\frac{p_a(x)}{1-p_a(x)}\ = 0.5logit(p_a) \tag{7}$$

3.3. Results

3.3.1. Residual and null-residual patches

The process of random extraction of null-residual patches was applied to each of the fire events at R_4, R_8, R_{16}, R_{32}, and R_{64}. The random extraction of null-residual patches at 4 m spatial resolution for an individual fire event (F01) is shown in Figure 3.5. The null-residual patches were expected to mimic the residual patches in number, size, shape, and orientation. However, the algorithm sometimes failed to generate a suitable area for placing the residual patches; this is particularly true for large residual patches. The spatial extent of some of the residual patches was large; larger patches are harder to randomly fit within the desired burned areas because they are likely to overlap areas of existing residual patches or other unburnable land cover types, which are not suitable for placing residual patches. As a result, the random simulation was not able to completely mimic all the residual patches, and hence the total number of null-residual patches extracted for most the fire events was not equivalent to the total number of residual patches (Table 3.4). The table also shows that the number of residual patches within the fire events, across the five spatial resolutions was different. Table 3.5 also presents the extent of the residual and null-residual patches, and the proportion of fire footprint print occupied the patches at 4 m spatial resolution.

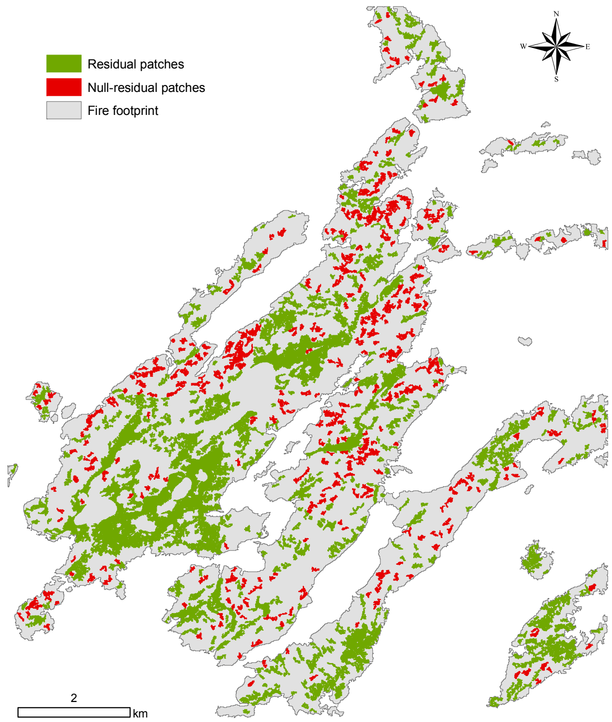


Figure 3.5. This map shows the distribution of the residual patches extracted from classified Ikonos image and the algorithmically simulated null-residual patches for a single fire event (F01), at 4 m spatial resolution.

Table 3.4. The total number of existing residual patches obtained from Ikonos images (R) and the total number simulated null-residual (NR) patches at five spatial resolutions; this is regardless of size, shape, and orientation of the residual patches.

	Spatial resolutions									
Fire ID	4 m		8 m		16 m		32 m		64 m	
	R	NR	R	NR	R	NR	R	NR	R	NR
F01	469	276	442	357	440	416	400	389	364	361
F02	9	2	10	7	11	11	8	8	6	6
F03	6	3	10	7	12	12	11	11	6	6
F04	102	64	102	91	110	107	103	100	119	119
F05	170	137	187	166	185	178	170	165	150	145
F06	327	218	360	290	379	328	326	295	299	282
F07	59	52	82	78	72	70	71	70	72	72
F08	162	104	192	156	198	173	200	191	193	188
F09	3	2	2	2	3	3	3	3	4	4
F10	308	154	285	240	312	295	308	291	309	300
F11	40	33	38	35	52	51	50	49	56	55

Table 3.5. The total number of residual and null-residual patches obtained at 4 m spatial resolution, and the extent and the proportion of the fire footprint occupied by the exiting residual patches and simulated null-residual patches where R and NR represent residual and null-residual patches.

Fire ID	**Number of patches**		**Patch area**		**% of footprint occupied by patches**	
	R	NR	R	NR	R	NR
F01	469	276	641.53	128.47	14.18	2.84
F02	9	2	3.56	0.55	4.42	0.68
F03	6	3	3.96	1.31	4.92	1.63
F04	102	64	103.70	27.76	6.58	1.76
F05	170	137	344.90	90.75	15.09	3.97
F06	327	218	787.97	108.43	21.06	2.90
F07	59	52	90.90	30.70	9.67	3.26
F08	162	104	361.03	43.84	11.75	1.43
F09	3	2	1.86	0.65	3.22	1.13
F10	308	154	379.13	79.49	11.57	2.43
F11	40	33	26.28	14.31	3.65	1.99

3.3.2. Pair-wise correlation analysis of predictors

Prior to applying RF for explaining the importance of the predictor variables for the existence of residual patches, a pair-wise correlation analysis was conducted to assess the relationship among pairs of explanatory variables. Ideally, some of the variables (e.g., topographic variables of RI, slope, and elevation) would be highly correlated among each other and may show high redundancy among them. Hence, it was important to examine the correlation among the variables to avoid redundancy in the dataset. To accomplish this, all pair-wise correlation coefficients were computed among the seven explanatory variables, across five spatial resolutions. A correlation coefficient measures the strength and direction of a linear relationship between two variables; it ranges from -1 to +1 and the diagonal elements are always 1. The correlation coefficients of all pairs among seven variables were computed for all fire events, but the following tables show the correlation matrices for selected (large) fire events (F01, F06, F08, and F10) at 16 m spatial resolution. The results showed that almost all the variables, with few exceptions (Table 3.6), are not significantly correlated to each other, suggesting that a multivariate approach to data reduction would not be productive. Thus, the occurrence of residual patches was explored in relation to the seven explanatory variables.

Table 3.6. Correlation coefficients among the explanatory variables of selected fire events; a means to determine whether multivariate approach to data (variable) reduction is required (Bold cells indicate relatively significant correlation among the variables).
.

F01	WL	WA	BV	EL	SL	RI	LC
WL	1.00						
WA	0.07	1.00					
BV	0.23	0.32	1.00				
EL	-0.14	0.27	0.21	1.00			
SL	-0.09	0.10	0.05	0.24	1.00		
RI	0.09	0.00	0.02	0.07	0.47	1.00	
LC	0.02	0.18	0.04	0.05	0.04	0.02	1.00

F06	WL	WA	BV	EL	SL	RI	LC
WL	1.00						
WA	**0.83**	1.00					
BV	0.20	0.16	1.00				
EL	0.32	0.37	-0.17	1.00			
SL	-0.17	-0.12	-0.12	0.24	1.00		
RI	-0.21	-0.21	-0.16	0.25	**0.68**	1.00	
LC	0.03	0.07	0.09	-0.01	0.03	0.0	1.00

F08	WL	WA	BV	EL	SL	RI	LC
WL	1.00						
WA	-0.18	1.00					
BV	-0.04	-0.10	1.00				
EL	-0.12	0.11	-0.07	1.00			
SL	-0.01	0.11	-0.02	0.11	1.00		
RI	-0.06	0.12	-0.02	0.06	**0.58**	1.00	
LC	-0.14	0.15	-0.05	0.06	0.05	0.01	1.00

F10	WL	WA	BV	EL	SL	RI	LC
WL	1.00						
WA	0.19	1.00					
BV	0.30	0.10	1.00				
EL	-0.24	0.21	-0.02	1.00			
SL	-0.13	-0.08	-0.08	-0.16	1.00		
RI	-0.13	-0.11	-0.10	-0.16	**0.86**	1.00	
LC	0.20	0.17	0.13	-0.10	-0.07	-0.07	1.00

3.3.3. Sensitivity of the model to n_{tree}

The choice of n_{tree} and m_{try} can be important for computing the importance values of the predictors (Genuer et al. 2010). The sensitivity of the variable's importance values to n_{tree} based on a single fire event (F01), across R_4, R_8, R_{16}, R_{32}, and R_{64} was explored; hence subsequent variable importance analysis would be performed based on the optimal n_{tree} obtained. The sensitivity of RF to n_{tree} *and* m_{try} was examined based on the OOB error values as shown in Table 3.3 (above). Also, the importance values of the predictors were obtained using different n_{tree} values at m_{try} 3 (Figure 3.6); boxplots are based on the 10 n_{tree} values considered. The behaviour of variable importance was expected to be affected considerably when larger values of n_{tree} were considered; as larger value of n_{tree} provides more stable classification and hence variable importance (Liaw and Wiener 2002; Prasad et al. 2006). However, this study discovered that the effect of taking a larger n_{tree} value was less visible and the variability of variable importance across the n_{tree} values considered was not considerable for most of the predictors (Figure 3.6).

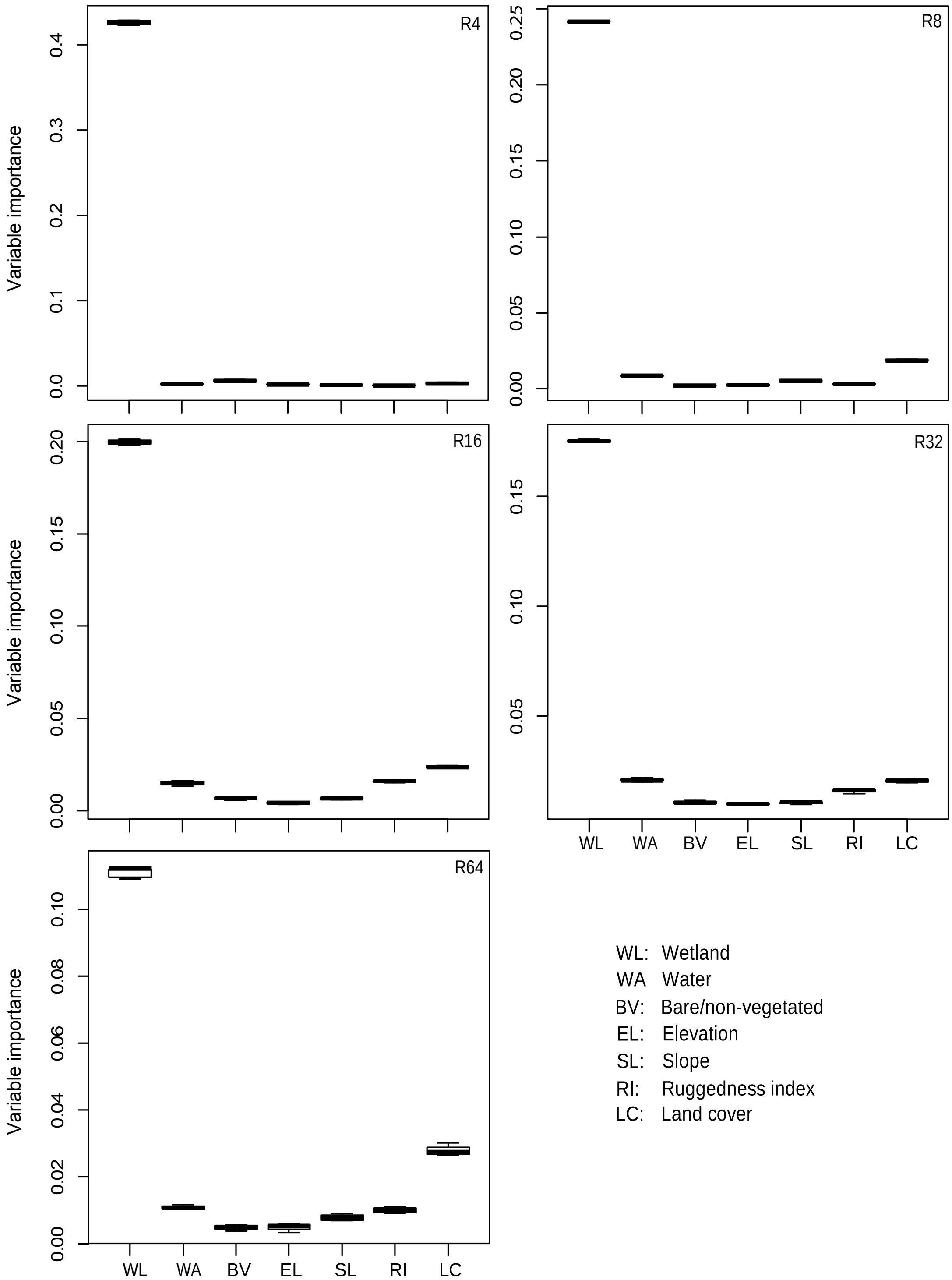


Figure 3.6. Box plots for predictor's importance values (i.e., sensitivity or variability of predictor's importance values to n_{tree}); each box in the plots is based on the importance scores obtained across different n_{tree}). The importance scores are obtained using the permutation accuracy measure shown in Equation 3).

3.3.4. Spatial variable importance of predictors: scale effect

The quantification of a variable's importance is essential for ranking the relative importance of the variables and selecting the predictors that best explain the response variable. The variables were first ranked by sorting their importance values in a descending order and plot the variable importance obtained across the five spatial resolutions against the predictor variables for each fire event. Figure 3.7-

Figure 3.9 show the variable importance scores obtained for each event; boxplots are based on the importance score computed across R_4, R_8, R_{16}, R_{32}, and R_{64}. For a predictor variable to be considered important and informative to explain the response variable, a rule of thumb was introduced by Strobl et al. (2009b). The rule states that a variable can be considered informative and important if its conditional importance value is above the absolute value of the lowest negative scoring variable. The rationale for this rule of thumb is that the importance of irrelevant variables varies randomly around zero (Strobl et al. 2009a). Based on the rule of thumb, the importance scores for most of the predictors except wetland varied around zero, indicating that the predictors were less informative to explain the occurrence of residual patches.

The concentration and distribution of wetlands within the fire perimeter contributed considerably for the existence of residual patches throughout most of the fire events. The results showed that distance to wetlands is the most important predictor variable to explain residual occurrence for most of the fire events (e.g., Figure 3.7-

Figure 3.9), thus supporting the hypothesis on variable importance in relation to wetlands. This was not however the case for some of the fire events (e.g., F06, F07, and F09) where the importance values of all the predictors was close to zero. Similarly, the importance scores of the categorical variable, land cover, were close to 0, suggesting the relative low importance of land cover variable for residual patch occurrence. With the exception of wetland variable which stands out as the most important predictor, the importance values (for all predictors) are close to zero. This may suggest that the existence of residual patches is associated with a complex interaction of geo-environmental factors. Therefore, it was important to evaluate the marginal effect of each of the variables to explain the probability of residual patch occurrence.

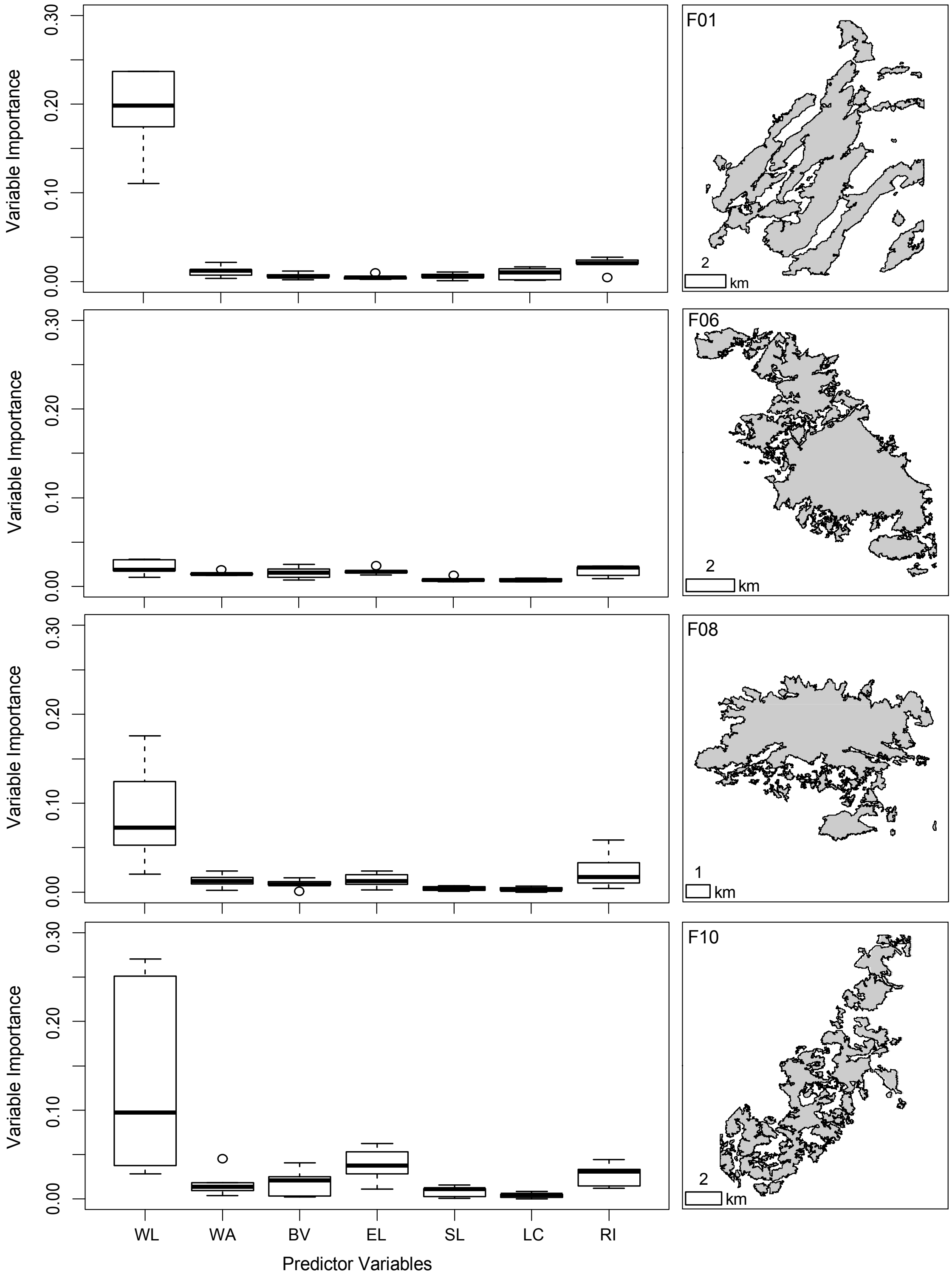


Figure 3.7. Box plots for the relative importance of the predictor variables considered in this study, for the large sized fire events; each box in the plots is based on the importance values computed across different spatial resolutions.

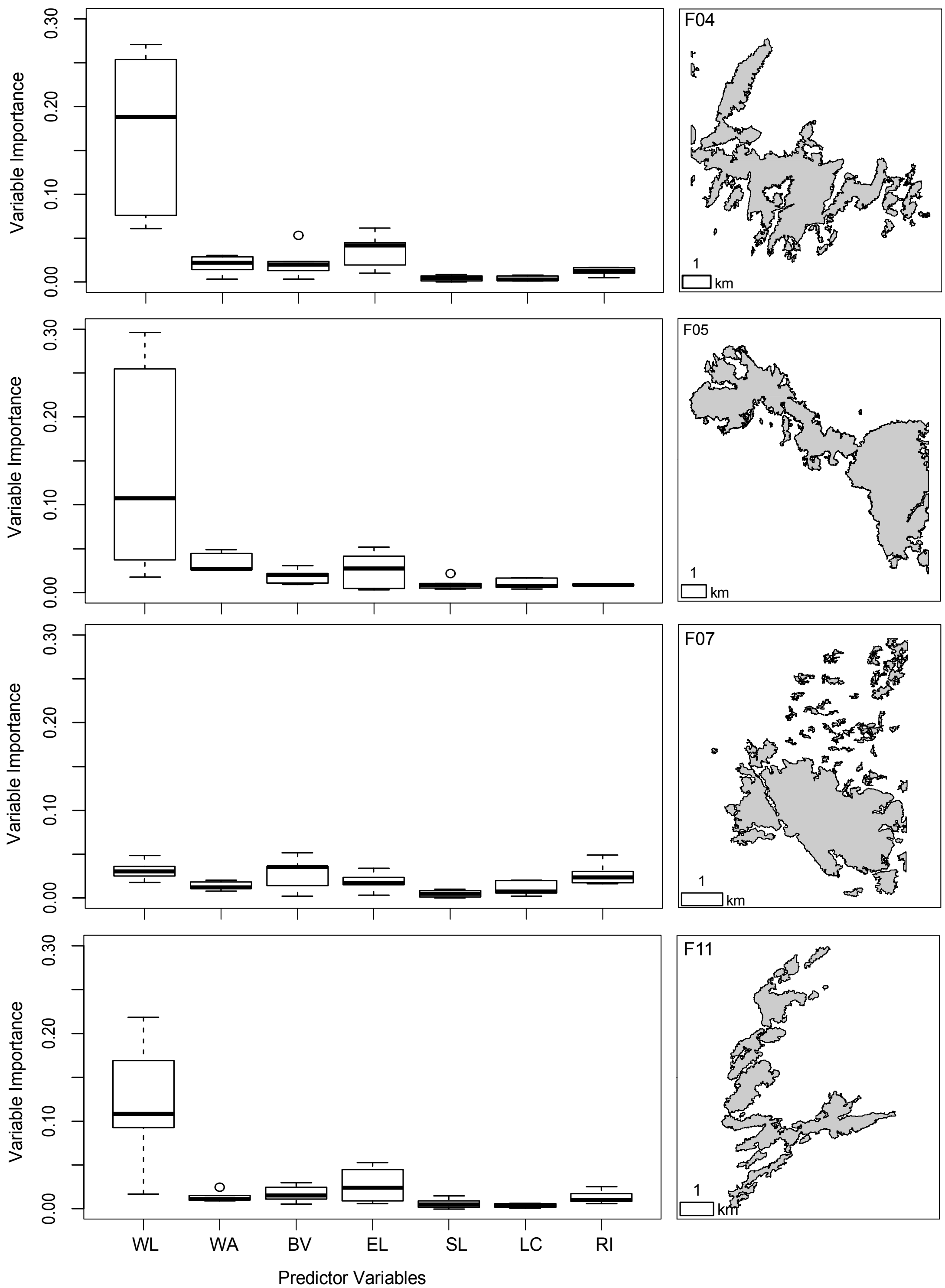


Figure 3.8. Box plots for the relative importance of the predictor variables considered in this study, for the medium sized fire events; each box in the plots is based on the importance values computed across different spatial resolutions.

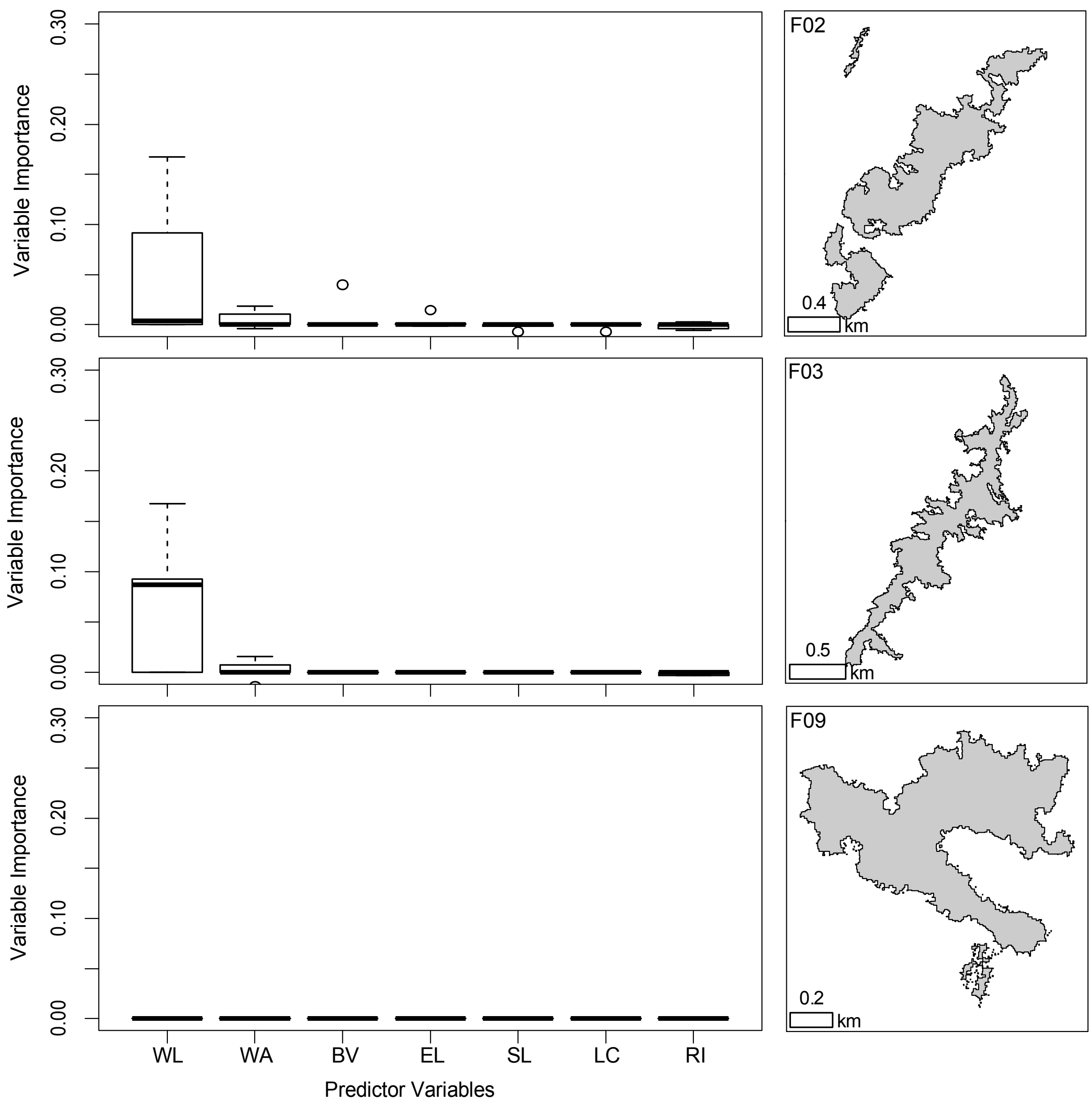


Figure 3.9. Box plots for the relative importance of the predictor variables considered in this study, for the small sized fire events; each box in the plots is based on the importance values computed across different spatial resolutions.

The effect of grain size on variable importance was examined by executing RF and computing the importance values at five spatial resolutions. The figures shown above presented the variability of predictors' importance across five spatial resolutions; boxplots are based on the five spatial resolutions (R_4, R_8, R_{16}, R_{32}, and R_{64}). The variability indicates that the effect of changing grain size on the importance values. Compared with other predictors, the importance of the most important predictor variable (i.e., distance to wetlands) was highly variable. The change in grain size affects not only the variability of the importance values, but also the relative ranking of the predictors.

3.3.5. ANOVA analysis on importance values

A statistical analysis of the measures of variable importance was undertaken using One-Way Analysis of Variance (ANOVA) test. This was applied to determine if the variable importance scores were significantly different based on the predictors. The null hypothesis for the ANOVA test was that the variable importance scores (i.e., mean importance values across five spatial resolutions) for all the predictors are equal while the alternative hypothesis stated that at least one of the means is different from the others. The test was computed for a 95% confidence level (α = 0.05). If there is a significant statistical difference among the means, the null hypothesis is rejected; and hence one of the means is different from the others. However, the test does not tell which groups (variables) are statistically different from one another; it can only tell if there is a difference. In this specific scenario, a post-hoc test is required to provide pair-wise tests of mean differences amongst the groups (variables).

Based on Strobl's rule, all the predictors appeared to explain residual patches, but there has been a slight change in the importance scores among the variables. The result of ANOVA confirmed this view that a significant statistical difference was present amongst the means (relative importance of the variables is different), allowing us to reject the null hypothesis ($p < 0.01$). Therefore, a Tukey post-hoc test was computed to provide pair-wise comparisons of the means (Table 3.7). The table provides the statistical test and *p-values* of the pairs of variables (only between the most important predictor – wetland and the remaining predictors) at 95% confidence interval ($p < 0.05$) for the eleven fire events. The pair-wise test based on the other predictors did not produce a statistically significant difference; indicating that the differences among the other predictors (*BV, EL, LC, RI, SL, and WA*) was not significant. The wetland variable was the most informative predictor for F01, F04, F05, F08, and F10; but its importance scores was close to zero for F06 (Figure 3.7) and F07 (Figure 3.8). Similarly, the post-hoc test output indicates that the differences between wetland and the other variables (except for *R*I and *SL* for F06) were not significant for F06 and F07.

Table 3.7. A post-hoc test based on Tukey test (Bold cells indicate statistical significant difference); *WL*- wetland, *WA*- water, *BV*-bare/non-vegetated, *EL*- elevation, *SL*- slope, *RI*-ruggedness index, and *LC*- land cover

	p-value			
Pairwise variables	**F01**	**F06**	**F08**	**F10**
WL- BV	**0.0000**	0.6018	**0.0004**	**0.0053**
WL- EL	**0.0000**	0.8577	**0.0009**	**0.0293**
WL- LC	**0.0000**	0.9019	**0.0061**	**0.0113**
WL- RI	**0.0000**	**0.0043**	**0.0001**	**0.0014**
WL- SL	**0.0000**	**0.0079**	**0.0001**	**0.0021**
WL- WA	**0.0000**	0.4418	**0.0008**	**0.0051**

	p-value			
Pairwise variables	**F04**	**F05**	**F07**	**F11**
WL- BV	**0.0000**	**0.0066**	0.9984	**0.0001**
WL- EL	**0.0001**	**0.0119**	0.5899	**0.0007**
WL- LC	**0.0000**	**0.0029**	0.9964	**0.0001**
WL- RI	**0.0000**	**0.0033**	0.1189	**0.0000**
WL- SL	**0.0000**	**0.0032**	**0.0182**	**0.0000**
WL- WA	**0.0000**	**0.0235**	0.2396	**0.0001**

	p-value		
Pairwise variables	**F02**	**F03**	**F09**
WL- BV	0.2441	**0.0065**	**0.0062**
WL- EL	0.1480	**0.0065**	**0.0062**
WL- LC	0.0948	**0.0032**	**0.0057**
WL- RI	0.0960	**0.0065**	**0.0062**
WL- SL	0.0931	**0.0065**	**0.0062**
WL- WA	0.1834	**0.0085**	**0.0667**

3.3.6. Marginal effect of predictors on probability of occurrence

According to RF's variable importance measures, I discovered that distance to wetland was the most important predictor for all the fire events except for F06, F07, and F09 while the importance values of most of the predictors were close to 0. Additionally, it has been stated that RF is not a tool for traditional statistical inference, and the variable importance measures in RF has been used to subjectively identify relatively important predictor variables for interpretation (Cutler et al. 2007). Therefore, it was useful to investigate the relative effect of each the variables for predicting the response variable. The marginal effect of the predictor variables on class prediction (i.e., residual patches probability of occurrence) was studied using a PDP. The plots that summarize the contribution of selected predictors to class probability (on the largest fire events: F01, F06, F08, and F10) are shown in Figure 3.10- Figure 3.13. For a binary classification (i.e., the presence or absence of residual patches), the *y*-axis on partial dependence plots is presented in logit scale (Equation 7). The figures show not only the marginal effect of each predictor variable, but also the effect of changing grain size on partial dependence of a response variable on the explanatory variables.

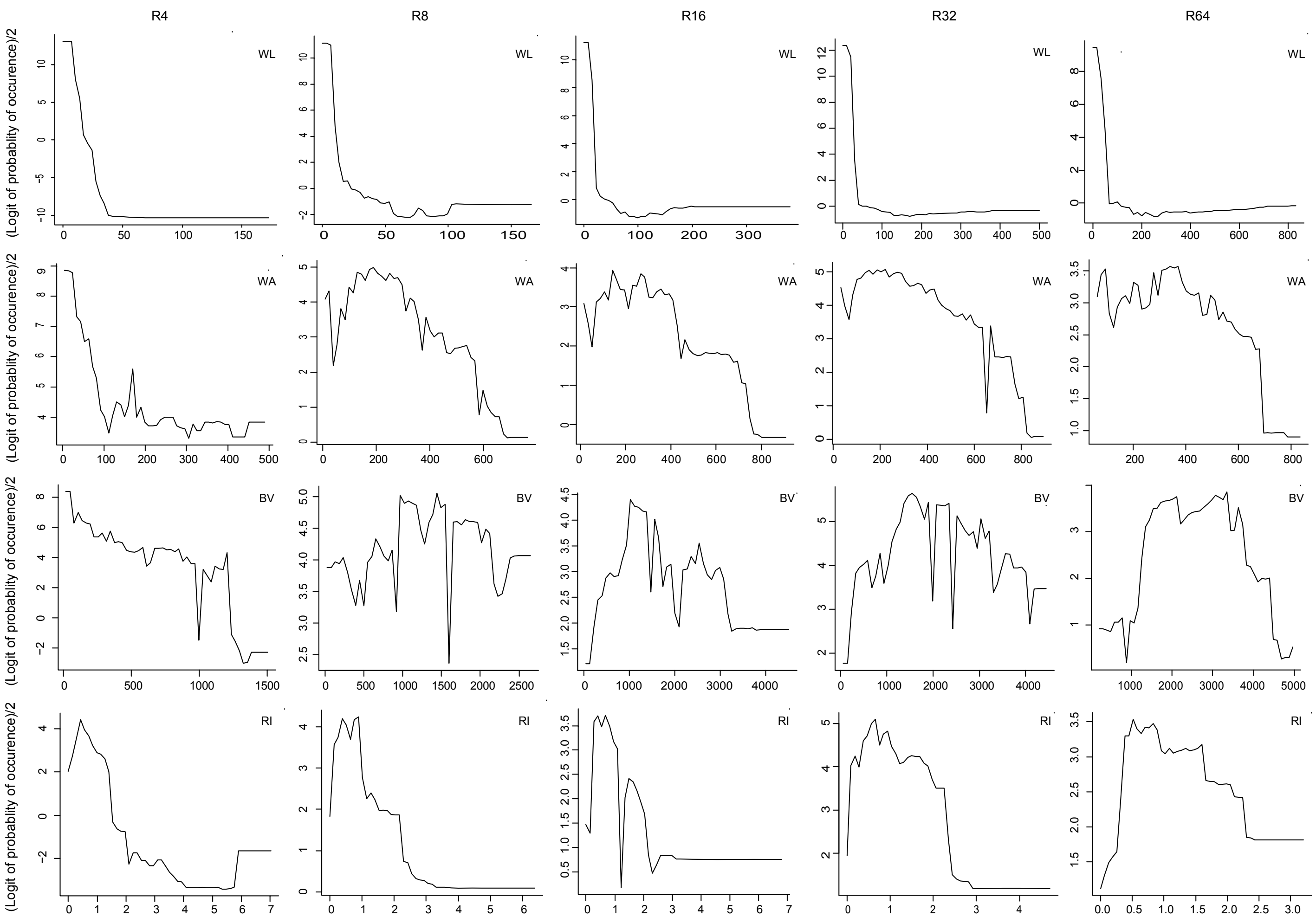


Figure 3.10. Partial dependence plots for selected predictor variables for random forest predictions of the presence of residual patches for F01 at R_4, R_8, R_{16}, R_{32}, and R_{64}. Partial dependency is the dependence of the probability of presence on one predictor variable after averaging out the effects of the other predictor variable in the model. The x-axis of each plot indicates the explanatory variables (WL – distance to wetland, WA – distance to surface water, BV – distance to bedrock and non-vegetated areas, and RI – ruggedness index) while the y-axis is a half of the probability of occurrence (Equation 7).

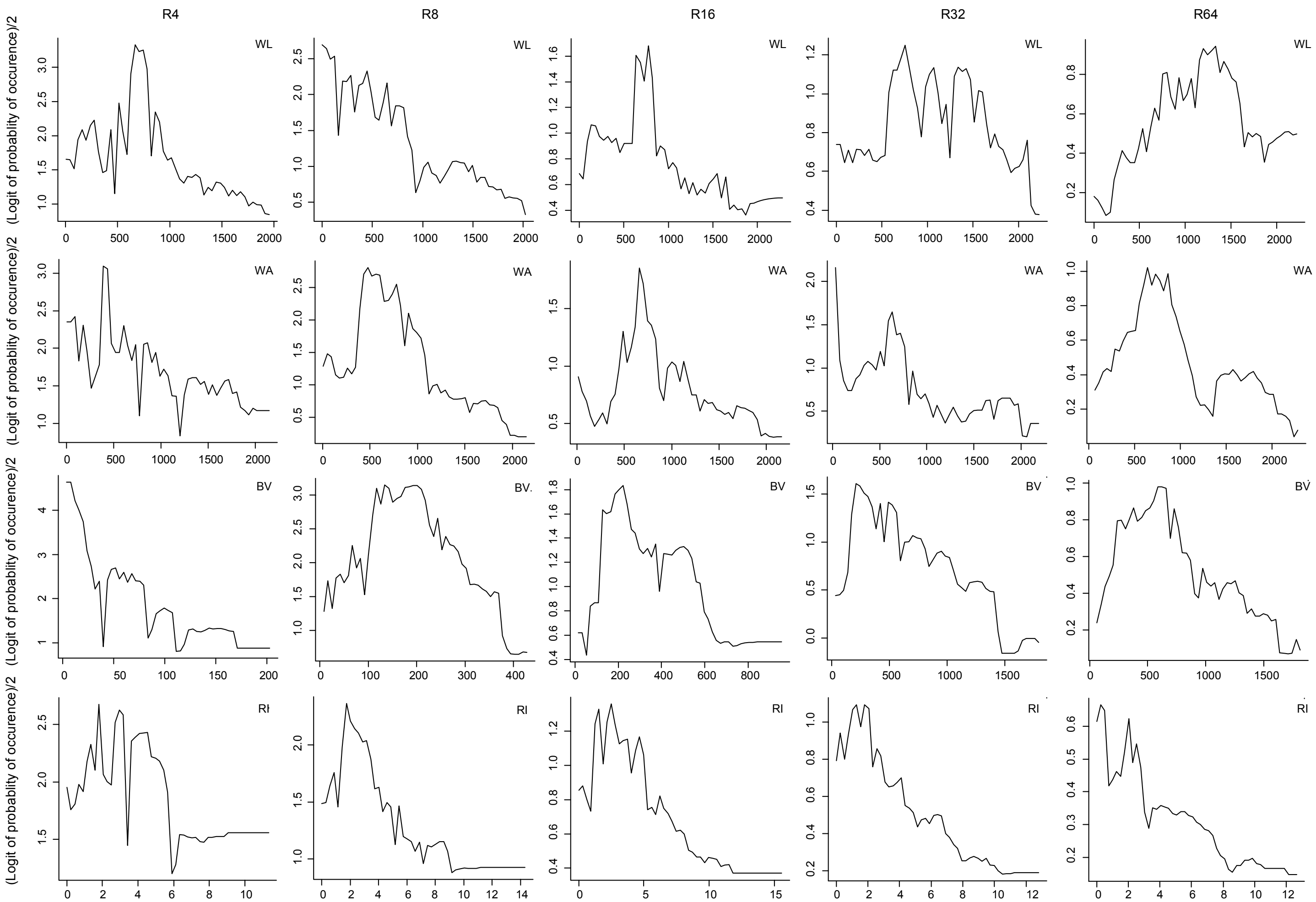


Figure 3.11. Partial dependence plots for selected predictor variables for random forest predictions of the presence of residual patches for F06 at R_4, R_8, R_{16}, R_{32}, and R_{64}. Partial dependency is the dependence of the probability of presence on one predictor variable after averaging out the effects of the other predictor variable in the model. The x-axis of each plot indicates the explanatory variables (WL – distance to wetland, WA – distance to surface water, BV – distance to bedrock and non-vegetated areas, and RI – ruggedness index) while the y-axis is a half of the probability of occurrence (Equation 7).

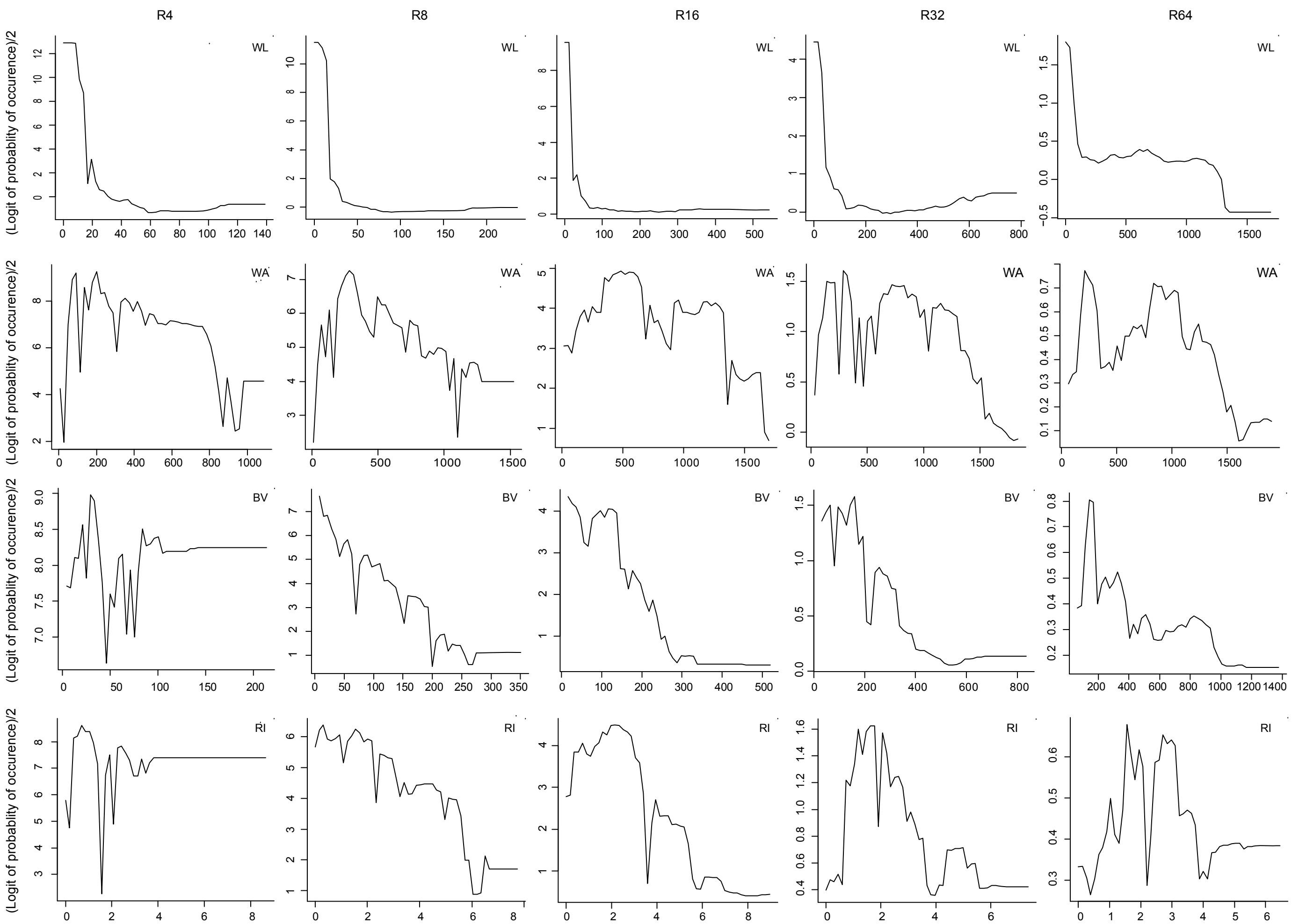


Figure 3.12. Partial dependence plots for selected predictor variables for random forest predictions of the presence of residual patches for F08 at R_4, R_8, R_{16}, R_{32}, and R_{64}. Partial dependency is the dependence of the probability of presence on one predictor variable after averaging out the effects of the other predictor variable in the model. The x-axis of each plot indicates the explanatory variables (WL – distance to wetland, WA – distance to surface water, BV – distance to bedrock and non-vegetated areas, and RI – ruggedness index) while the y-axis is a half of the probability of occurrence (Equation 7).

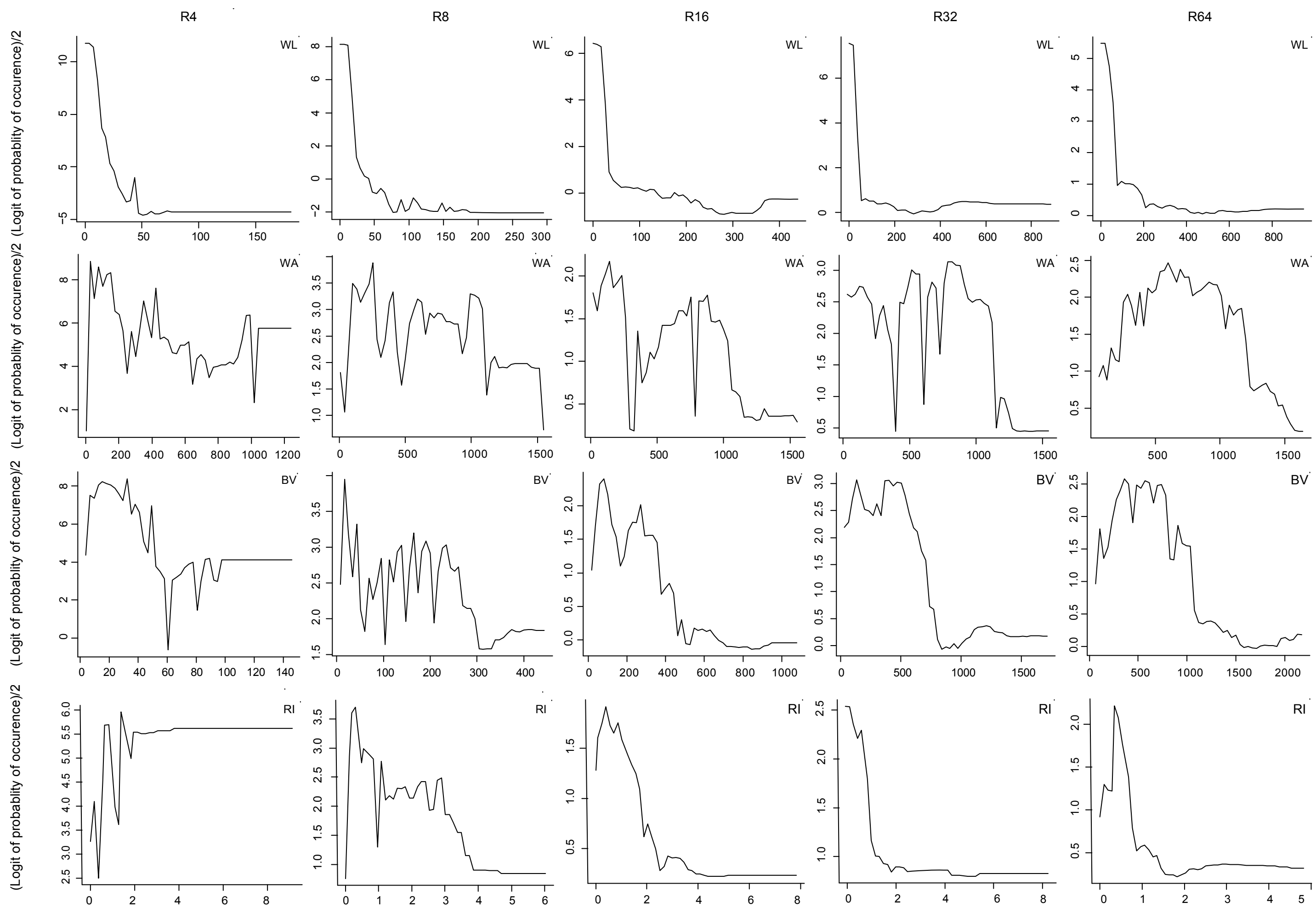


Figure 3.13. Partial dependence plots for selected predictor variables for random forest predictions of the presence of residual patches for F10 at R_4, R_8, R_{16}, R_{32}, and R_{64}. Partial dependency is the dependence of the probability of presence on one predictor variable after averaging out the effects of the other predictor variable in the model. The x-axis of each plot indicates the explanatory variables (WL – distance to wetland, WA – distance to surface water, BV – distance to bedrock and non-vegetated areas, and RI – ruggedness index) while the y-axis is a half of the probability of occurrence (Equation 7).

3.4. Discussion

3.4.1. Residual and null-residual patches

Owing to the presence of large residual patches and unburnable areas within the fire footprint, the random simulation was not able to completely mimic the existing residual patches. The total number of residual patches simulated across the five spatial resolutions was different; this is attributed to the availability of suitable space within the fire perimeter. The aggregation method (independent as opposed to 'iterative' and majority rule as opposed to random rule-based method) used to resample the classified image might have also contributed in the class imbalance between the number of residual and null-residual patches. Studies have shown that different aggregation methods may have considerable effect on patterns of landscape structure (Wichham et al. 1995, He et al. 2002; Wu et al. 2002). For example, for landscapes with greater local heterogeneity, the results of aggregation method might be less adequate (Benson and MacKenzie 1995). On the other hand, important properties of landscapes are not preserved under some aggregation methods (Wu et al. 2002).

3.4.2. Spatial variable importance of predictors

For improving forest management, harvest planning and emulating natural disturbances, sustainable forest management practices in Ontario must consider disturbances as necessary agents of change, not as elements to be excluded entirely. The NDPE guide also provides standards and guidelines to emulate fire disturbances and determine the type and proportion of residual patches to be retained during harvesting. This study was part of an effort to develop a consistent and repeatable method to improve our understanding of the patterns and characteristics of post-fire landscape structure. The study addressed the relative importance of the predictor variables, and their marginal effect for residual patch occurrence in a given landscape.

The tendency for a fire to spread and for a residual patch to occur is influenced by several factors (e.g., terrain, natural firebreaks, vegetation types, pre-fire characteristics and weather conditions). However, in data mining applications, the input explanatory variables are seldom equally relevant to explain the response variable (Hastie et al. 2009); there is a tendency for certain predictors to become more informative than others. Similarly, I discovered that certain predictor variables (e.g., wetlands) were more informative to separate residual patches from the null-residual patches while topographic and land cover variables were less important.

The impact of terrain related variables on fire behaviour and the occurrence of residual patches have been documented in various studies (e.g., Viegas 1993; ESA 2002; van Wagtendonk 2004; Linn et al. 2007; Meddens et al. 2008; Cuesta et al. 2009; Madoui et al. 2010). The terrain related variables play an important role in the occurrence and patterns of residual patches. Despite this, the results of this study revealed that the topographic factors (slope, elevation, and RI) were among the least important predictors to explain residual patches across most of the fire events. This does not actually reflect the hypothesis that topographic variables would be informative for residual patch occurrence. Based on the DEM used in the study and the associated parameters (e.g., RI), the sites are characterized by relatively flat relief. The RI category was initiated by Riley et al. (1999) to quantify and categorize topographic heterogeneity into seven classes (Table 3.8).

Table 3.8. Classification scheme for a Terrain Ruggedness Index (RI) that quantifies topographic heterogeneity (Riley et al. 1999). Based on the RI values computed, a terrain is broadly categorised to one of the classification schemes.

RI Category	**RI values in m**
Level	0 - 80
Nearly level	81 -116
Slightly rugged	117 - 161
Intermediately	162 – 239
Moderately rugged	240 – 497
Highly rugged	498 – 958
Extremely rugged	958 – 4367

In this study, the RI values derived from the DEM ranges from 0 m to 38 m (Figure 3.14), indicating that the sites are within the first category of the Riley's classification where the topography is considered to be a level. However, we should be aware of the potential biases originated in DEM where RI values are computed and interpreted (Riley et al. 1999). The spatial resolution of the DEM available for the study area was coarse compared with the elevation data available for the southern Ontario where a 10 m spatial resolution DEM is available. The spatial resolution of the DEM varies from a minimum 0.75 arc seconds (~ 20 m) along a profile in the south-north direction and to a maximum 3 arc seconds (~ 90 m) in the east-west direction. Since the relief in the region is characterized by flat plains, undulating uplands areas, and dissected uplands with ridges and escarpments (Baldwin et al. 2000), having a much higher spatial resolution DEM can have an impact on the analysis.

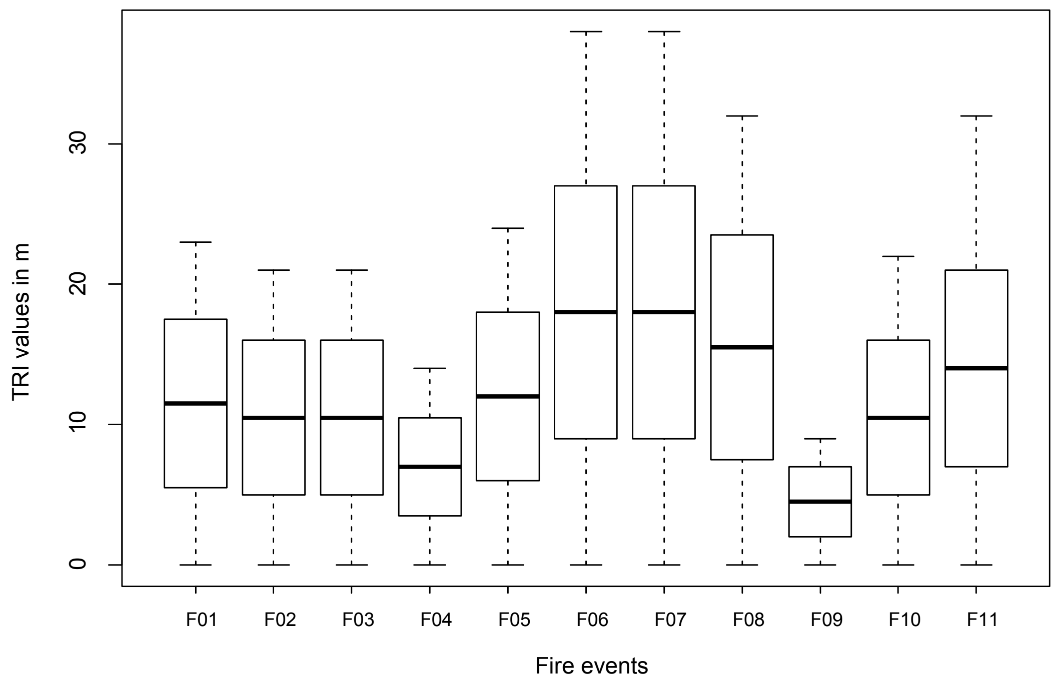


Figure 3.14. A box plot that shows the variability of the RI values computed for each fire event; each box in the plot is based on the RI values computed at different spatial resolutions.

Furthermore, several studies have indicated that the importance of natural firebreaks (e.g., surface water and wetlands) for the occurrence and distribution of residual patches (Turner et al. 1997; Perera et al. 2007; Cuesta et al. 2009). Specifically, the occurrence of residual patches in relation to proximity to surface water was investigated in (Madoui et al. 2010; Dragotescu and Kneeshaw 2012). In a study conducted by Dragotescu and Kneeshaw (2012) residual patches tended to be concentrated in closer proximity to surface water and there have been a uniform distribution of residual patches near surface water, specifically near lakes. Similarly, in this study, it was expected that predictor variables related to natural firebreaks (e.g., surface water, wetland, bedrock and non-vegetated areas) would have high importance values. Despite their importance to residual patch abundance and distribution, the study demonstrated that distance from surface water was less informative to explain the post-fire residual patches compared to other natural firebreak features (e.g., wetlands). Similarly, the importance values for the other natural firebreak feature (i.e., distance to non-vegetated lands) were close to zero for most of the fire events; this is despite the anticipation that the variable would be relevant to explain the residual patches. The small importance values associated with distance to water and non-vegetated areas is likely because the proportion of water and non-vegetated lands within the fire events was not substantial.

Considering the relative ranking of the importance scores, I discovered that distance to wetland was the most important predictor variable across the fire events with few exceptions (F06, F07, and F09). In a similar study by Madoui et al. (2010) the existence of residual patches did not associate with wetlands. The low importance values associated with the wetland could be associated with the abundance and distributions of wetland in the area. This was evident in F06 and F07 where the proportion and distribution of wetlands within the fire events was not considerable (less than 1%); besides the existing wetlands are concentrated on the periphery (and outside) of the fire perimeter or surface water.

Studies have indicated that the occurrence of residual patches is often associated with the dominant land cover types in the landscape (Kafka et al. 2001; Burton et al. 2008). However, there are some cover types such as deciduous forest and sparse conifer that dominate the existence of residual patches despite their low abundance in the landscape (Kafka et al. 2001; Madoui et al. 2010). The results from the previous chapter showed that the abundance of residual patches in the eleven fire events is attributed not only to land cover types with high abundance in the landscape (e.g., sparse and dense conifer) but also to land cover types that are less prone to high severity fire (e.g., treed wetland) (§0). This depends on the fuel condition and moisture content of cover types. Besides, the abundance of diverse land cover types in the landscape was expected to maximize the importance scores of the land cover variable. However, the importance scores for the land cover variable were close to zero across all the fire events, suggesting that land cover variable was less informative to explain the overall residual patch occurrence in the sites.

The effects of changing grain sizes on characterizing the patterns of post-fire landscape structure have been addressed in the previous chapter. The study concluded that the sensitivity to scale change varied greatly among the metrics used to characterize the patterns of residual patches and spatially across the fire events. Similarly, the effect of grain size was assessed in relation to the importance scores of the predictor variables used to explain the residual patches. The study indicated that the importance values varied across the five spatial resolutions and there has been a change in relative ranking of the predictors. Despite the variability in importance values and a change in relative ranking of some of the variables, natural firebreak features (distance to wetlands) remain the most important predictor variable across all fire events except for F06, F07, and F09. The slight change in predictors' importance ranking reinforces the idea that the parameters that explain the residual patches should be examined at each spatial resolution. Thus, it is not easy to develop a simple scaling rule to predict variable importance values across the gradient of scales. The scale effect on predictors' importance values (and rankings) can eventually affect variable selection. Finally, it is important to note that the variable importance measure considered in this study is interpreted as a relative ranking of significant predictors. The absolute values of the importance values should not be interpreted or compared over different studies (Strobl et al. 2009a); the assessment of a predictor's importance should be site specific.

3.4.3. Marginal effect of predictors on probability of occurrence

The marginal effects of the four predictors revealed different interesting patterns. First, the relationships between individual predictor variable and probability of residual presence are nonlinear; the plots showed a monotonic or erratic relationship. Second, there appears to be a distance decay effect, which describes the effect of distance on spatial distribution and interactions. The distance decay effect reflects how diversity is spatially distributed and states that the interaction between two locations declines as the geographic distance that separates them increases (Morlon et al. 2008). Geographic distance is the most important parameter that affects the diversity of a community in a landscape. The results reflected the distance-decay function because the interaction between natural firebreak features (wetlands and water) and residual patches declined as the distance between them increases. Third, the marginal effects of the most important variable (i.e., WL) exhibit a decreasing monotonic trend for some of the fire events (F01, F08, and F10); this is in spite of the differences in the logit scale. For the fire events considered in the study, the probability of residual occurrence increases monotonically with increasing distance from wetlands but levelled off at a specific distance, depending on the grain size (e.g., in Figure 11, it levels off at 40 m for R_4). This supports the idea that high density or concentration of residual patches is associated with closer proximity to wetlands. It also supports the view that areas with a higher moisture regime, such as wetlands have the potential to retain post-fire residual patches (Cuesta et al. 2009; Dragotescu and Kneeshaw 2012). In a study undertaken by Madoui et al. (2009) there was a failure to find a relationship between residual patch occurrence and wetland existence, suggesting the environment was part of the fuel.

Water courses in the form of lakes and rivers are often considered to be barriers for fire spread. Perera et al. (2009b) examined residual vegetation in proximity to water courses in which the occurrence of residual patches did not associate with the proximity to surface water and wetlands. In this study, lakes of different size and rivers exist within the studied fire events, although the abundance and distribution varies across the fire events. The response of residual occurrence in relation to other natural firebreak parameter (i.e., WA) indicated that the existence of residual patches occur more in closer proximity to surface water and decreases with increasing distance from water. In a similar study, Dragotescu and Kneeshaw (2012) found a relatively uniform distribution of residual patches in closer proximity to surface water. Yet, some erratic patterns are evident from the plots that show the marginal effect of WA. This is in agreement with Madoui et al. (2009) findings where most of the fires studied in the boreal forest of Western Quebec show no evidence of spatial association with surface water, although in some of the fires the presence of residual patches were associated with the proximity of water. There are many possible reasons why the results did not exclusively reflect the claim that residual patterns are likely to occur in closer proximity to surface water; example differences in characteristics of fires studied and the abundance and distribution of surface water in the area. Another plausible reason could be pertained to the shape of the water courses. For example, Dragotescu and Kneeshaw (2012) discovered a tendency for residual patches to be concentrated near meandering rivers than linear ones.

Fourth, the abundance and spatial distribution residual are often associated with rugged terrain and residual patches are less likely to occur in flat terrain. Dragotescu and Kneeshaw (2012) observed that there are almost no residual patches on relatively flat areas. The findings of this study are not in agreement with this, as residual patches tended to exist in relatively flat areas. One of the plausible explanations is related to the coarse spatial resolution of the DEM used to generate the topographic variables. Based on the ruggedness index computed, the sites are located in a relatively flat topography, but this may not give a comprehensive description on the impact of terrain. Fifth, the effect of scale on the predicted probability of residual occurrence is apparent across the plots. While there was a slight trend for some of the predictors (e.g., *WL*), the magnitude of the probability plots varied for all the predictors with changing grain size. The non-predictable and erratic partial plots, across different grain sizes, reflect that it is not easy to develop a simple scaling law to predict patterns at different scales. In summary, the partial responses for residual occurrence for the most influential variables (i.e., *WL*) demonstrated that residual patches occurred closer to wetlands (e.g., within 100 m from the wetland area for F01 at R_{64}). The variation in the magnitude and general trends of the plots also indicates that the occurrence of residual patches is attributed to various geo-environmental factors that interactively affect their existence. It is important to note that partial plots can show general trends, but may not provide a comprehensive description (Friedman 2009). It was necessary to investigate the combined effects of the predictor variables for residual patch occurrence; hence the next chapter (Chapter 4) examines the combined effect of the predictor variables to explain the occurrence and distribution of residual patches using a predictive modelling approach.

3.5. Summary and conclusions

The heterogeneous characteristics of post-fire landscape structure have been examined in relation to various factors including the abundance of natural firebreak features, topographic features, and other pre-fire forest characteristics (such as land cover types). An ensemble learning method (RF) was used to explore the relative importance and marginal effects of the various geo-environmental variables that explain the existence of residual patches within 11 boreal fire events. This part of an effort to develop a consist approach for examining the variables that best explain the existence of residual patches, and accordingly informing forest managers the potential of the techniques developed in this study for subsequent forest management practices.

One of the objectives of this chapter was to assess the factors that explain the existence of residual patches and the sensitivity of the importance scores to scale change. Based on the findings of this study, the following conclusions are drawn. First, in data mining applications, predictor variables are seldom equally important; certain predictors are more discriminant than others. I discovered that certain predictor variables (specifically distance to wetlands) were the most important predictor to separate residual patches from the null-residual patches. This was true across most of the fire events except for F06, F07, and F09 where the relative importance of all the variables varied close to zero. The high importance values of the dominant predictor (distance to wetlands) are associated with their relative abundance and distribution within the burned landscapes. Second, topographic variables are a contributing factor in shaping the patterns of post-fire landscape structure, specifically in the formation of residual patches. Based on the RI scheme computed in this study, the study sites were generally classified as nearly level; hence topographic variables used in the study were found to be less informative to explain the presence of residual patches. However, the local topographic variability within the burned landscape may not have been reflected in the study due to the coarse spatial resolution of the DEM used in the study. Third, owing to the distinctive spatial patterns of landscapes at different scales, a single scale description of landscape patterns may provide partial or misleading information. The effects of analytical scale (i.e., spatial resolution) on determining the importance of each of the predictor' variables were assessed; the importance of the variables was examined at (R_4, R_8, R_{16}, R_{32}, and R_{64}). Changing the grain size affects the importance values and relative ranking of the predictors; it can be inferred that the configuration of patch characteristics were sensitive to changing grain size. The sensitivity to scale change provides an opportunity to understand the multiple-scale characteristics of a given landscape. The study noted that assessing the importance of a predictor should be site specific (i.e., fire event level in this context).

Random Forest is not a tool for traditional statistical inference, and the variable importance measure in RF has been used to identify predictor variable for interpretations. The study was also investigated the marginal effect of the predictor variables on the occurrence of residual patches and it was inferred that: 1) the relationships between individual predictor variable and probability of residual presence were non-linear; 2) the marginal effect of the most important predictor (WL) exhibited a decreasing trend; the occurrence of residual patches within a disturbed landscaped tended to occur more in closer proximity to natural firebreak features, specifically to wetlands. Although the probability of occurrence varied with the spatial resolutions considered, residual patch occurrence were prevalent within 100 m from the wetlands; and 3) the effect of scale on the predicted probability of residual occurrence was visible. The erratic patterns of the plots and the variation in the magnitude of the plots with changing grain size make deriving a simple scaling rule difficult.

The importance scores computed using permutation accuracy provides a relative importance of an individual variable; they may not provide a comprehensive description of the overall effect of the predictors. Similarly, the PDP only shows how the response variable changes as you change the predictor variable; it may not provide a comprehensive description of a variable's effect on the prediction. Despite this, the results of this study indicate that RF is a repeatable and broadly application approach for determining the relative importance of various factors and assessing their marginal effect on predicting the response variables. Although it is beyond the scope of this study, the approach integrated in this chapter can also paved a way for forest managers and policymakers to consider the method implemented in this chapter for implementing a disturbance-based forest management practices as the approach allows policymakers and forest managers to answer the question 'where to retain a residual patch?' while implementing forest harvesting operations.

4. Spatially explicit model predicting residual vegetation patch existences within boreal wildfires

Abstract

Wildfires are frequent boreal forest disturbances in Ontario and emulating them with forest harvesting has emerged as a common forest management goal. Since wildfires typically contain a considerable number of unburned residual patches of various size, shape, and composition, the study presents means for learning their characteristics to improve the subsequent emulation of wildfires with forest harvest planning. A method for developing probability maps for the existence of residual vegetation within wildfire dominated landscapes is presented. The study uses the Random Forests ensemble learning approach to predict the occurrence and distribution of residual vegetation patches based on selected predictor variables: proximity to wetlands, surface water, old burns, or non-vegetated areas, in conjunction with site characteristics comprising slope, elevation, a ruggedness index, and land cover type. Satellite derived data for 11 fire events is partitioned into training and validation data using a hold-out validation approach; with data records from a single fire event is used for validation while data records from the remaining 10 fire events are used to construct and calibrate the model. The predictive power of the model is examined using a fixed-probability threshold and threshold-independent measures of model performance. The predictive performance of the model ranges from “strong” model at R_4 to “marginal” model at R_8, R_{16}, R_{32}, and R_{64}, with high (and reasonable) discrimination ability for one of the largest fire events (F01); yet low prediction accuracy (“weak” model) is exhibited for another large fire event (F06). The lowest predictive performance is observed for the smallest fire events (F02, F03, and F09).

Keywords: fire disturbance, residual and null-residual patches, predictor variables, random forest, predictive models, predictive performance, and spatial prediction

4.1. Introduction

4.1.1. Spatial language: residual and null-residual patches

The abundance of residual patches within wildfires and understanding their patterns is a central theme for emulating fire disturbances with harvesting practices. The use of occurrence models for such forest resource management practice reflects the accessibility of presence-absence data and digital (spatial) information pertained to various environmental variables such as land cover types, natural firebreak features, and topographic variables (Nielsen et al. 2005). The presence-absence data in this study are referring to the presence and absence of residual patches, described as residual and null-residual patches respectively. The presence-data (residual patches) were extracted from existing classified Ikonos images while the absence-data (null-residual patches) were derived using a simulation procedure. The approaches used to extract the residual and null-residual patches are presented in §3.2.2.

4.1.2. Predictive distribution Modelling

Knowledge about the geo-environmental factors that define the patterns of post-fire landscape structure has been a central issue in understanding forested landscape and fire disturbances (Guisan and Zimmermann 2000; Manel et al. 2001). Accordingly, the previous chapter assessed the relative importance and marginal effects of different predictor variables that explain residual vegetation patches. This enabled me to investigate the relative importance of an individual variable but it may not provide a comprehensive description of the overall effects of the predictors. Therefore, there was a need to discover the combined effects of the variables using predictive modelling approaches. The term predictive (distribution) model has been used to describe probabilistic models that explain the relationship between the occurrence and distribution of species (spatial elements) and a set of environmental variables (Jepsen 2004; Magness et al. 2009). It often employs statistical and machine-learning techniques to unravel the complex interactions between a species distribution and environmental variables (Munoz and Felicisimo 2004) and produce spatially explicit predictive maps (Beauvais et al. 2006). In this instance, a wide range of spatially explicit predictive models are currently in use to predict and assess the distribution of various spatial elements.

For example, a model has been developed to predict habitat suitability (Fielding and Bell 1997); fish species distribution (Olden et al. 2002), plant biodiversity, forest composition, and structural diversity (Frescino et al. 2001), plant functional type (Pearce and Ferrier 2000; Zaniewski et al. 2002; Anderson et al. 2003). All models used for spatial predictions integrate three components: the assumed theory, the type of data used and the ways in which the data are collected, and the statistical methods (and theory) applied (Austin 2007). The theory is based on the premises that species distributions can be predicted from mapped environmental variables (Beauvais et al. 2006) while the statistical relationship between a response and a number of explanatory variables is also generally specified by regression models (Guisan and Zimmerman 2000; Jepsen 2004). However, some of the models are parametric in which they assume a Gaussian relationship or linear relationships (Frescino et al. 2001; Munoz and Felicisimo 2004). In parametric models such as generalized linear models (GLM) the relationship between response and predictors are assumed to be linear (Munoz and Felicisimo 2004) while real world effects often exhibit complex spatial dependency and non-linear relationships (Evans and Cushman 2009). An alternate approach (i.e., non-parametric methods) that captures the complex and non-linear relationship was thus required (Hastie et al. 2001; Austin 2002). Some of the non-parametric models that gained popularity in spatial predictions include generalized additive model (GAM) (Frescino et al. 2001), CART (Iverson and Persad 1998), and RF (Prasad et al. 2006). Several studies have also been undertaken to assess and compare the predictive capabilities of different statistical and rule-based methods (Pearce and Ferrier 2000; Wilfried et al. 2003; Munoz and Felicisimo 2004). These studies indicate that the results of the modelling depend on 1) the adequacy, quality, and spatial resolution of environmental data, 2) mathematical procedures (formulations), 3) the scale at which analyses are undertaken (Wilfried et al. 2003), 4) the number of input variables, and 5) expert knowledge (Peters et al. 2007).

4.1.3. Random Forests

Random Forests (RF) are among the most widely known predictive models that have been used in distribution modelling. The predictive modelling based on RF (also known as machine learning) aims to generate the most accurate estimates of some quantity or event. Such models are not generally meant to be descriptive and are usually not well-suited for inference; they are rather probabilistic or stochastic methods. A predictive model based on RF constructs a classification and regression tree by successfully splitting data based on single predictors. Each binary node, split, forms a branch in the decision tree and trees are grown without pruning. The model utilizes bagging, a technique that builds a large number of trees and averages the output. The prediction with RF algorithm begins with the selection of many bootstrap samples from the data, and for each bootstrap sample it generates a classification tree (Figure 4.1). A spatial prediction is produced for each classification tree and it eventually builds an ensemble of CART tree classification predictors using a majority vote for final prediction (Magness et al. 2008). A detailed description of RF is provided in §3.1.2.

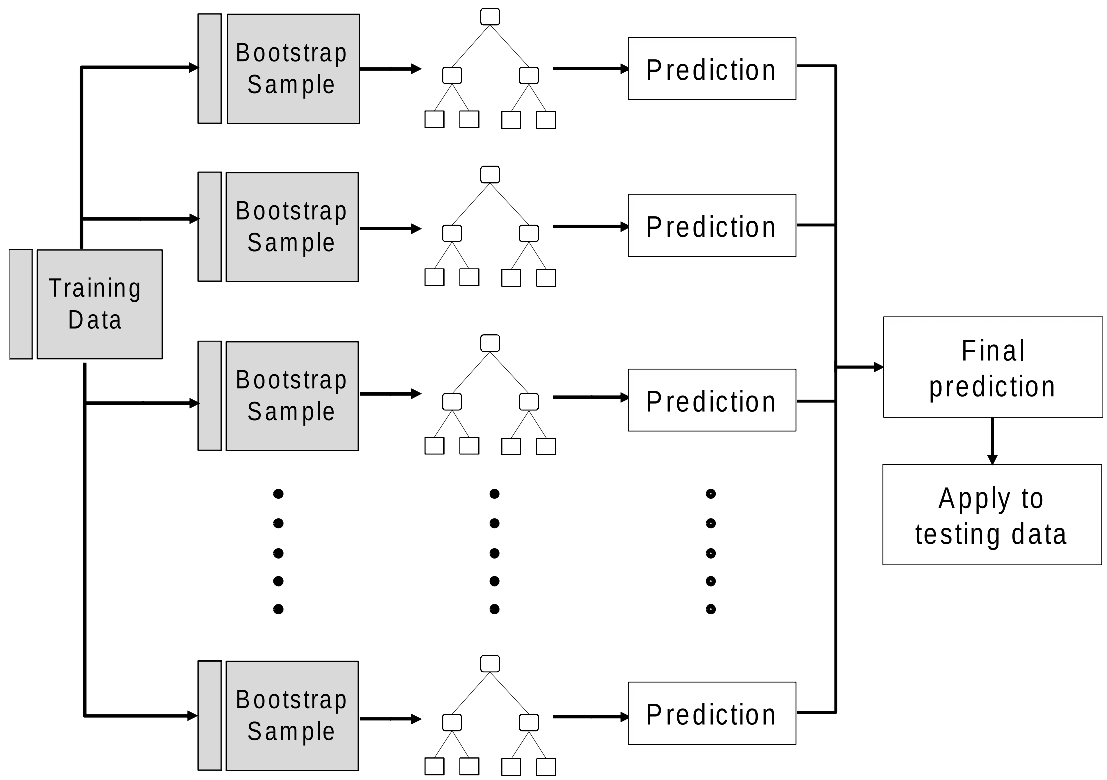


Figure 4.1. Graphical depiction of RF model implementation,1) the model generates multiple classification trees for each bootstrap sample of the training data, and 2) it combines the results from multiple models for final prediction using a majority voting.

4.1.4. Research framework

Several studies have documented the presence of residual patches within a fire disturbed landscape, and few studies have also speculated on agents behind their existences. A wide range of studies (e.g., Vera 2001; Ryan 2002; Perera et al. 2007; Cuesta et al. 2009) have also been undertaken to assess the parameters that govern the occurrence of post-fire residual patches using different statistical techniques. The studies indicated that the complex interactions of various environmental factors acting at different scale could determine the existence of residual patches. Eberhart and Woodard (1987), for example, indicated that the probability of residual patch occurrence increases with the abundance of natural firebreak and topographic features. Epting and Verbyla (2005) suggested that the variability in fire weather, elevation, and fuel conditions can affect the patterns of post-fire forest characteristics. Variations in wind direction

(Rowe and Scotter 1973; Madoui et al. 2009) and presence of certain land cover types (Mermoz et al. 2005; Madoui et al. 2009) can also affect the occurrence of unburned areas. Similarly, Dragotescu and Kneeshaw (2012) found that residual patches are associated with proximity to surface waters and other physiographic features. However, most of these studies did not consider the combined effects of the variables which would interactively affect the occurrence and distribution of post-fire residual patches.

The combined effects of the various environmental variables can be evaluated using predictive modelling approaches (Jepsen 2004); hence potential areas where residual patches are likely to occur can be identified. However, little has been directed at assessing the combined effects of the geo-environmental factors and the predictive ability of statistical modelling techniques for residual patch occurrence. In this chapter, a non-parametric tree-based learning technique (RF) was used to develop a spatially explicit model, and predict the likely occurrence of residual patches by integrating topographic, natural firebreak features, and land cover variables. Moreover, the configuration of post-fire patch characteristics, and the magnitude and variability of the environmental factors are sensitive to a change in spatial resolution (Wiens 1989). Yet, there is a considerable uncertainty regarding the appropriate scale at which analyses are undertaken. In a predictive model, the scale at which prediction models are performed depends on the model structure, the purpose of the model outputs, and the spatial resolution of the environmental data (Wilfried et al. 2003). Because of this scale multiplicity, scale holds the key to understanding the predictive performance of models.

In this chapter, a non-parametric predictive model (RF) was applied to address the following objectives: 1) develop a spatially explicit model for predicting residual patch occurrence within fire events, 2) assess how well the model predicts the likelihood of residual patch occurrence across different fire events within the same ecoregion, as determined by measures of model performance, 3) produce spatially explicit probability maps that show the potential areas where residual patches are likely to occur, and 4) determine the discriminative ability (predictive accuracy) of the model as function of scale (i.e., sensitivity of the model to changing grain size). The study was based on the assumption that 1) the patterns of residual patches are explained by the environmental data, and 2) the response variable (i.e., presence-absence of residual patches) is related to the predictors in a non-linear fashion, and non-parametric models are suitable under such hypotheses. The novelty of this study is the use of RF predictive model for understanding and predicting the occurrence of post-fire residual patches in boreal wildfires, as far as the study is concerned.

4.2. Methods

4.2.1. Study area

A method that generates probability maps for the existence of residual patches within wildfire dominated landscape is developed based on the data records from the 11 fire events occurred in northwestern Ontario as described in §2.2.1. The 11 fire events vary in size, intensity, and severity but they all occurred within the same ecoregion (2W). The extent and geographic location of each of the fire events along with a brief description of climatic conditions and vegetation species is also presented in §2.2.1.

4.2.2. Landscape data: residual and null-residual patches

The study considered a binary response variable: the presence and absence of residual patches, described as residual and null-residual patches respectively. The residual patches are defined based on the work of Remmel and Perera (2009) in which the NDPE guide was used to define and extract the residual patches. Data pertaining to the absence-data were obtained using a simulation technique in which the null-residual patches were simulated and extracted using an algorithm developed in house. The absence-data are randomly simulated (and placed) within the burned landscape in which the shape, size, and orientation of the null-residual patches mimic the residual patches. The details about the residual and null-residual patches used for developing the predictive model are presented in §3.2.2.

4.2.3. Environmental variables for modelling

The occurrence of residual patches is often attributed to various environmental factors (e.g., weather, ignition source, vegetation, topography, and other pre-fire characteristics) that interactively affect fire behaviour and post-fire forest characteristics. However, the quality, adequacy, and spatial resolution of these environmental variables are imperative for successfully implementing statistical models for prediction (Wilfried et al. 2003). Therefore, the selection of predictors should be undertaken with caution to improve the interpretability of the models; this involves using existing knowledge of physiography and environmental process (Austin 2007). For developing a model that is valid for prediction and extrapolation, the variables should also be consistent spatially and temporally (Jepsen 2004). In this study, the environmental variables used for prediction were selected based on published literature on the response of residual patches to environmental gradients and the availability of appropriate digital coverage of the fire events. The prediction was eventually developed in relation to different predictors, which are obtained from different sources: 1) topographic variables (slope, RI and elevation) from digital elevation models; 2) vegetation cover type from existing pre-fire land cover maps, and 3) distribution of natural firebreaks (surface water, wetland, and non-vegetated areas) from existing maps. A detailed description of the environmental variables is provided in §3.2.3.

4.2.4. Predictive model selection

Various studies have demonstrated the use of statistical (and machine-learning) methods to explain the relationship between species occurrence and environmental variables using regression methods (Guisan and Zimerman (2000), including logistic regression, discriminant functions, CART, and ensemble trees. The application of some of the approaches (e.g., discriminant functions, logistic regression, and GLM) is limited because of the multivariate normality or linear relationship assumption (Edwards et al. 2007; Evans and Cushman 2009). Other statistical models, such as CART and GAM, are more flexible and better suited to handle nonlinear relationships between species and environmental gradients (Hastie et al. 2001; Munoz and Felicisimo 2004). Specifically, tree-based classification approaches (e.g., CART) have gained popularity in landscape ecology (De'ath and Fabricius 2000), but they are associated with the problem of over-fitting and parameter selection. This has led to the development of ensemble methods that overcome the problem of over-fitting and obtain better predictive performance than the standard classification trees. The ensemble method has risen in prominence and is increasingly used for various environmental mapping and modelling applications (Gislason et al. 2006; Cutler et al. 2007; Mellor et al. 2013). In this chapter, a predictive model based on the random forest ensemble-learning approach was applied to develop a spatially explicit predictive model and produce probability maps of residual patch occurrence. The ensemble method was used in this study because it: is a nonparametric and nonlinear classifier that does not require any assumption on data distribution, it has high predictive performance and is computationally efficient. Additionally, unlike most of the predictive models such as logistic multiple regression, GLM, and GAM, RF model allows the integration of categorical variables in the prediction process. These unique features make an ensemble method a powerful tool for relating response and predictor variables, and hence predicting the occurrence and distribution of spatial elements.

4.2.5. Model calibration

The relationship between response variable (residual and null-residual patches) and the environmental variables are modelled using RF as implemented in R (R development core 2013; Liaw and Wiener 2002). The RF model is an ensemble classifier, a statistical procedure based on multiple decision trees used to predict a response variable according to explanatory variables (Breiman 2001). RF makes no assumption on the type of relationship between response and explanatory variables, so it can handle very complex relationships involving interaction and nonlinearity across a response variable. The RF algorithm was used to build a predictive model for which there are some user-defined parameters (e.g., number of trees and number of variables used to split the nodes) that require some adjustment. The details on RF model calibration and implementation is provided in §3.2.4. The model was calibrated by combining data records from the 11 fire events and applying a data splitting using a hold-out approach (Figure 4.2) where data records from a single fire event (e.g., F11) is held-out for testing while data records from the remaining fire events are used for training the model. The procedures undertaken to partition the data into training and test, and calibrate the predictive model using RF algorithm is also given in Figure 4.2.

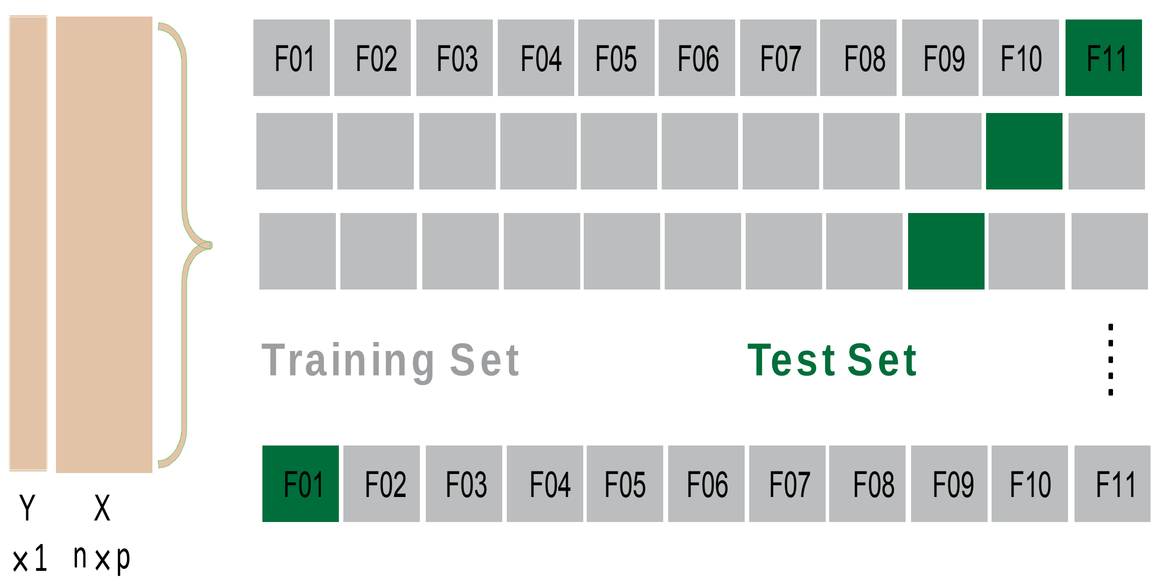


Figure 4.2. Overview of the data partitioning (hold-out or k-fold partitioning) approach implemented in this study for the allocation of cases to training and testing data sets. The data records from a single fire event (e.g., F01) is used for testing while data records from the remaining 10 fire events (e.g., F02 to F11) were used for model construction and calibration.

4.2.6. Model validation

The inherent uncertainty of predictive models needs to be evaluated and quantified (Beauvais et al. 2006). This has been an integral component of any model development and is useful for determining the suitability of a model for specific applications, and compare different modelling techniques (Pearce and Ferrier 2000). The predictive ability of a model (i.e., how accurate a model should be) depends on the conditions to which the model is applied, the types of questions asked, and the alternatives available (Jepsen 2004). However, a model with high predictive performance can generally be used for predicting changes under alternate future scenarios, and informing resource management decisions (Beauvais et al. 2006).

In a presence/absence model, there are two possible prediction errors: false positives (Type I error) and false negatives (Type II error) (Fielding and Bell 1997; Anderson et al. 2003). These errors can result from insufficient sample size, measurement error, and insufficient spatial resolution in the mapped environmental predictors (Pearce et al. 2001). In this study, the prediction error was evaluated based on the discrimination ability of the model to correctly distinguish between positive and negative records (i.e., residual and null-residual patches respectively). This has traditionally been expressed using a confusion matrix as shown in Figure 4.3 (Fielding and Bell 1997). The quadrant of the matrix is populated by cross-tabulating the observed and predicted category of each point in the evaluation set. Elements *a* and *d* in the quadrant are considered as correct classifications where *a* indicates the number of positive sites (residual patches) correctly predicted and *d* denotes the number of negative sites (null-residual patches) correctly predicted. The elements of *c* and *b* are usually interpreted as omission and commission errors respectively

		Actual	
		Present	Absent
Predicted	Residual (Present)	a	b
	Null-residual (Absent)	c	d

Figure 4.3. The derivation of the confusion matrix used as a base for measuring the performance of presence-absence models. The table cross-tabulates observed and predicted patterns: a) true positive; b) false positives; c) false negatives; and d) tree negatives.

One specific way of evaluating the predictive performance of a model is to split the data into training and testing, which are respectively used to develop and validate the model. However, there is no standard rule for splitting the data into training (calibration) and testing (validation) set (Beauvais et al. 2006). Fielding and Bell (1997) summarizes the different approaches that have been used to allocate cases for training and testing; including re-substitution, bootstrapping, randomization, prospective sampling, and *k*-fold partitioning (hold-out or external methods). However, a classic approach to evaluate the accuracy of a predictive model is to compare the model with independent data (i.e., data not used to develop the prediction model). Refaeilzadeh et al. (2008) also indicated that one of the natural approaches of model validation is the use of a hold-out validation approach with independent data, and this has been used in this study to assess the prediction accuracy of RF model. Given the 11 fire events, the data records from an individual fire event (e.g., F01) was held-out for testing while the records from the remaining fire events (i.e., F02 to F11) were used for training purposes .

4.2.6.1. Performance with a fixed-probability threshold

The use of predicted maps for various applications may not be captured in a single map accuracy value; several measures of accuracy should be incorporated (Moisen and Frescino 2002). Some of the measures of model performance are reviewed in (Fielding and Bell 1997; Liu et al. 2009). Each of the measures tends to emphasize on a particular aspect of model performance, and hence serves a specific purpose (Beauvais et al. 2006). Some of these global measures of model performance (e.g., Table 4.1) can be computed from a 2 by 2 contingency table of predictions and observations shown in Figure 4.3. The simplest and most widely used measure of prediction accuracy is the percent correctly classified (*PCC*) but model assessment using the overall accuracy, with no indication on the present or absent success might be misleading. Therefore, the overall measure of accuracy can be broken into present success (Sensitivity – *Sn*) and absent success (Specificity – *Sp*) (Table 4.1). . The former, also known as true positive fraction, refers to the proportion of presence (i.e., residual patches) correctly predicted; the later (true negative fraction) is the proportion of absence (null-residual patches) correctly classified. The three indices, which are also referred to as fixed-probability threshold measures, capture a bit of the information on model performance and when presented together they provide most users a good sense of model quality. In this study, the validity of the RF model was initially assessed using the three measures of model performance: *PCC*, *Sn*, and *Sp* (Table 4.1). All classification analyses were carried out in the R statistical package (R development core Team 2013).

Table 4.1. Potential measures of presence-absence model's performance. The measures of model's performance are derived from the confusion matrix shown in Figure 4.3. The formulae are based on correctly predicted positive occurrences (a), falsely predicted positive occurrences (b), falsely predicted negative occurrence (c), and correctly predicted negative cases (d).

Measure:	**Calculation**
Percent correctly classified (PCC)	$PCC = \frac{(a+b)}{(a+b+c+d)}$
Present success rate (*Sn*)	$Sn = \frac{(a)}{(a+c)}$
Absence success rate (*Sp*)	$Sp = \frac{(b)}{(b+d)}$

4.2.6.2. Threshold-independent measures of model performance

The fixed-probability threshold is based a single cut-off value, a value that is used to translate predicted probabilities into a binary (0 and 1) class where the default threshold is 0.5 (Jepsen 2004; Beauvais et al. 2006). Yet, the choice of an appropriate threshold value is difficult, often arbitrary, and affects the measures of model performance. It does not necessarily provide a more accurate accuracy measure (Manel et al. 1999). Therefore, a more universal approach, threshold-independent indices (i.e., methods based on broad spectrum of threshold values), are needed (Pearce et al. 2001). Liu et al. (2009) summarizes some of the threshold-independent accuracy measures of model performance, and one of the most widely used measures is receiver operating characteristics (ROC) curves (Fielding and Bell 1997; Jepsen 2004; Beauvais et al. 2006; Peters et al. 2007). This was originally developed by signal processing and medical researchers (Zweig and Campbell 1993; Peters et al. 2007), and has recently been integrated into distribution modelling for assessing model's performance (Jepsen 2004).

The ROC curve provides a graphical depiction of model's discrimination ability over a range of threshold values (Pearce and Ferrier 2000). It is obtained by plotting all true positive fractions (sensitivity values; on y-axis) and false positive fractions (1- specificity; on x-axis) over all available thresholds (Zweig and Campbell 1993; Fielding and Bell 1997). A model with perfect discrimination ability has an ROC plot that passes through the upper left corner, representing perfect sensitivity (true-positive fraction = 1) and perfect specificity (false-positive fraction = 0) (Figure 4.4). The theoretical plot for a test with no discrimination (i.e., a completely random guess or chance of performance) is a 45° diagonal from the lower left corner to the upper right corner (Zweig and Campbell 1993; Pearce and Ferrier 2000; Fielding and Bell 1997).

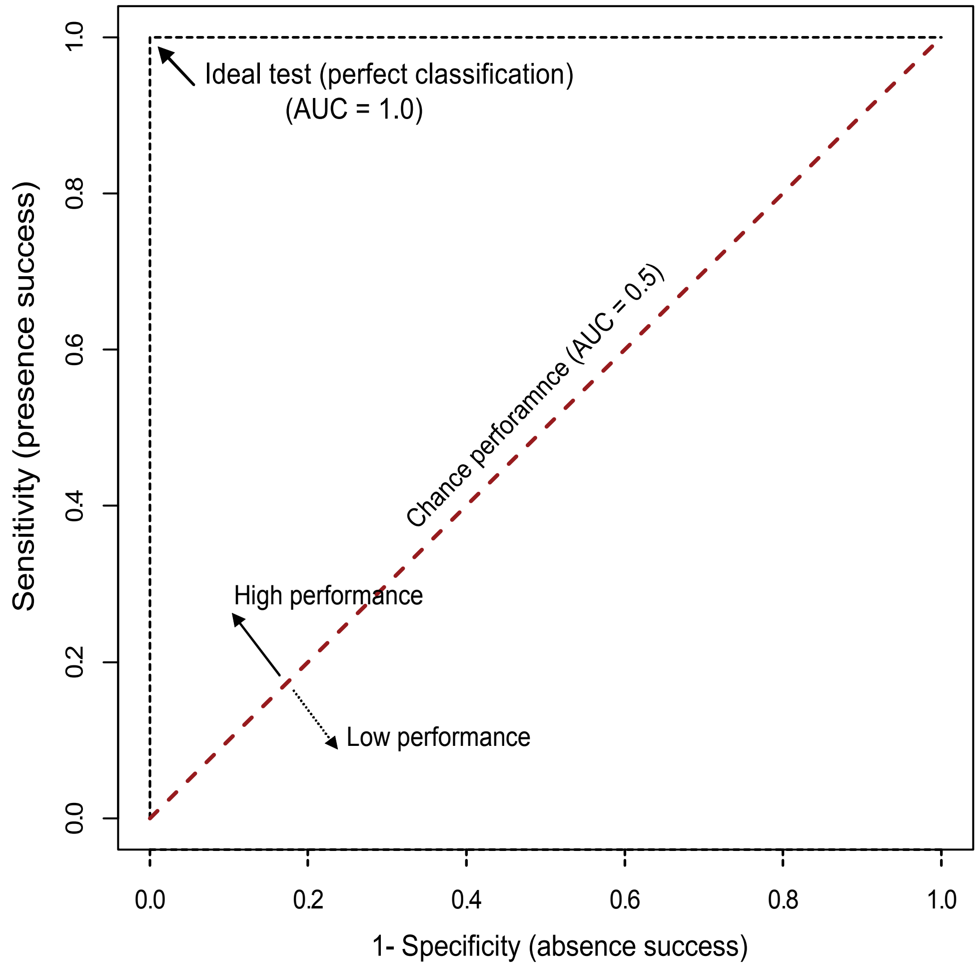


Figure 4.4. Hypothetical example of ROC graph in which the sensitivity (true positive proportion) is plotted against the false positive proportion for a range of threshold probabilities. A perfect model follows left of the axis and top of the plot while the 45° line represents the sensitivity and false positive values expected to be achieved by chance alone for each decision threshold.

The ROC plot for assessing model performance has received a considerable attention because 1) it is simple, graphical, and easy to understand visually, and 2) of its discrimination ability of the presence-absence over a wide range threshold values (Zweig and Campbell 1993). Therefore, the ROC curve was generated to assess the predictive performance of the model, independently of a specific threshold set to classify the data records into residual and null-residual patches. In order to construct ROC curves, the predicted probabilities of residual occurrence across the events (and spatial resolutions) were used to generate several confusion matrices, one for each possible cut-point. A cut-point represents a threshold probability above which the residual patch is modelled to be present (Peters et al. 2007). Pearce and Ferrier (2000) noted that large numbers of sensitivity and false positive pairs (i.e., based on a threshold interval of 0.01) would result in a better fit. Thus, the threshold interval at 0.01 across the predicted probability range was used to produce ROC plots over 100 threshold values evenly spaced across the range of available predicted values (from 0.0 to 1.0).

However, comparing ROC curves directly from the plot has never been easy and is subjective (Eunsik and Wenbao 2011); a single index that describes the discrimination ability of a model is required (Zweig and Campbell 1993; Pearce and Ferrier 2000). The area under the resulting ROC curve, which is referred to as AUC, is then considered as an indicator (discrimination index) of model's performance. The AUC provides a single measure of model's ability to distinguish between residual and null-residual patches, independent of a specific threshold value (Munoz and Felicisimo 2004; Peters et al. 2007; Refaeilzadeh et al. 2008). The AUC is expressed as a proportion of the total area of the unit square defined by the false positive and true positive axes (Pearce and Ferrier 2000), with high AUC values (i.e., large areas under the curve) indicates a high predictive performance of a model. ROC plots for each of the fire events using R were produced; for each of the ROC curve the AUC value was also computed. As a general rule, the AUC value ranges from 0.5 for a model with no discrimination ability to 1.0 for models with perfect discrimination ability (Table 4.2). In order to test whether each of the ROC index computed was significantly greater than 0.5, a statistical test based on Wilcox test was also computed.

Table 4.2. Classification of AUC values for assessing model performance (source: Swets 1988). The presence-absence models can be categorized as strong, marginal, or poor model based on the AUC values.

AUC values	**Description**	**Remarks**
0.5	Random guess	Discrimination ability of a model is equivalent to the one obtained by a random model (i.e., random assignments of predicted values to sites)
0.5 – 0.7	Low accuracy (Poor discrimination ability) – weak model	Sensitivity rate is not much more than the false positive rate
0.7 – 0.9	Reasonable discrimination ability – marginal model	Useful application
> 0.9	High accuracy (Good discrimination ability) Strong model	The sensitivity rate is high relative to the false positive rate

4.2.7. Spatial prediction

The advancement in computing power and GIS technology has enhanced the possibility of generating predicted probability maps; rather than having as abstract formulae or qualitative description of model output (Beauvais et al. 2006). The ability to generate such spatially explicit predicted maps has also been one of the goals of predictive modelling. However, the predicted maps do not show the actual distribution of residual patches; they are rather cartographic representations based on probabilities of residual occurrence (Guisan and Zimmermann 2000). Owing to the variation in modelling techniques, there is no standard procedure for expressing predictive models in a (digital) map form (Beauvais et al. 2006). Yet, some potential approaches (e.g., *RSAGA* in R) have been implemented to convert predicted probability values into digital (*ASCII* grid) or any GIS readable format maps.

The prediction using RF produces deterministic (binary) response classes or predicted probabilities (i.e., matrix of class probabilities). However, most environmental attributes are inherently continuous and classifying spatial elements into discrete (deterministic) classes of presence and absence yields a simplistic view of the landscape (Evans and Cushman 2009). This limits our ability to examine the continuous nature of residual patch occurrence in a disturbed landscape. In this study, the occurrence of residual patches is represented as a separate probability surface rather than a mosaic of discrete patches that are implicitly assumed to be categorically discrete. The predictive probability maps that show the potential areas where residual patches are likely to occur, at each grid of the fire footprint, were also produced as *ASCII* format in R, and imported into *ArcMap* for display and analysis.

4.3. Results

4.3.1. Prediction performance with fixed-probability threshold

The three measures of accuracy that are based on a fixed probability threshold and integrated in this study are *PCC*, *Sn,* and *Sp* (Table 4.3); the latter two indices (*Sn* and *Sp*) are incorporated because the overall accuracy (*PCC*) alone does not provide information about the commission and omission errors included in the predictions (Frescino et al. 2001). As shown in Table 4.3, the overall accuracy at predicting residual and null-residual patches always exceeded 50% for all fire events (at R_4, R_8, R_{16}, R_{32}, and R_{64}) except for three fire events (F02, F03, and F06; at R_{32} and R_{64}). The *PCC* values also varied among the fire events (and with spatial resolutions). At R_4, the overall accuracy was relatively high, with *PCC* value greater than 80% for all of the fire events except for F06, F07, and F11. The highest model accuracy was observed for one of the largest fire events (i.e., F01; 96.52%); yet the lowest *PCC* was obtained for another large fire event (i.e., F06; 68.41%). The accuracy for the other two largest events (i.e., F08 and F10) was also high, with *PCC* value of 86.13 and 89.63% respectively. As is evident from the table below, such a trend tends to continue, but with a slight variation with increasing grain sizes. The overall accuracy generally tends to decrease with increasing grain size, indicating the sensitivity of model to changing grain sizes; model's predictive performance was high at finer spatial resolutions.

Table 4.3. The discrimination performance of the presence-absence models to distinguish between residual and null-residual patches based on fixed-probability threshold measures: Percent correctly classified (PCC), Sensitivity (Sn) and specificity (Sn) are presented.

		Fire events (Statistically independent dataset for assessing predictive performance)										
		F01	F02	F03	F04	F05	F06	F07	F08	F09	F10	F11
R4	*PCC*	96.52	90.47	88.88	92.96	83.56	68.41	76.99	86.13	80.95	89.63	69.14
	Sn	99.69	100.00	85.71	97.16	77.48	81.14	88.20	91.75	92.30	98.72	80.00
	Sp	91.46	0.00	100.0	80.64	93.70	20.23	43.26	62.01	62.50	68.94	53.84
R8	*PCC*	75.54	85.37	86.49	82.61	73.43	68.78	69.73	80.16	69.23	70.49	67.67
	Sn	96.90	93.10	85.00	79.08	63.18	79.85	86.02	89.65	100.00	92.41	98.38
	Sp	36.05	66.67	88.23	87.78	86.52	19.37	30.25	45.85	50.00	38.79	17.54
R16	*PCC*	61.92	68.62	67.34	70.45	53.64	57.35	63.37	70.26	75.00	58.57	50.60
	Sn	94.83	53.57	68.18	72.35	39.95	66.04	78.67	90.89	80.00	86.39	87.78
	Sp	15.49	86.95	66.67	67.86	70.41	29.54	29.16	10.51	71.42	13.28	8.00
R32	*PCC*	68.02	48.57	31.42	61.91	51.97	41.15	59.42	63.58	53.84	61.81	57.46
	Sn	86.79	50.00	5.00	68.16	66.67	30.91	80.92	89.56	62.50	78.23	94.65
	Sp	37.56	47.05	58.82	53.14	32.93	69.25	23.87	15.71	40.00	26.81	3.00
R64	*PCC*	66.96	37.50	30.77	60.34	40.90	48.81	53.16	59.96	70.00	58.48	52.36
	Sn	81.76	22.22	14.28	38.57	17.51	43.67	68.55	75.57	40.00	65.13	60.33
	Sp	49.40	57.14	50.00	82.57	73.52	60.22	33.77	34.09	100.00	46.58	40.17

High values of overall accuracy can be deceptive (Edwards et al. 2007); hence the other alternate measures of accuracy (*Sn* and *Sp*) should be included for a better model assessment. The sensitivity value was relatively higher for all fire events, ranging from 77% (for F05) to 100% (for F02) at R_4, but the trend was not similar throughout fire events and across the five spatial resolutions.

4.3.2. Threshold-independent measures of model performance

Before calculating the ROC curves, the discriminatory power of the model was assessed visually by comparing the distribution (or variability) of predicted probabilities across the five spatial resolutions (Figure 4.5- Figure 4.7). The graphs indicate that the median values for sites at which residual patches were present are relatively higher than those for null-residual patches particularly at R_4; confirming the model's ability to discriminate residuals from null-residual patches. This is not however the case for some of the observations (e.g., F06) where the model had difficulty in distinguishing positive records (residual patches) from negative records (null-residual patches). The graphs also show that the variability in predicted values was relatively low at R_4, indicating that the model had good discrimination ability at finer spatial resolutions than coarser spatial resolutions. The refinement of the values predicted by the model also varies across the events (and the spatial resolutions) with prediction values ranging from 0 to 1.0.

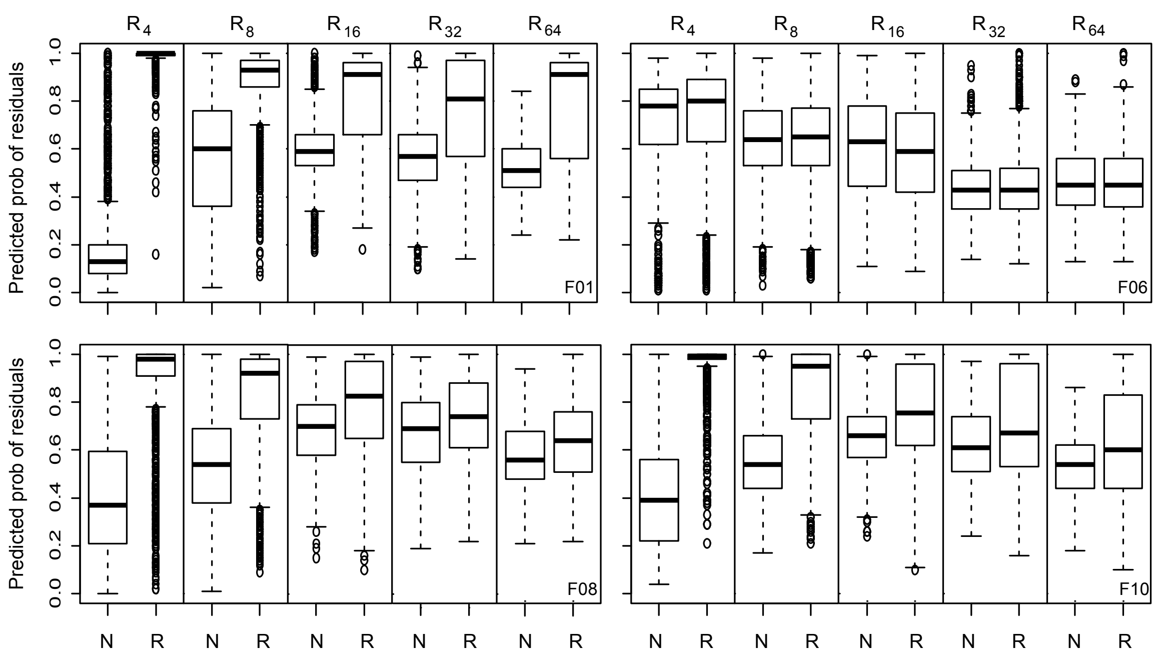


Figure 4.5. The discrimination ability of the model to distinguish between residual and null-residual patches: variability of predicted probability values associated with residual (R) and null-residual (N) patches, for the large sized fire events.

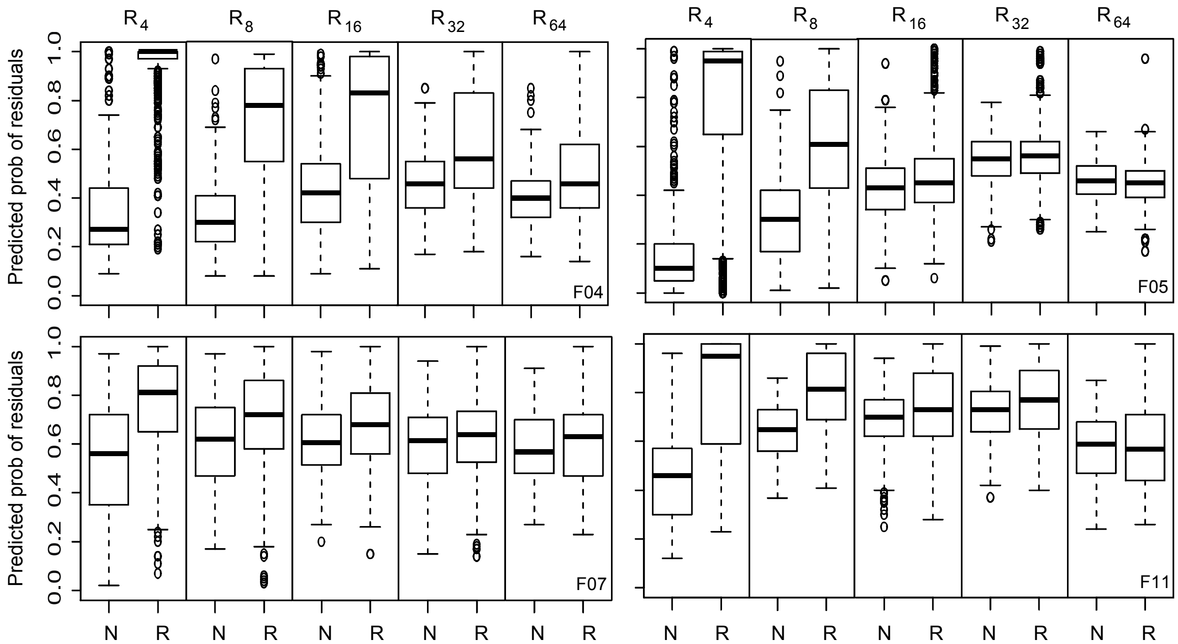


Figure 4.6. The discrimination ability of the model to distinguish between residual and null-residual patches: variability of predicted probability values associated with residual (R) and null-residual (N) patches, for the large sized fire events.

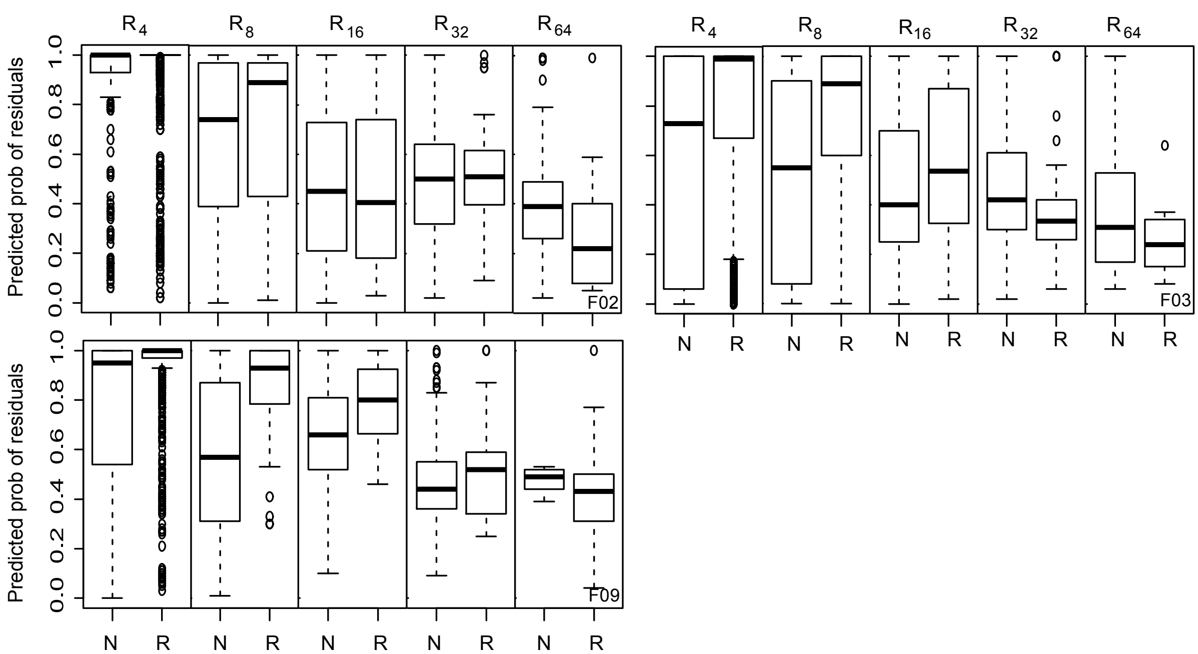


Figure 4.7. The discrimination ability of the model to distinguish between residual and null-residual patches: variability of predicted probability values associated with residual (R) and null-residual (N) patches, for the large sized fire events.

To further examine the discrimination capability of the model, the relationship between the proportions of observed presences correctly predicted (sensitivity) and the proportion of observed absences incorrectly predicted (1 – specificity) was graphically summarized using ROC curves (Figure 4.8). A model that perfectly predicts the residual patches generates an ROC curve that follows the left axis and top of the plot, whilst a model with predictions that are no better than random produces an ROC curve that follows a 45° diagonal from the lower left corner to the upper right corner. The shape of the ROC shown in Figure 4.8 describes the trade-off between true positive and false positive rates as the threshold probability is changed (Pearce and Ferrier 2000). A plot lying above and to the left of another plot indicates greater observed accuracy (Zweig and Campbell 1993); such trend was evident in the ROC curves shown in Figure 4.9 with changing grain sizes. The curve for some of the fire events (F01, F04, F05, F08, and F10) at R_4 was closer to the perfect discrimination. However, it is subjective and not easy to assess and compare the predictive accuracy directly from the ROC curves (Refaeilzadeh et al. 2008). The validity of the model has to be assessed based on the estimated areas under the respective ROC curves (AUC).

The AUC provides a summary measure of a model's predictive accuracy; the ROC curve with the larger area is, on average, more accurate (Pearce and Ferrier 2000). The AUC values shown in Table 4.4 provide a measure of model's ability to discriminate between locations where residual patch of interest is present or absent. The model had the highest discrimination accuracy with an index value of 0.995 for F01 at R_4, suggesting the model could discriminate between residual and null-residual patches 99.5% of the time. At R_4, RF model had also AUC values greater than 0.7 for all the cases except for F02 and F06; the model's discrimination ability at this scale was evaluated as marginal model (having reasonable ability) to strong model (with excellent discrimination ability) based on the rule of thumb set by Swets (1988). The lowest accuracy was observed for F02 and F06, with AUC values of 0.629 and 0.563 respectively. Yet, the predictive model at R_4 had significantly higher discrimination ability ($p < 0.05$) for all fire events.

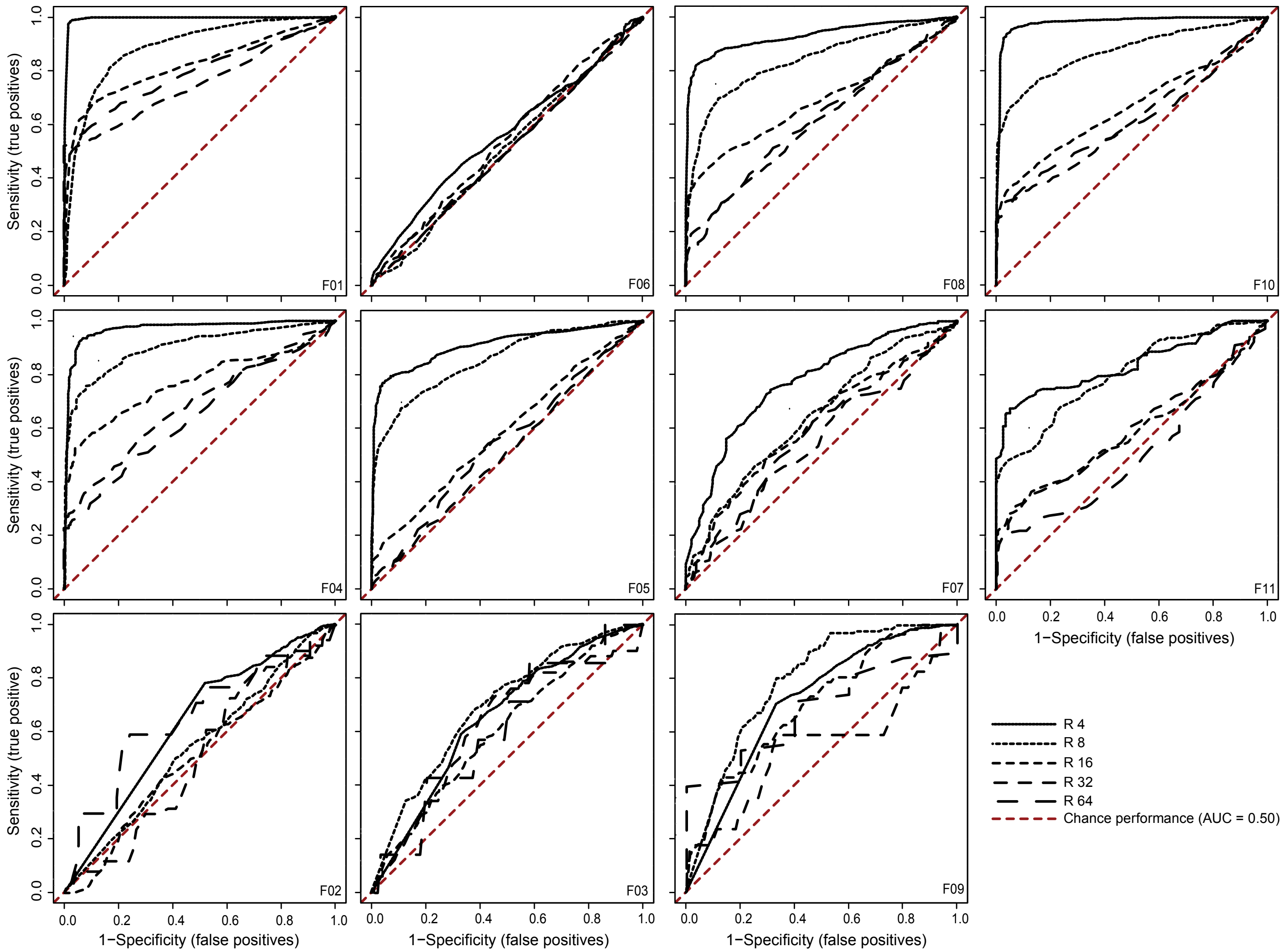


Figure 4.8. Graphical representation of the predictive performance of RF model for the 11 fire events; each ROC curve in the plot depicts model's performance at specific spatial resolution.

Table 4.4. Quantitative measures of model's predictive performance based on AUC values where AUC values between 0.9 and 1.0 show strong model, values between 0.7 and 0.9 marginal model, and values between 0.5 and 0.7 are poor models. The p-value tests the significance of the area under a ROC curve (bold cells indicate statistically not significant AUC values; statistically not different from random prediction).

		Spatial resolutions									
		R_4		R_8		R_{16}		R_{32}		R_{64}	
	Fire ID	**AUC**	***ρ-value***	**AUC**	***ρ-value***	**AUC**	***ρ-value***	**AUC**	***ρ-value***	**AUC**	***ρ-value***
Large fire events	F01	0.995	0.0000	0.886	0.0000	0.816	0.0000	0.749	0.0000	0.793	0.0000
	F06	0.563	0.0000	0.506	**0.3230**	0.544	**0.9999**	0.503	**0.4195**	0.500	**0.4986**
	F08	0.933	0.0000	0.844	0.0000	0.685	0.0000	0.605	0.0000	0.613	0.0000
	F10	0.981	0.0000	0.874	0.0000	0.659	0.0000	0.616	0.0000	0.611	0.0000
Medium fire events	F04	0.970	0.0000	0.902	0.0000	0.771	0.0000	0.688	0.0000	0.643	0.0000
	F05	0.910	0.0000	0.854	0.0000	0.584	0.0000	0.512	**0.2185**	0.546	**0.9879**
	F07	0.770	0.0000	0.642	0.0000	0.611	0.0000	0.555	0.0130	0.567	0.0159
	F11	0.837	0.0000	0.799	0.0000	0.588	0.0009	0.590	0.0006	0.503	**0.4654**
Small fire events	F02	0.629	0.0000	0.537	0.0000	0.507	**0.6446**	0.507	**0.4161**	0.648	**0.9788**
	F03	0.647	0.0000	0.688	0.0000	0.601	0.0000	0.622	**0.9981**	0.611	**0.8404**
	F09	0.710	0.0000	0.786	0.0000	0.699	0.0000	0.551	**0.2349**	0.677	**0.9115**

Similarly, the model has a reasonable discrimination ability at R_8 for all events ($0.786 < AUC < 0.886$) except for F02, F03, F06, and F07 while high AUC value was retained for F04 (AUC = 0.902). The RF model at R_{16} had also low or poor discrimination ability for most of the events ($0.507 < AUC < 0.699$), but exhibited a reasonable discrimination ability for F01 and F04, with AUC index values of 0.816 and 0.771 respectively. At the coarser spatial resolutions (i.e., R_{32} and R_{64}), the model displayed discrimination rates less than 0.7 for all except F01, and the model did not perform significantly better than random for most of the fire events.

4.3.3. Scale effect

The accuracy measures (i.e., AUC) derived from the optimized probability thresholds markedly increased the accuracy derived at the conventional probability threshold of $p = 0.5$ (i.e., fixed-probability threshold). The predictive performance of the model was assessed at different grain sizes to determine whether residual patches could be modelled more successfully at specific grain sizes (e.g., at finer spatial resolutions). A comparison of the RF model across scale gradient revealed that some distinctive patterns in model performance were observed as a function of scale (grain size). At the finest scale (R_4), for example, the model provided better discrimination, but the predictive accuracy decreases with increasing grain size in a consistent power-law relationship for all large and medium sized fire events (except for F06), with a coefficient of determination > 80% (Figure 4.9). Such a monotonic decreasing trend with a power law relationship was not observed for the smallest fire events – F02, F03, and F09 (Figure 4.9).

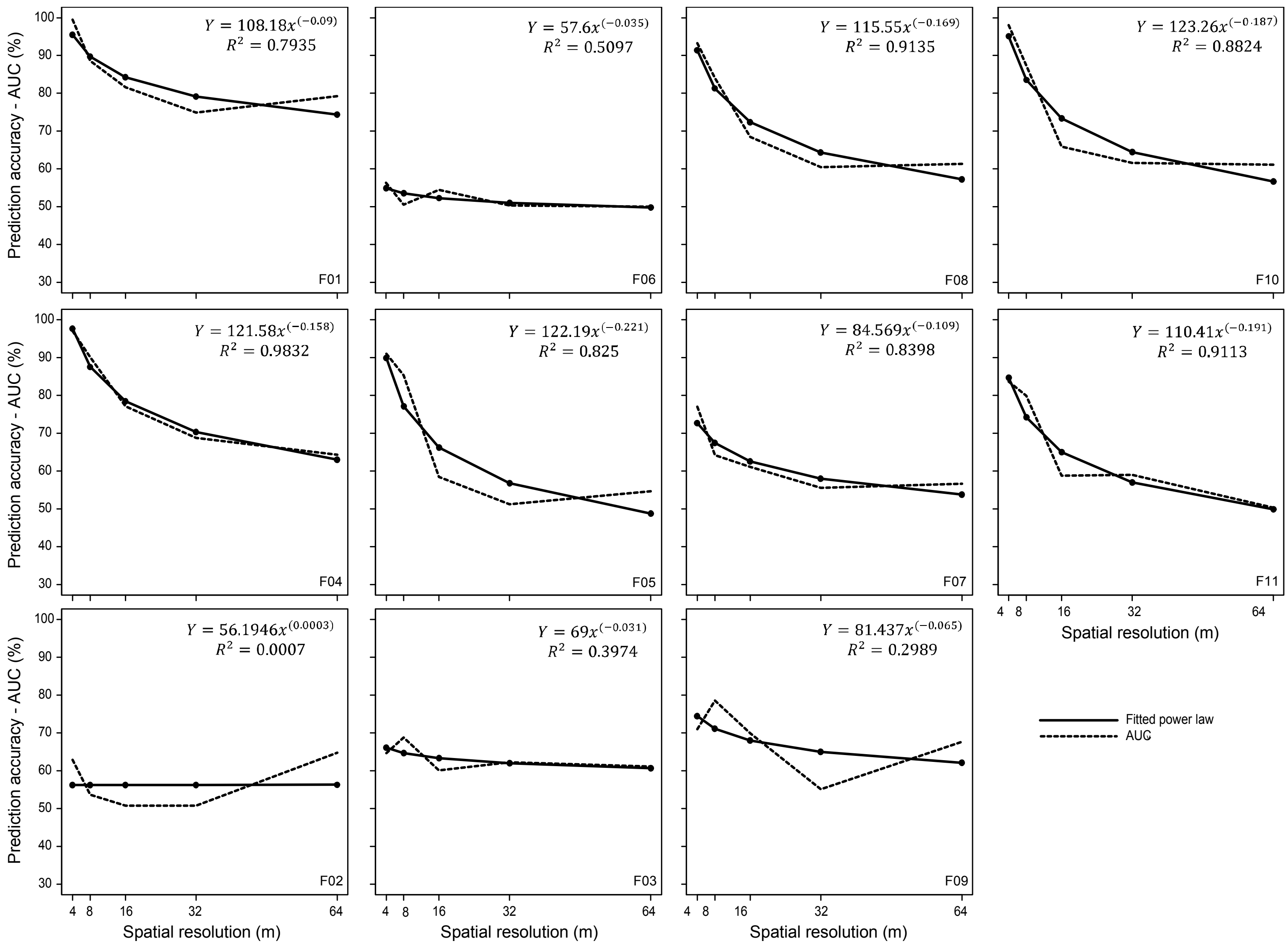


Figure 4.9. The predictive performance of the model as a function of scale: the dashed line in each plot shows the AUC values computed at different spatial resolutions while the solid line is the best fit model.

4.3.4. Spatial prediction of residual patches

While the statistical validity and accuracy of the model is important, graphical representation of the model output (i.e., spatially explicit predicted maps) are desired for examining the spatial distribution of spatial elements (Fielding and Bell 1997). The predicted probability maps of residual patch occurrence were obtained for all fire events at the given spatial resolutions. In this section, the results from R_{32} are only graphically summarized in the predicted probability maps shown in Figure 4.10- Figure 4.14; the maps are given at 32 m spatial resolution because it is the closest to how a remote sensing device based on Landsat imagery would view the landscape. The output from the models is a probability value scaled from 0 to 1 for each grid cell, with predictions closer to 1 indicating greater chance of residual patch occurrence.

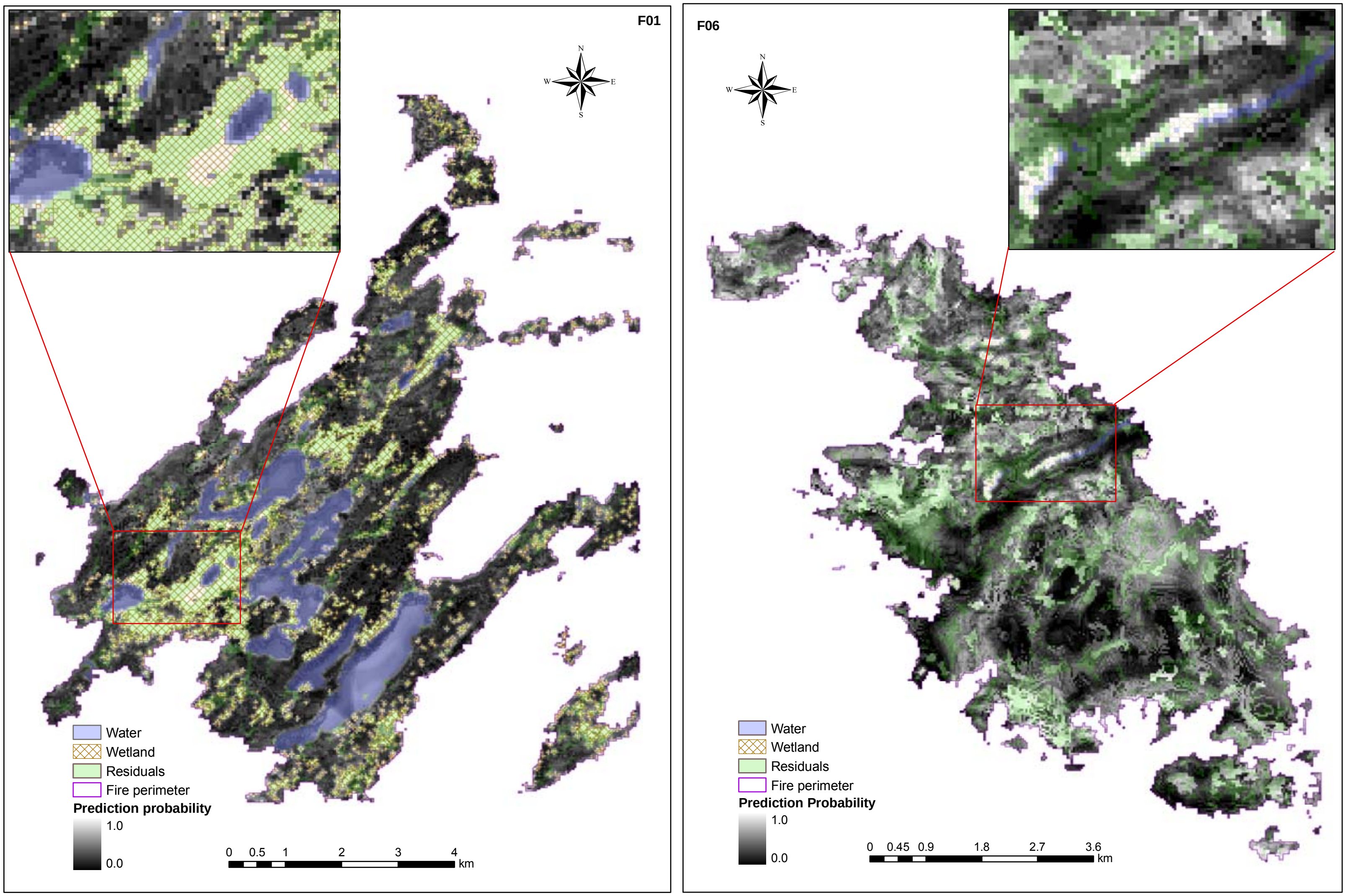


Figure 4.10. Predicted probability maps of residual patch occurrence for large sized fire events (F01 and F06) at R_{32}; lighting shading – greater chance of residual patch occurrence, cross-hatched areas – distribution of wetlands, and light blue – distance of water.

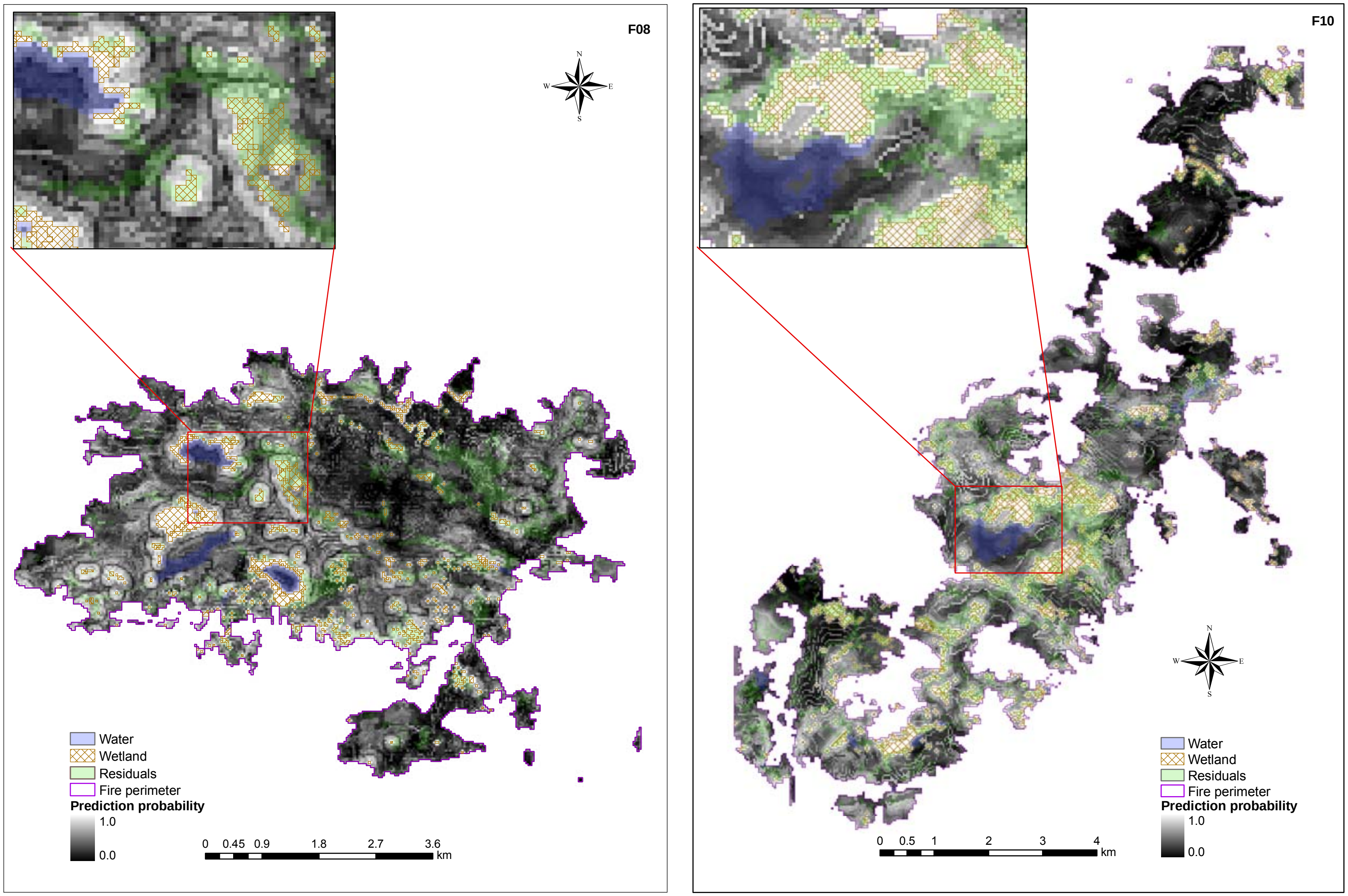


Figure 4.11. Predicted probability maps of residual patch occurrence for large sized fire events (F08 and F10) at R_{32} lighting shading – greater chance of residual patch occurrence, cross-hatched areas – distribution of wetlands, and light blue – distance of water.

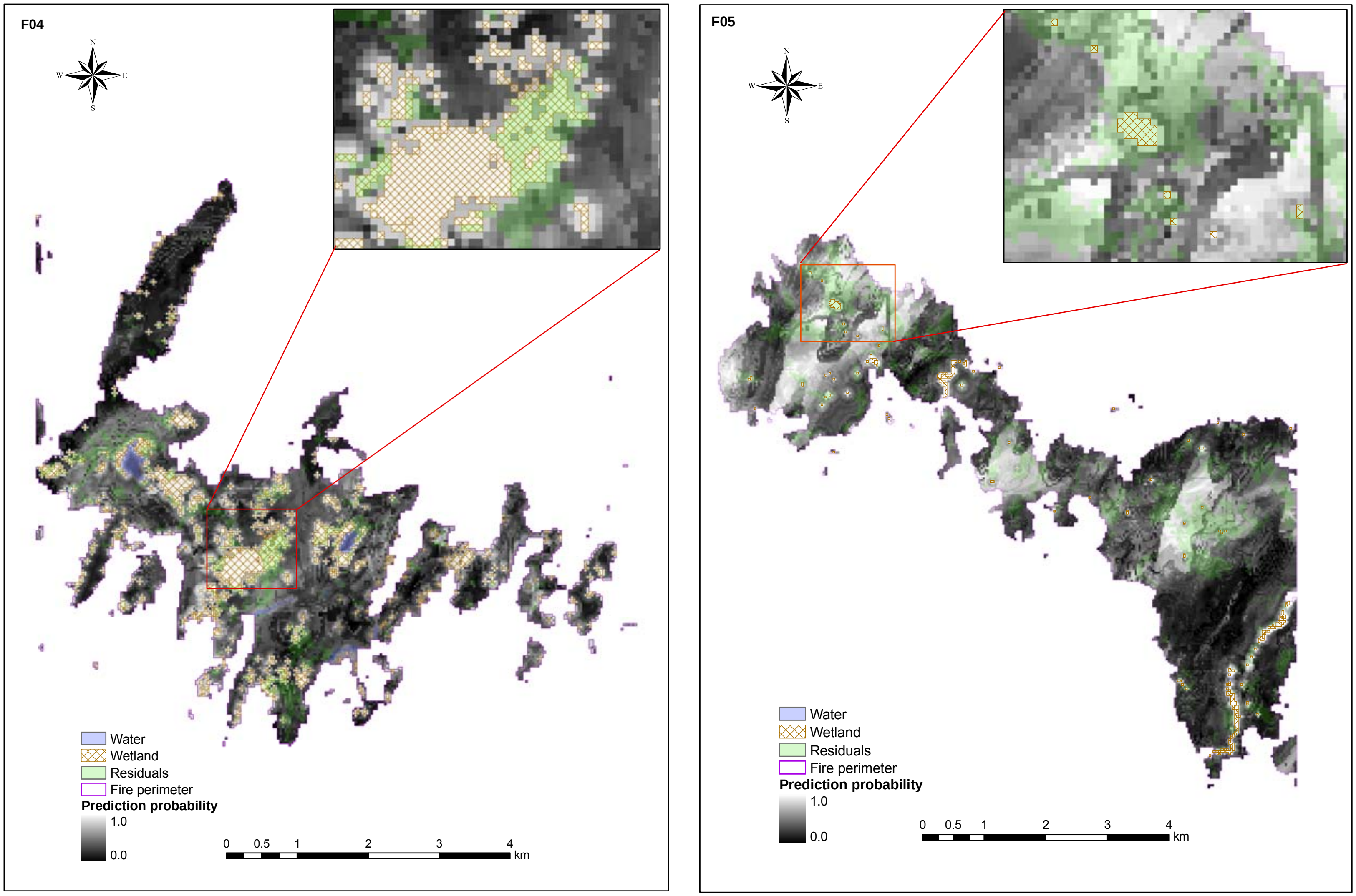


Figure 4.12. Predicted probability maps of residual patch occurrence for medium sized fire events (F04 and F05) at R_{32} lighting shading – greater chance of residual patch occurrence, cross-hatched areas – distribution of wetlands, and light blue – distance of water.

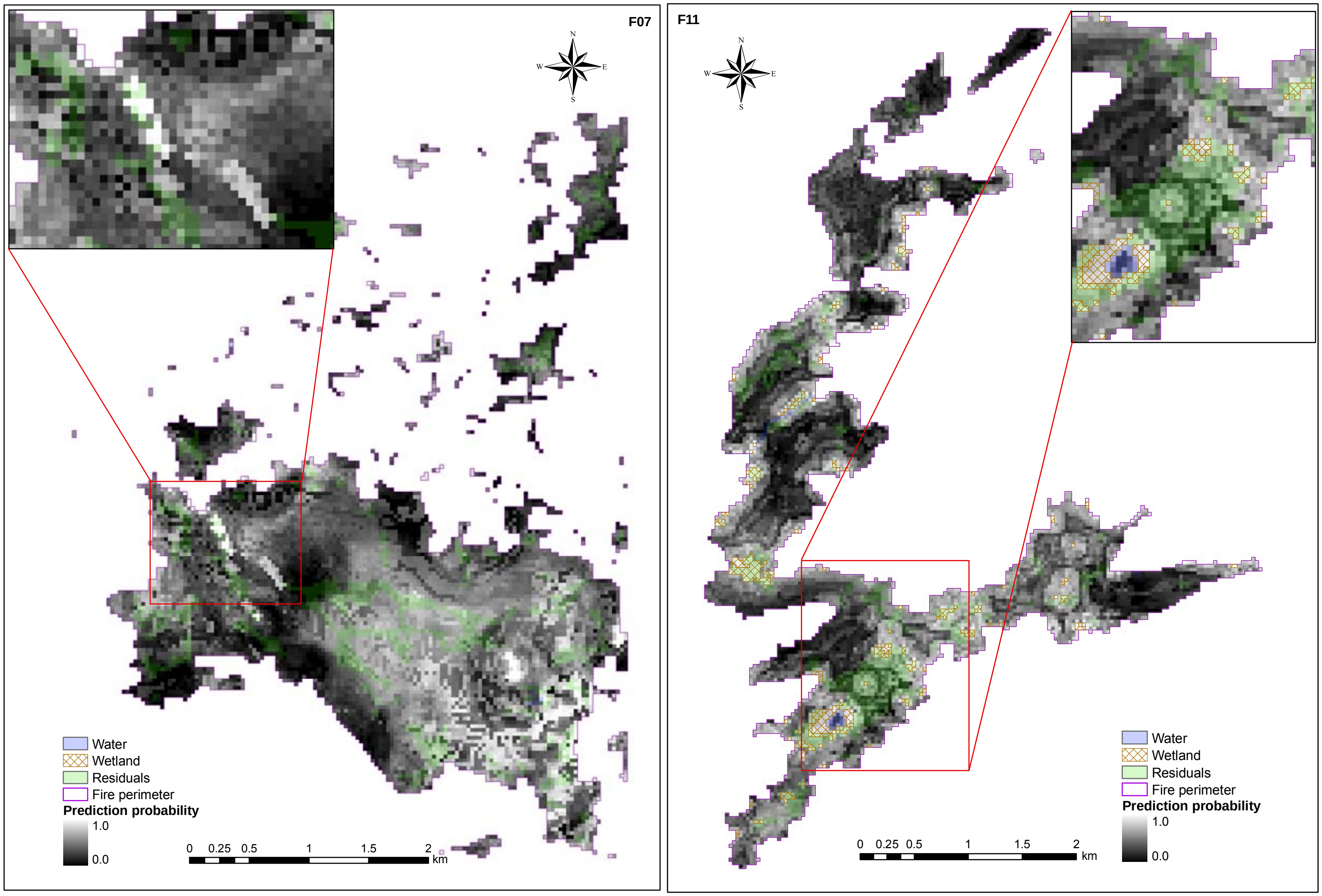


Figure 4.13. Predicted probability maps of residual patch occurrence for medium sized fire events (F07 and F11) at R_{32} lighting shading – greater chance of residual patch occurrence, cross-hatched areas – distribution of wetlands, and light blue – distance of water.

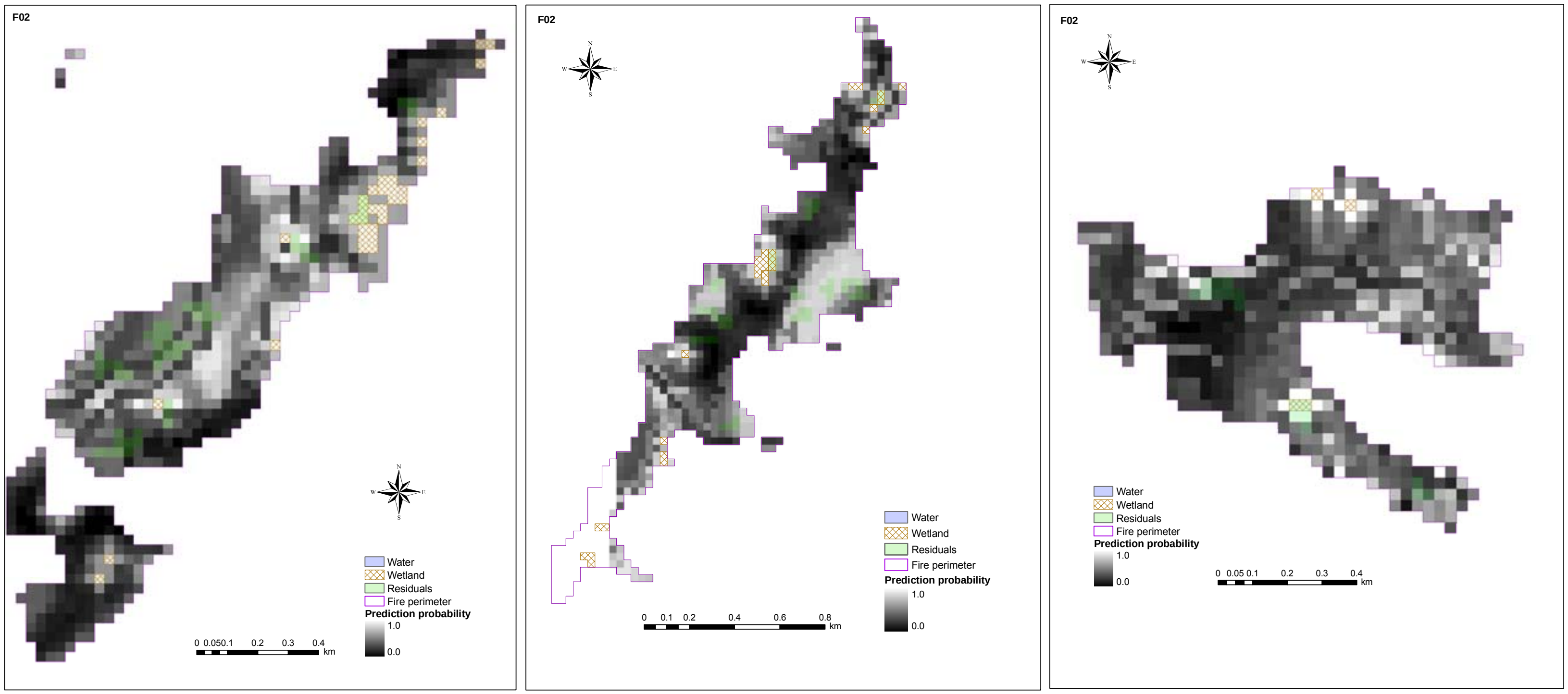


Figure 4.14. Predicted probability maps of residual patch occurrence for small sized fire events (F02, F03, and F09) at R_{32} lighting shading – greater chance of residual patch occurrence, cross-hatched areas – distribution of wetlands, and light blue – distance of water.

4.4. Discussion

4.4.1. Model performance with fixed-probability threshold

The RF model provides OOB error statistics, which are indicative of model fit, but not necessarily predictive performance of a model. I chose to perform an external cross-validation (hold-out validation) approach to provide a statistically independent measure of model performance; as effective and correct model assessment has real significance to distribution modelling (Manel et al. 2001). The results of the fixed-probability threshold, specifically the overall accuracy, indicated that the model did a reasonable job to discriminate between residual and null-residual patches. The results from the other two alternative measures of accuracy also indicated that 1) the model has successfully discriminated residual patches from null-residual patches, particularly at finer spatial resolutions; 2) the accuracy was marginally higher for the residual patches than null-residuals, suggesting a tendency for the model to misclassify null-residual patches as residual patches. This has led to a slight overestimation of residual patches in the prediction. There was also a situation in which the accuracy tended to be higher for null-residual patches, indicating a slight underestimation of residual patches for certain fire events (e.g., F03 and F05 at R_4; F03, F04, and F05 at R_8; F02 and F05 at R_{16}); 3) the predictive accuracy of the model showed different patterns as a function of scale; the ability of the model to correctly classify residual and null-residual patches decreases with increasing grain sizes. Given the focus of this study (i.e., predicting residual patch occurrence), the model has successfully predicted residual patches with high accuracy (sensitivity) values. Overall the results for most of the fire events were appealing, but being a fixed-probability threshold the results from the three indices need to be evaluated further. Although the overall prediction success (based on a fixed-probability threshold) is widespread in determining classification accuracy, Manel et al. (2001) noted that this measure is inherently misleading and fails to consider the prevalence effects (i.e., the frequency of occurrence of a target element). The validity of the model performance was thus assessed using an alternative approach (ROC curve) that takes into account the prevalence effects and threshold-independent approach.

4.4.2. Model performance with independent measure of performance

The predictive accuracy based on RF model was generally well supported and significant at $p < 0.05$, for the finest spatial resolutions (R_4, R_8, and R_{16}). As a general rule of thumb, AUC values greater than 0.9, and values between 0.7 and 0.9 indicate a strong and marginal model respectively. The values reported here indicated that the model has successfully predicted the occurrence of residual patches particularly at R_4 but the accuracy decreased with increasing grain sizes. The high prediction accuracy at R_4 suggests that at this particular scale (and for this data), residual patch occurrence is influenced by a combination of topographic variables, natural firebreak features, and land cover variables. The trend was not true for some of the fire events (F02, F03, F06, and F09); a question that arises is why the model did not perform well for these specific fire events. Given the prediction approach (hold-out validation); one can claim that the fire events (F02, F03, F06, and F09) are different from the rest in relation to the variables considered. One important finding is that three of the observations (F02, F03, and F09) are the smallest fire events, and one plausible reason for the low prediction accuracy could be attributed to the sample size. The extent of the fire events is less than 100 ha; yet the number of patches or records in the evaluation dataset was small (less than 10 records for an event). Edwards et al. (2007) argued that predictive models usually attain more accurate predictions with increased sample size. Similarly, Pearce et al. (2001) found that the performance of a model for rarer species, with less than nine records in the evaluation dataset, was poorer than those with large number of records. Schwarts et al. (2006) also argued that few observation of species distribution resulted in low statistical power of a prediction. Despite the low prediction accuracy, the accuracy (for F02, F03, and F09) was statistically significant at R_4 and R_8.

The ability to discriminate among the large fire events (except for F06) ranges from marginal to strong model outcome, suggesting that the distribution of residual patches appears to be explained by the predictors incorporated in the model. However, particular attention may need to be devoted to improving the predictions for certain fire events (e.g., F06), as the model had generally low predictive performance. One potential reason could be the differences within a landscape as a function of the environmental attributes and burn severity. Describing a low predictive performance of a model, Burton et al. (2008) argued that there are inter-regional or inter-landscape differences among disturbed landscapes. Another reason for such poor prediction accuracy (for F06) may also rely on the distribution of unburnable cover types within the fire perimeter. In the previous chapter, the variable importance assessment indicated that wetland is the most important predictor that explains the occurrence of residual patches for most of fire events except for F06. The proportion of this predictor (wetland) for F06 was not substantial (only < 0.1% of the fire footprint at R_4); this can have a profound effect on the prediction accuracy at this specific event. Therefore, for a model to predict spatial elements with an excellent discrimination ability (strong model), the parameters (e.g., wetland) that determine the residual patches should occupy a considerable portion of the fire footprint.

4.4.3. Sensitivity of model performance to scale change

The results of the significance test revealed that 1) the AUC index at R_4 was found to be significantly better than that expected from a random model ($p < 0.05$); this is regardless of the poor performance observed for some of the fire events; 2) the discrimination ability of the model at R_8 was statistically significant for all events except for F06 ($p = 0.3230$); 3) the model's predictive performance was poor and statistically not significant for F06 across the gradient of scales considered except at R_4; 4) conversely, the model performed significantly better than random for F01, F04, F07, F08, and F10 across all the grain sizes; this is in spite of the low AUC values obtained for some events at certain grain sizes (e.g., F04 and F08 at R_{32} and R_{64}); and 5) the performance of the model to predict residual patches decreases with grain size, and hence the statistical significance was sensitive to changing grain sizes. Despite the variations in the significance test of the model for predicting residual patches across gradient of scales, the study suggested that a scaling rule can be developed to predict residual patches across gradient of scales for certain fire events except F02, F03, F06, and F09 where low predictive accuracy was attained for these events over multiple scales. Yet, a robust scaling rule that can be used to predict residual patches across all the fire events may not be able to be established given the variables considered in the study.

4.4.4. Spatial predictions

Predictive models for predicting presence-absence are used increasingly in landscape ecology, particularly for conservation planning (Manel et al. 2001), but the analysis of residual patches in the study sites is often based on information obtained from classified maps. The maps represent residual patches as categorical type in a discrete mosaic of patches (Evans and Cushman 2009). These maps are different from those presented in this study because they are based on assigning locations into discrete patches in which a given patch is believed to share the same type, rather than a continuous probability of occurrence. Such classified patch-scale maps of vegetation have long been the foundation of natural resource management and the science of landscape ecology (Evans and Cushman 2009). However, it is more appropriate to represent residual patch occurrence as continuous surface (i.e., probability of occurrence and occupancy), rather than as a mosaic of discrete patches. The approach adopted in this study is spatially explicit, and has power to predict probability of occurrence of residual patches continuously across the fire events. It was also reported that the evaluation of predictive performance has already made an important contribution to the application of distribution models in regional conservation planning (Pearce et al. 2001).

Broadly speaking, the prediction accuracy based on the RF model was able to discriminate residual from null-residual patches, with significant AUC values. An important question is how well the model predicts residual patches beyond the extent of residual and null-residual patches (i.e., prediction not constrained to a fire footprint). The model was applied to generate 'potential' residual patch distribution maps within the fire footprint; such maps are cartographic representations of the probability of occurrence of residual patches. Despite the variation in the prediction accuracy, visual interpretation of the predicted maps showed that the model was able to identify potential areas where residual patches are likely to occur. This suggests that the variables incorporated in the study were good indicators of residual patch occurrence. Yet, residual patches tended to occur substantially in areas dominated by wetlands This supports the findings reported in the previous chapter where distance to wetlands was found to be more informative to explain residual patch occurrence. This is not surprising in the sites where the abundance and distribution of wetlands was prevalent, and is one of the driving mechanisms for residual patch distribution. In a similar study, the importance of the topographic variables was less informative. Topographic variables such as slope, RI and elevation were computed from coarse spatial resolution data, which could be too coarse to be useful at the event level; the data may be useful when measured at the fire regime scale. In summary, visual inspection of the probabilities underlying each prediction indicated that 1) high prediction probability was associated with the existence (abundance) of wetlands; 2) residual patches cannot be retained within surface water, and this has successfully been reflected in the predicted maps where surface water were associated with low (zero) probability values; and 3) high probability values were observed in areas where residual patches do not exist; residual patches tended to occur in areas beyond the extent of the observed residual patches.

4.4.5. Future directions

The distribution of residual patches seemed to be associated to all environmental variables, but the study revealed that the marginal effect of wetlands was prevalent. The validity of the model is likely to be affected when it is applied to an environment where the abundance (and distribution) of certain parameters (e.g., wetlands) is limited. Therefore, questions remain on the response of the residual patches to other environmental gradient and whether variables not represented in the study but may have an impact on model's predictive power. The development and addition of other predictors that may explain post-fire forest characteristics would improve the model. Frescino et al. (2001) indicated that integrating human attributed variables (e.g., harvesting operations and wildfire suppression) can affect the prediction of post-fire forest attributes. In a study by Wilfried et al. (2003), it was also noted that the predictive accuracy of the model was attributed to the human impacts on species distribution, including forestry activities and fire suppression. However, the study sites are not actively suppressed or harvested; hence adding human related predictors do not affect the effectiveness of the model. Therefore, the focus should be on other environmental variables such as forest age composition, fuel type composition, and weather variables that would show variability at local geographical scale.

Moreover, the results of the study showed that a predictive model based on RF algorithm is flexible enough to identify the potential areas where the residual patches would potentially occur. This confirms the findings of other studies (e.g., Mellor et al. 2013) that the ensemble classifier can be used to learn complex non-linear relationship. However, the performance of the RF model is based on the assumption that training data is representative of residual and null-residual patches across the study area, and is statistically independent of the test data. An important next step in assessing the model's performance is to undertake an independent assessment of the implemented model from sites located in a different ecoregion. The eleven fire events studied in this section are located within the same ecoregion, and hence the model is implemented in a different ecoregion for independent assessment of RF model. This would improve our understanding of the characteristics of the fire events, and model's robustness to predict spatial elements in different ecoregions. Therefore, the next chapter focuses on the application of the RF model developed in this study in different ecoregion (i.e., an 'independent' site).

4.5. Summary and conclusions

Satellite-derived information pertaining to the presence of residual patches can be obtained using different remote sensing techniques. These data (i.e., presence-only) are often used as a base for spatial prediction, but the use of a predictive model using presence-only data has failed to provide a better predictive performance. The presence-absence data are required to successfully apply a model because these reflect the natural distribution of spatial elements. However, information related to absence data is not readily available; this prompted the need to apply a computer simulation approach to algorithmically generate 'pseudo' absence (i.e., null-residual patches). Yet, models designed based on presence-absence are profoundly affected by class imbalance, which eventually over- or under-estimate the majority and minority classes respectively. RF model has the option to balance the error rates in unbalanced data by either adding class weighing parameters or sampling techniques (down- or over- i.e., RF) based on presence-absence data (for better predictive performance). In this study, a simulation algorithm was rather developed to extract ‘pseudo’ absence (i.e., null-residual patches) within the burned landscape. Also, to avoid the effect of class imbalance, the absence-data (i.e., null-residuals) were extracted in a way they mimic the presence (residual) data in size, shape, and number. Using the presence-absence data (residual and null-residual patches) and environmental variables, a presence-absence model was implemented to identify the potential areas where residual patches are likely to occur.

One of the goals of this analysis was to develop a spatially explicit model for predicting residual patch occurrence and examine its predictive performance for producing pixel-scale maps of probability of occurrence. Given that goal, the RF was an effective classifier in predicting the probability of patch presence in response to the environmental variables considered. This is in spite of the low prediction accuracy observed in some instances (either due to small sample size or low abundance of some of the variables that explain the residuals). This reflects the view that a model with good discrimination ability is the one that correctly discriminate between presence and absence in the evaluation dataset, irrespectively of the reliability of the predicted probabilities (Pearce and Ferrier 2000). This study generated spatially explicit probability maps to identify areas where residual patches are likely to occur. Based on the findings, it can be inferred that 1) given the predictor variables, the predictive performance of RF model was reasonable enough to determine the occurrence of residual patches within a burned landscape, 2) the identified variables did a reasonable job to explain residual patches and model was flexible enough to identify potential areas where residual patches are likely to occur, 3) high prediction probability of residual patches is likely to occur within or in closer proximity to wetlands, 4) for a presence-absence model to predict residual patches sufficient sample is required as insufficient sample affects the performance of a predictive model, 5) the predictive power of the model decreases with increasing grain size; yet a robust scaling that determines the patterns across the gradient of scale for all fire events may not be established, and 6) for a model to predict spatial elements with an excellent discrimination ability (strong model), the parameters that determine the residual patches should occupy a considerable portion of the fire footprint.

Therefore, having all the desired parameters for prediction, the RF model was a robust modeling approach for predicting residual patch distribution from presence/absence data, which is in agreement with previous studies undertaken based on RF models (e.g., Edwards et al. 2007; Evans and Cushman 2009; Dahinden 2011; Mellor et al. 2013). For all the merits of RF in prediction, its interpretability is limited; it is a black-box and does not provide set of rules that are often obtained from CART (Evans and Cushman 2009). However, RF excels at identifying predictor variables and visually characterizing the relationship between predictor variables and predicted classes (Hastie et al. 2001). The ability and validity of a predictive model, given certain variables, to identify potential areas where residual patches are likely to occur is useful to 1) improve our understanding of the characteristics of post-fire residual structure in natural conditions, 2) inform management about the forest management activities that should be undertaken, including the type and nature of the post-harvest residual patch retention, and 3) provides policy makers guidance for emulating fire disturbance patterns with forest harvesting operations.

5. An evaluation of the predictive performance of the Random Forest model for residual patch existence in the Red Lake Fire, Ontario

Abstract

The study presents a method for developing spatially explicit probability maps for the existence of residual patches within a burned landscape. Using the Random Forest ensemble method, I develop a set of rules that explain residual patch occurrence based on selected predictor variables. I then implement the rules (akin to inverting the learning algorithm) to build maps of likely residual stand locations. Initially, satellite derived data from eleven fire events (from the same ecoregion) are partitioned into training and validation using a hold-out approach. The performance of the model is then assessed using independent data from the extensive RED084 fire event that was not involved in training the predictive tool as validation data while data records from 11 fire events are used for developing the model. The model is assessed using a fixed-probability threshold and threshold-independent measure at five spatial resolutions (R_4, R_8, R_{16}, R_{32}, and $R_6$4 m). The model has a reasonable or high predictive performance ('marginal' or strong' model outcome) for most of the fire events within the same ecoregion. However, the predictive power of the model is low for the independent fire event (RED084). Additionally, the patterns of the residual patches and the importance of various factors that explain their existence are assessed. Similar to the previous study, the responses of the landscape metrics are grouped into three categories: monotonic and predictable response, monotonic change with no simple scaling relationship, and non-monotonic change with erratic responses. The results also indicate that the predictors tend to interactively affect the residual occurrence, but natural firebreak features, specifically wetland and water, are among the most important predictors.

Keywords: RED084, fire disturbance, residual and null-residual patches, predictor variables, random forest, predictive models, predictive performance, and spatial prediction

5.1. Context

5.1.1. Study area

In 2011 Ontario encountered an extreme fire season, with 1334 fires disturbing over 6,354 km^2. This exceeds the total burned area between 2000 and 2010 (6,312 km^2), with most of the fires occurring in northwestern Ontario (OMNR 2011; Paysal et al. 2011). One of the largest fire incidents (i.e., the focus of this chapter) is the RED084 fire event, northeast of Ear falls. Based on the classification of terrestrial ecoregions (Hills 1961), the RED084 falls within the Cat Lake ecoregion – 3S (Figure 5.1), which contains 6.7% of the province land area (William et al. 2009). The region is mainly dominated by coniferous forests in which 29.9% of the region is occupied by this cover type; followed by sparse forest (23.5%), water (14.8%), mixed forest (10.2%), and treed wetland (4.3%) (William et al. 2009). The area is also contained with Ontario's forest management zones in the AOU, areas where legal harvesting operations area permitted (Figure 5.1).

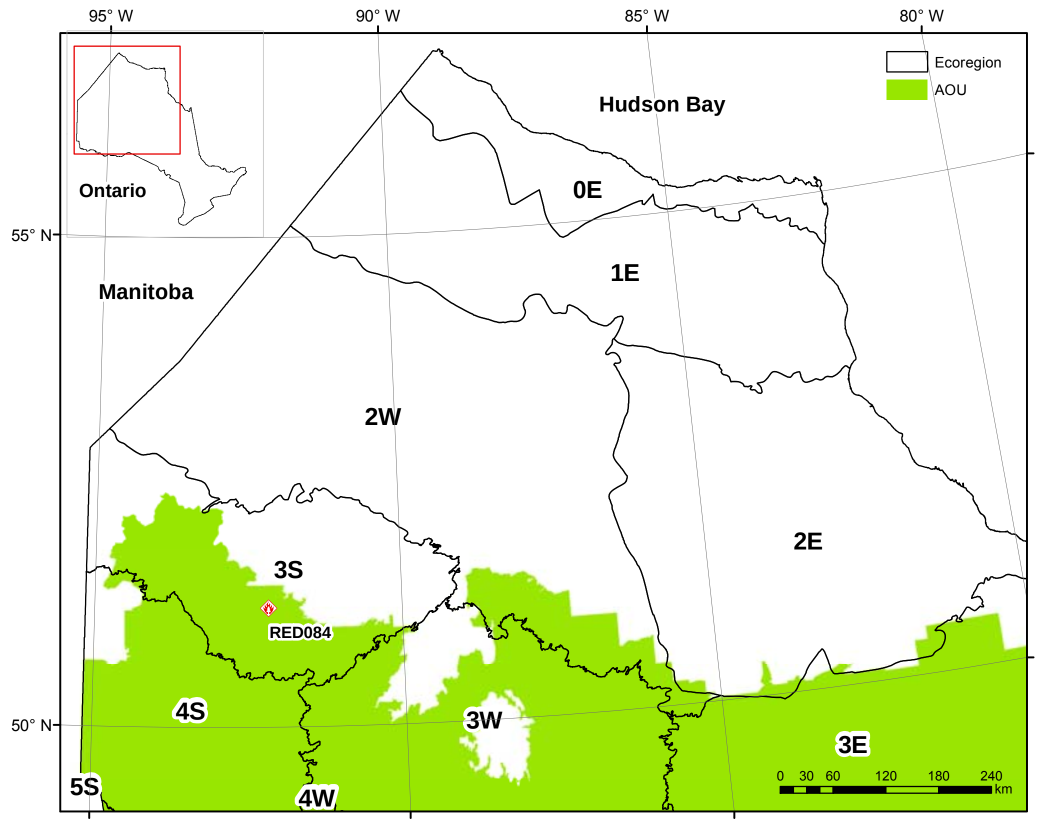


Figure 5.1. Location of the RED084 fire event in relation to Ontario's ecoregion boundaries and the Area of Undertaking (AOU) within which forest management zones exist.

The area around the RED084 is covered mainly by mature coniferous forests, typified by black spruce and jack pine. Small pockets of balsam fir, trembling aspen, mixedwood forests and treed bog are also scattered throughout the area. The area was mostly covered by 80 to 120 year old trees interspersed with small areas of older or younger stands (Baysal et al. 2011). Additionally, the region is sparsely populated by humans, and it is within the AOU where forest management practices are permitted. There is also an increasing interest in the expansion of forest management activities with roads and harvesting operations are spreading.

5.1.2. Red Lake fire event

The ecoregion, in which the RED084 is contained, often experiences intense and frequent fire disturbances as the landscape is characterized by shallow substrates and a periodically dry climate (William et al. 2009). Lighting ignites most of the fires in the region, including the RED084; the probability of human-ignited fires is relatively rare (Baysal et al. 2009). The RED084 fire was reported on 10 July 2011, northeast of Ear Falls, and efforts have been made to suppress it. Despite the effort s, the fire continued to burn for a month until it was naturally extinguished on 2 August. The fire occurred in an area that had experienced a small fire event in 1999 in which an area of 56 km^2 was burned northwest of the RED084.

The weather in the region (northwestern Ontario) is generally characterized by long, cold, and dry winter, and short, warm and moist summers, with annual temperature that varies from 3°C to -4°C and mean annual precipitation ranging from 600 to 900 mm (Runesson 2011). The summer rainfall is between 244 and 299 mm (William et al. 2009). However, the weather record from the closest station in the 30 days before the ignition indicated that the temperature ranges from 18°C to 27.5°C, with 27.5°C temperature record at the time of the ignition. Moreover, little rain was recorded in the month prior to the fire, and a rainfall intensity of 3 mm was measured at the estimated time of the ignition. Such weather conditions influence the spread and intensity of fire considerably. The weather conditions and wind prior to the ignition of the fire create large areas of wind throw (i.e., abundance of dead trees), which contributed to local fire intensity. All these triggered a large fire event with a total burn area of 54,828 ha as mapped by the Ministry of Natural Resources (MNR), including all land and water area inside the mapped fire perimeter. For further information on the nature and patterns of the RED084 fire event refer to (Baysal et al. 2011).

Although the burned area is within the AOU, most of the RED084 area had not been actively harvested. The fire is also within MNR intensive fire management zone, and many efforts were made to suppress it. However, the fire continued to burn persistently until it was naturally extinguished. The RED084 is selected as independent fire event to evaluate the approaches and techniques implemented in Chapters 2, 3, and 4, specifically to test the predictive performance of a model constructed using data records from the 11 fire events. The RED084 fire event is used as validation data because 1) like the 11 fire events, the RED084 is ignited naturally by lighting, 2) despite the attempts to suppress it, it continued to spread until it was naturally extinguished; such naturally ignited and suppressed fires are likely to provide considerable insights to natural fire processes and patterns, and 3) it is contained within a different ecoregion – 3S (different from the 11 fire events which makes the RED084 a potential (independent) site for validating a model's predictive performance.

5.1.3. Research framework

The existence of post-fire residual patches can be mapped using remotely sensed imagery coupled with field observations. This provides only a snapshot of residual patch occurrence, but timely and spatially explicit information on residual patch occurrence is required for effective forest resource management (Beauvais et al. 2006). Studies have indicated that the abundance and distribution of residual patches is explained by complex interactions of various geo-environmental factors (Cuesta et al. 2009; Cui et al. 2009). However, the combined effects of the environmental variables should be examined using statistical models to better understand residual patch occurrence. The use of models to understand the combined effects of the environmental variables provides spatially and temporally explicit assessments pertaining to residual patches. Spatially explicit information related to post-fire forest characteristics is also essential for developing land management policies in forested landscapes and assessing the significance of residual patches. Specifically in Ontario, understanding the characteristics, composition, and spatial arrangements of residual patches and the reason(s) why certain residual patches exist within a fire disturbed landscape has become a primary requirement for emulating forest disturbances, emerging as a general forest management goal within burned landscapes (Perera et al. 2009a). This requires a thorough examination of the pattern and characteristics of residual patches, type and proportion of post-harvest residual patches that have to be retained and the geographical locations where post-harvest residuals are likely to occur. Therefore, this chapter focuses on implementing the predictive model developed in the previous chapter and assesses its validity for explaining the patterns and characteristics of residual vegetation patches in a given landscape. This provides a ground for researchers to consider the repeatable approach implemented in this study to define a set of rules that can be considered for retaining residual patches within harvest units. The approach can also help forest managers to design residual patches within harvesting layout planning to emulate fire disturbances.

Spatially explicit information related to residual patch occurrence can be obtained using different approaches (e.g., predictive modelling techniques) and this is becoming an increasingly important tool in natural resource management (Guisan and Zimmerman 2000; Beauvais et al. 2006). The approaches often employ statistical techniques to model the presence and absence of residual patches at observed sites in relation to environmental variables; thereby allowing the predicted probability of occurrence of spatial objects at unobserved locations or extrapolated across large areas (Pearce and Ferrier 2000; Zaniewski et al. 2002; Anderson et al. 2003; Brenning 2005). However, for such models to be useful for resource management their predictive performance has to be evaluated as accurate predictions are always necessary. The level of error associated with predictive models can be identified and quantified using independent data;

data not used in the calibration of the model. This is a vital step in the process of model development, and assists in determining suitability of a model for specific applications, comparing modelling techniques, and identifying aspects of a model that need improvements (Peace and Ferrier 2000, Pearce et al. 2001).

In the previous chapter, a predictive model based on RF was constructed using data records from 11 fire events contained within the same ecoregion. The model was constructed and evaluated using a hold-out approach; given the 11 fire events, the data records from an individual fire event is hold-out for testing while the data records from the remaining 10 fire events are used for calibrating the model. In this chapter, the predictive accuracy is evaluated using independent data from the RED084 fire event. The predictive model is implemented to address the following objectives: evaluate the predictive performance of the model using independent data; how well the model predicts the likelihood of residual patch occurrence in the RED084 fire event, and generate spatially explicit probability maps and provide information about the distribution of forest stands that escaped burning. I aimed to determine which of the predictor variables have the greatest influence on the occurrence of residual patches. Moreover, I hypothesized that 1) a scaling rule can be established to characterize the patterns of residual patches and compare measures of landscape metrics across multiple scales; 2) the occurrence of residual patches in the RED084 is associated with closer proximity to natural firebreak features, and 3) the existence of residual patches are also related to certain land cover types that are dominant in the landscape or less prone to burning.

5.2. Methods

5.2.1. Remote sensing data

Understanding the patterns and behaviour of wildfire forest disturbance and residual patches is recognized as being instrumental for ensuring effective forest management practices (Linke et al. 2007). This requires knowledge of post-fire forest characteristics (e.g., land cover or disturbance mapping) and the corresponding spatial arrangements (Clark and Bobbe 2007). Remote sensing has become an efficient technique for mapping such disturbance patterns and deriving spatially and environmentally relevant variables for predictive modelling. This has become crucial for resource managers as forest managers are increasingly relying on remotely sensed data for mapping and characterizing post-fire forest characteristics (Clark and Bobbe 2007; Coops et al. 2007). Moreover, ecological processes (i.e., in the form of wildfire disturbances) create certain patterns and heterogeneities in a landscape. This has often been assessed with landscape metrics, which are designed to characterize the geometric and spatial properties of mapped patterns (McGarigal et al. 2002). Remote sensing techniques have also been used for preparing classified or thematic land cover maps and examining spatial patterns.

A study has undertaken to map a footprint describing an individual fire event and develop an approach for extracting post-fire residual vegetation patches within the fire footprint (Remmel and Perera 2009). The study used high-spatial resolution satellite imagery that considers the actuality of gradual boundaries by assessing the fire-membership strength of each pixel prior to developing a footprint. Similarly, the perimeter of the RED084 fire event and the associated post-fire residual patches was investigated using Ikonos imagery. The spatial resolution of Ikonos provides high-resolution imagery (4 and 1 m in the multispectral and panchromatic bands respectively) and swath width = 11.3 km. The Ikonos images also have four multispectral bands corresponding to blue (455- 520 μm), green (510-600 μm), red (630 – 700 μm), and near infrared (760- 850 μm) channels. This level of detail would help mapping the heterogeneous (wildfire-disturbed) landscapes and characterizing the patterns of post-fire residual patches.

In order to map the extent of the fire footprint and the unburned areas within the fire perimeter, six Ikonos image scenes (Figure 5.2) were used; four image scenes were captured between 22 and 30 October 2011 while the other two image scenes were acquired on 12 July 2012. The images from 2012 are cloud free, and were incorporated to avoid excessive cloud cover of the 2011 images (scenes 3 and 4) (Table 5.1). The Ikonos images were supplied in GeoTIFF format, with each spectral band forming a unique file with the geospatial reference information. All images were registered to the Universal Transverse Mercator (Zone 15) projection using the 1983 North American Datum.

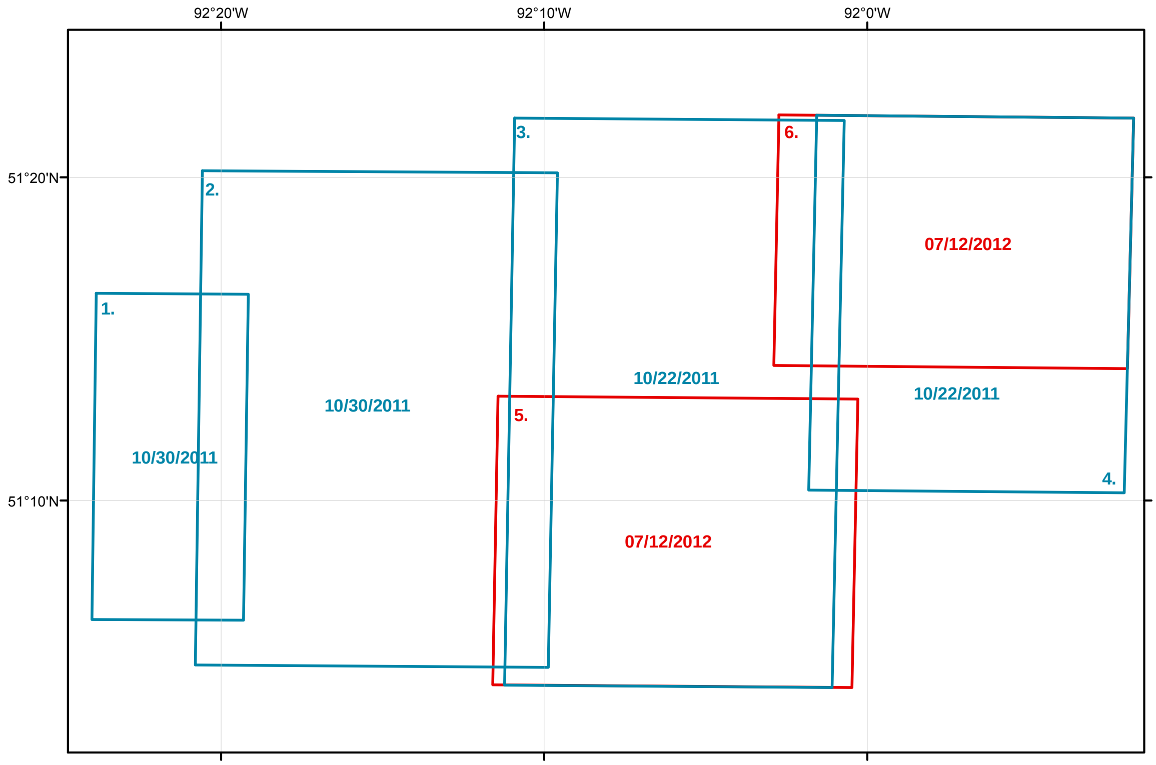


Figure 5.2. Outlines of the six Ikonos image scenes used to generate the extent of the RED084 fire footprint. The blue outlines show the scene for images acquired in 2011 while the red outlines are the extent of the cloud free Ikonos images from 2012.

Table 5.1. Atmospheric parameters used to apply Ground reflectance atmospheric correction module.

Image scene	Acquisition date	Satellite		Sun(solar)			Cloud cover (%)	Atmospheric information		Visibility (km)
		Elevation	Azimuth	Elevation	Solar zenith	Solar azimuth		Definition	Condition	
1	30/10/2011	62.22	154.46	24.41	65.59	168.19	1	Rural	Fall (Spring)	24.1
2	30/10/2011	67.61	144.10	24.39	65.61	168.30	4	Rural	Fall (Spring)	24.1
3	22/10/2011	78.84	176.40	27.32	62.68	170.09	1	Rural	Fall (Spring)	24.1
4	22/10/2011	86.82	134.65	27.28	62.72	170.20	5	Rural	Fall (Spring)	24.1
5	12/07/2012	67.65	293.80	59.65	30.35	161.32	0	Rural	Mid-lat. (Summer)	24.1
6	12/07/2012	64.47	320.24	59.77	30.23	160.84	0	Rural	Mid-lat. (Summer)	24.1

5.2.2. Image processing and classification

Prior to classifying the images into different land cover categories, the two image scenes from 2012 were geometrically corrected using an existing Ontario Base Map data as a reference. The correction was applied to perfectly overlay the 2012 images with the image scenes from 2011, with resulting root mean square error (RMSE) less than 0.5 pixels (1.6 m). Similarly, all the image scenes, from 2011 and 2012, were corrected for atmospheric haze and scattering using the ATCOR ground reflectance atmospheric correction module in PCI-Geomatics, and using the parameters described in Table 5.1. As the fire event was too large to be covered by a single image scene, all the images (six image scenes) were mosaicked to cover the extensive area of interest (RED084). The boundaries of each of the image scene were first edited to provide the cutline boundaries for mosaicking; the image scenes were then clipped based on the newly edited boundary. This is to optimize the efficiency of covering the entire fire event with image mosaicking and remove the excessive cloud cover from some of the image scenes.

An aerial reconnaissance survey of RED084 fire was conducted using an aircraft at a general altitude of 600 m on 26 September 2012, a predetermined flight path, and fire perimeter map (Figure 5.3). A total flight time of 2 hours was conducted to undertake a reconnaissance survey of parts of the RED084; almost one year after the fire was declared under control. The survey was conducted to get an idea of the nature and patterns of the fire (and post-fire forest characteristics) and collect validation data for classification error assessment. The survey was also supported by HD video footage acquired by a GoPro camera mounted under the wing of the aircraft. The visual observations and video recording were linked by GPS. Using GPS triggering, land cover types at 275 additional locations were identified (Figure 5.3) while flying overhead; the 275 sample locations were used for accuracy assessment.

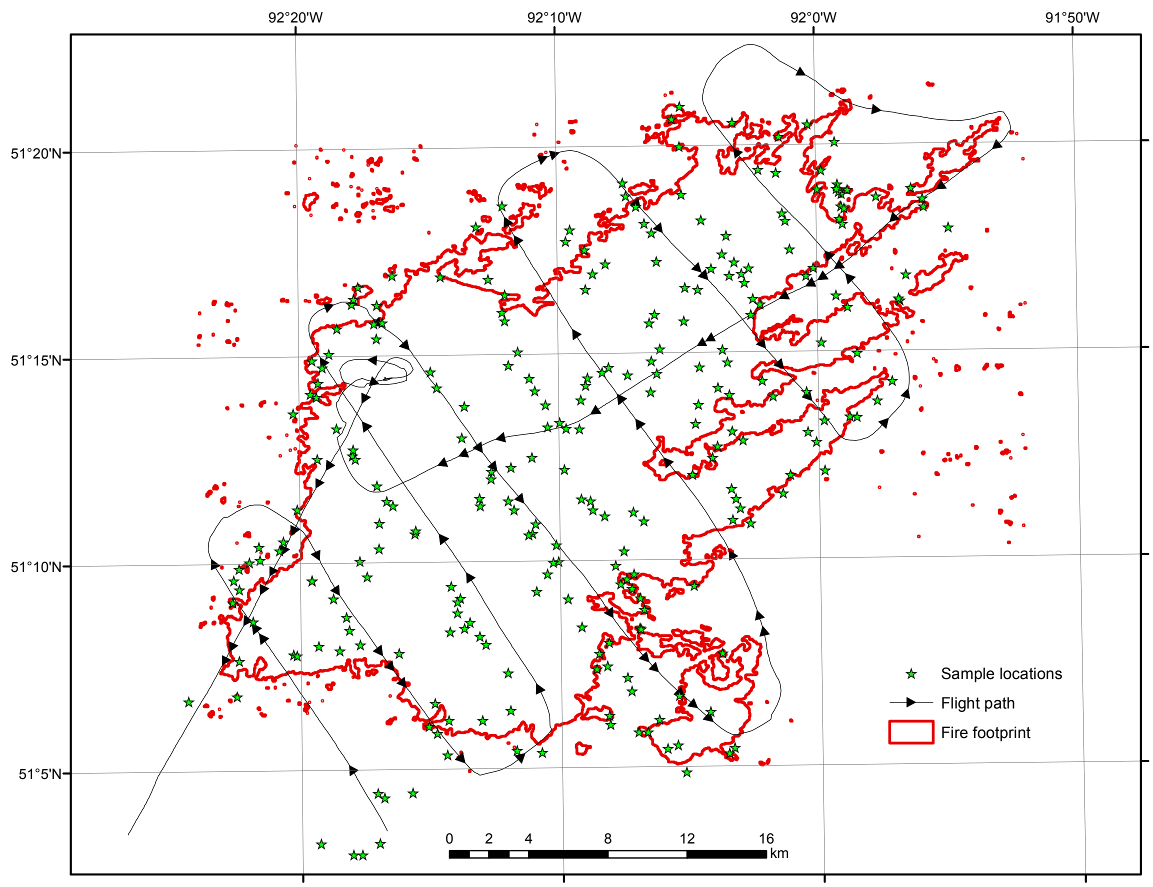


Figure 5.3. The extent of the RED084: the fire footprint, the path of the reconnaissance survey of RED084, and the distribution of the sample locations used for accuracy assessment.

In order to examine the nature of the post-fire forest characteristics and help image classification and assessment with validation data, fieldwork (ground survey) was also conducted in conjunction with the aerial survey from 23-27 September 2012. The burn severity, spatial heterogeneity within burned areas, abundance of unburned areas and natural firebreak features (such as wetlands, surface water, and non-vegetated areas) were recorded. Some photographic evidence of the residual patch occurrence in different physiographic settings is provided in Figure 5.4.

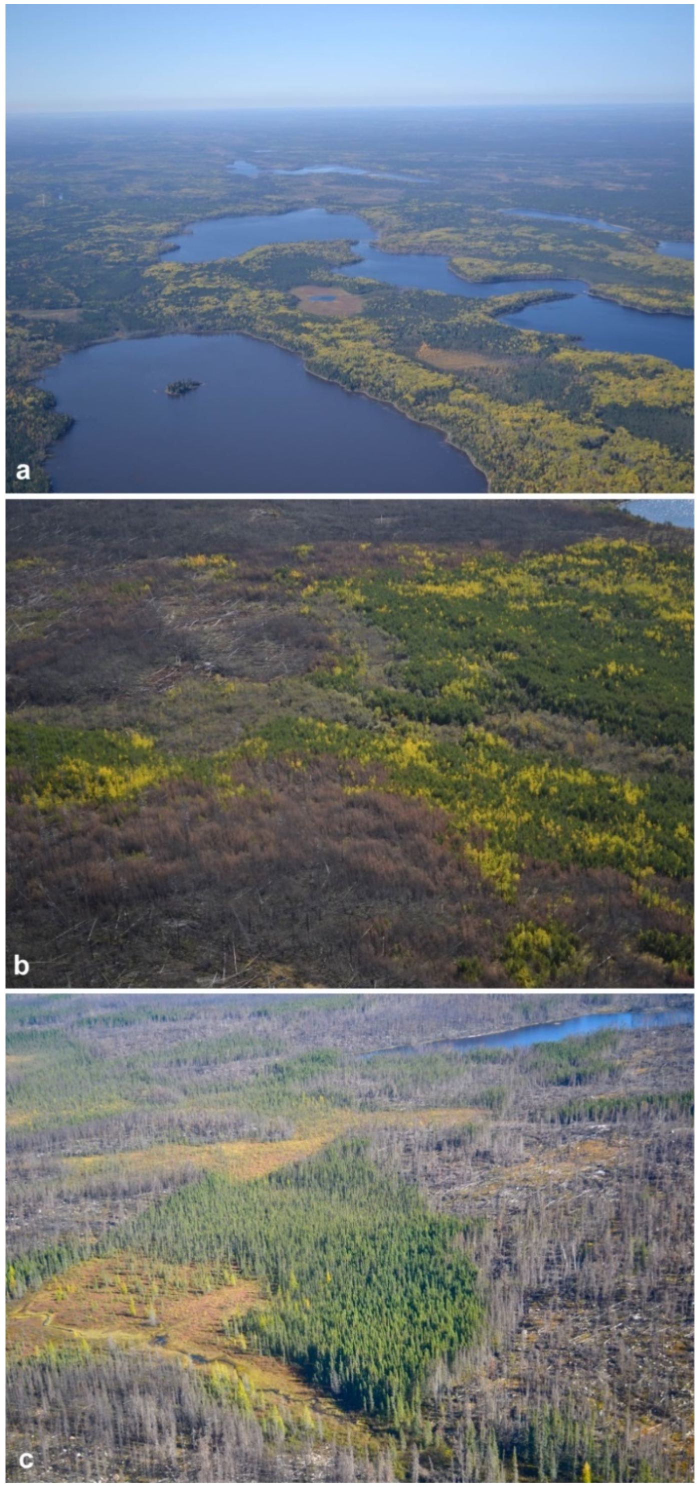


Figure 5.4. Residual patch occurrence in a variety of physiographic settings: a) abundance of natural firebreak features – surface water, b) residual patches of different cover types, and c) residual patches exist in closer proximity to natural firebreak features (e.g., wetlands).

A total of fourteen land cover classes are considered for image classification based on a similar study undertaken by (Remmel and Perera 2009), which is in parallel to the land cover classification schemes of the 2000 Ontario Provincial Land Cover Database (OMNR 2005). A brief description of each of the land cover categories is presented in Table 5.2. The process of extracting surface information from remote sensing data is often performed using image classification approach, which refers to the process of assigning pixels of continuous raster image to predefined land cover classes.

Table 5.2. Categories of land cover obtained from Ikonos image classification. These categories follow the classes of the 2000 Ontario Provincial Land Cover Database (OMNR 2005).

Land cover category	**LCID**	**Description**
Complete burn	CB	Vegetated areas burned over their full extent, showing little or no evidence of vegetation
Partial burn	PB	Vegetated areas burned over part of their extent, showing evidence of sparse or scattered vegetation
Old burn*	OB	Old burns where charring is still evident but regeneration appears
Dense conifer	DC	Dense, predominately coniferous forest that may include some minor component of deciduous species
Sparse conifer	SC	Sparse, predominately coniferous forest which may include some component of deciduous species
Deciduous	DE	Dense, predominately deciduous forest which may include some minor component of coniferous species
Alder shrub woodland	AS	Alder shrubs with some large trees occurring almost exclusively along watercourses
Low shrub	LS	Low shrub areas that may include grasses but do not support trees, found in proximity to lakes, on the deltas of watercourses, and on old burns
Treed wetlands	TW	Bogs and fens with tree cover
Open wetlands	OP	Bogs and fens without tree cover
Water	WA	Surface water; includes some extensive string bogs
Marsh	MA	Inundated areas with emergent vegetation adjacent to surface water
Bedrock and non-vegetated	BV	Areas with little or no vegetation, primarily bedrock outcrop
Cloud and shadow	CS	Image areas containing no usable data because of cloud and shadow effects

*Old burn is obtained based on a pre-fire land cover map.

Initially, unsupervised classification with ISODATA (Iterative Self-Organizing Data Analysis) algorithm was used to group pixels with similar spectral response into unique clusters. This was applied to help identifying certain feature classes for implementing a supervised classification approach. The mosaicked Ikonos image was then classified using a supervised image classification, by which representative image samples (termed “training sites”) of each land cover category were identified and integrated into spectral signature for each cover type. The training sites were identified and collected independently of the Ikonos images using the video footage and visual (field) observations. Spectral signatures were compiled from training sites, indicating representative feature vectors for each land cover class across the fire event. The spectral signature provides the typical spectral responses and their variability for each cover class, so that a maximum likelihood classification (MLC) algorithm could effectively assign each image pixel with the most probable land cover label. The MLC is widely accepted algorithm for image classification; it basically assigns land cover classes to pixels with similar spectral values. The mosaicked image was eventually classified into the 14 land cover categories given in Table 5.2. Once the classified map was generated, a 3×3 majority filter was applied for generalization and smoothing.

Accuracy assessment is a process used to estimate the accuracy of image classification by comparing the classified map with a reference map. This has been an integral part of image analysis as it allows one to determine the quality of the information derived from remotely sensed data. The best way to represent classification accuracy is in the form of an error matrix; using error matrix to represent accuracy is recommended and adopted as the standard reporting convention (Congalton 1991; Foody 2002). In this study, the assessment was undertaken using field observations where 275 sample validation points were collected from the reconnaissance survey and using a standard error matrix. The error matrix can be a starting point for a series of descriptive and analytical statistical techniques (Congalton 1991). The set of accuracy parameters used for assessment are 1) overall accuracy (*PCC*) – ratio of the total number of correctly classified pixels by the total number of pixels in the matrix, 2) producer's accuracy (*PA*) - represents the proportion of a reference pixel being correctly classified, and 3) user's accuracy (*UA*) – the proportion in which a pixel classified on the map actually represents that class on the ground (Congaton 1991).

5.2.3. Fire footprint and residual patch extraction

To mimic the multi-scale analyses implemented in the previous chapters, the classified imagery was subjected to spatial resampling, such that the block size represented the desired spatial resolutions: R_4, R_8, R_{16}, R_{32}, and R_{64}. The classified map was resampled into the desired spatial resolutions R_4, R_8, R_{16}, R_{32}, and R_{64} following the spatial aggregation used in the preview work (Remmel and Perera 2009) and discussed in (§2.2.2). Similarly, this study adopted an approach developed by (Remmel and Perera 2009) to map the extent of the fire footprint and extract the residual patches within the fire perimeter across five spatial resolutions.

OMNR's definition of residual patches comprises both insular and peninsular patches, and are defined based on their size and location in relation to the perimeter of the disturbed area (OMNR 2001; details on residual patch definition along with a graphic depiction is provided in §1.3.2). Since the focus of this study is on insular patches, hereafter described simply as residual patches; I only described the methods of extracting insular patches. The process of delineating and categorizing residual patches is based on an OMNR definition and a step-wise filtering criteria (Figure 5.5). The classified Ikonos image was first regrouped into burned and unburned pixels, and all unburned pixels within the fire perimeter were extracted and it was determined whether they contain burnable cover. This allowed us to generate "vegetated residuals". The "vegetated residuals" were then determined if they occur in clusters rather than individual pixels. Those conjoined pixels occurred in clusters were considered as patches, and this provides an estimate of the total number of residual patches that exist within the fire perimeter. As per OMNR's definition, 0.25 ha is the minimum size criteria for determining a residual patch. Therefore, clustered "vegetated" pixels that met the 0.25 ha threshold and occurred 1 pixel from the fire perimeter were defined as residual patches. This generates the desired residual vegetation patches that are used for further analyses.

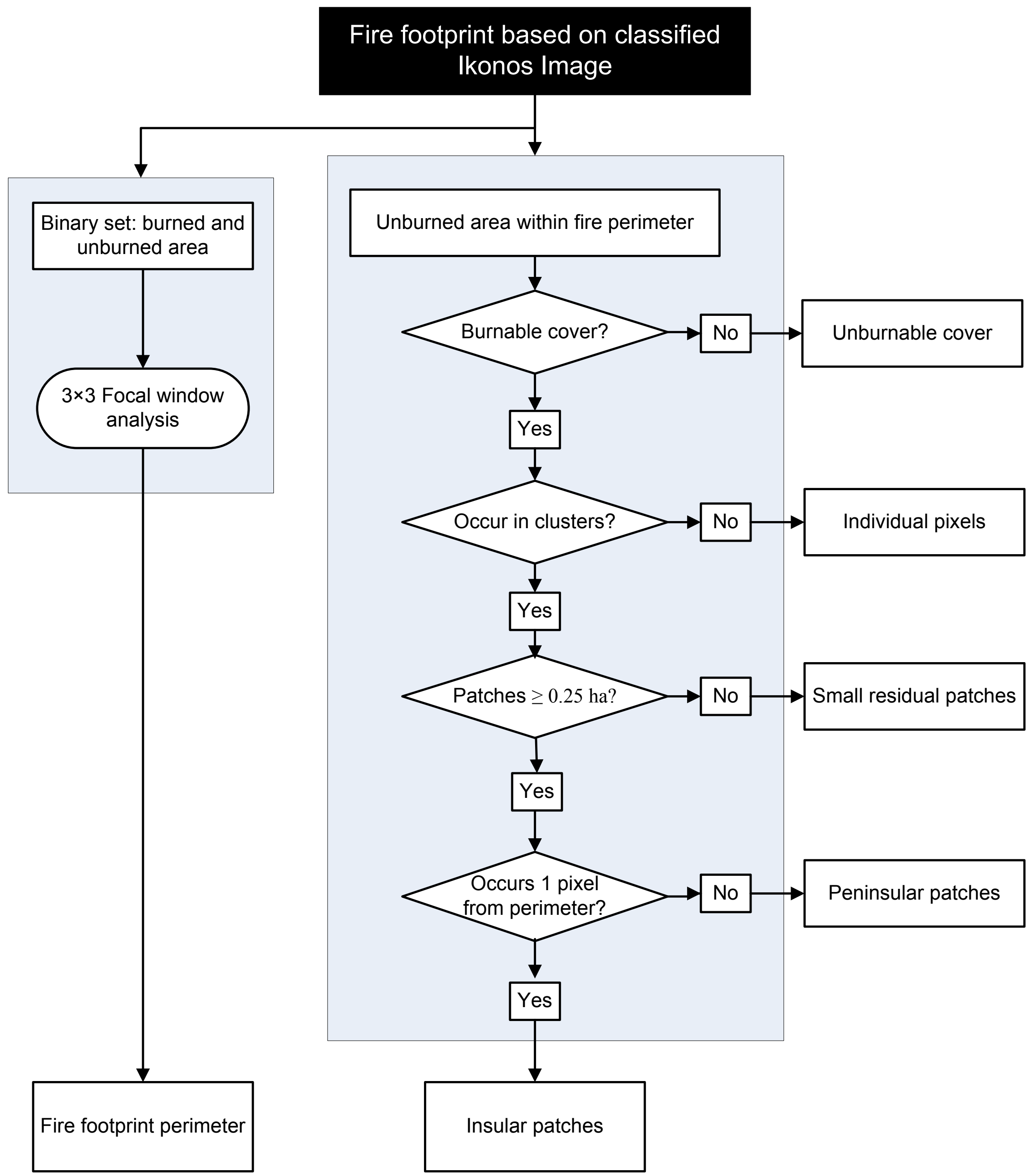


Figure 5.5. Stepwise criteria used to delineate fire footprint and extract post-fire residual patches within a fire-disturbed landscape (Modified from Remmel and Perera 2009).

In order to delineate the fire footprint, the 14 land cover classes were reclassified into a binary set of 0 and 1 where completely or partial burned = 1 while remaining feature classes = 0 for each pixel at R_4, R_8, R_{16}, R_{32}, and R_{64}. The later includes dense and sparse conifer, deciduous, alder shrub, low shrub, treed and open wetlands, water, marsh, and bedrock and non-vegetated areas. A 3×3 focal window was passed over the binary layers of the fire event; at each location, the focal sum was computed as a measure of each pixel's membership strength in the corresponding fire, resulting in values between 0 and 9. Low focal sums indicate few burned pixels within the focal neighbourhood while high focal sums indicate burned pixels surrounded by many other burned pixels. A pixel with a focal-sum value ≥ 1 indicates some probability of fire membership; these locations are coded as 1, and converted to vector representation that represents the fire footprint. A 1-ha filter was applied to eliminate external noise polygons prior to re-rasterizing and producing the desired fire footprint. For further details on the fire footprint extraction process, refer to Remmel and Perera (2009).

5.2.4. Spatial data analysis and modelling

The overall approach addressed in this chapter is based on the techniques and procedures implemented in Chapters 2, 3, and 4 where the spatial patterns and the occurrence probability of residual patches from 11 fire events that lie within a single ecoregion (2W) were investigated. This chapter is intended to apply and evaluate all the approaches on an independent fire RED084 from a different ecoregion (3S).

5.2.4.1. Spatial pattern analysis

The patterns of residual vegetation patches extracted from the classified Ikonos image was examined across the gradient of scales (R_4, R_8, R_{16}, R_{32}, and R_{64}), using 11 landscape metrics. The metrics were applied to explain the effects of scale on pattern analysis and the impact of land cover on residual patch characterization. The ways in which the metrics were computed and examined for residual patch pattern analysis, and the method of assessing land cover impacts on residual patch occurrence are discussed in §2.2.4 and §2.2.5. Moreover, the spatial association of the residual patches of RED084 with natural firebreak features, specifically surface water and the proximity of residual patches to the fire perimeter is investigated. The techniques used to generate concentric buffers and compute the proportion of residual patches within each ring are also described in §2.2.6.

5.2.4.2. Spatial variable importance assessment

The importance and marginal effects of different geo-environmental factors that explain the existence of residual patches in the RED084 fire event is assessed using a RF algorithm. The process is based on the presence-absence model, in which residual patches extracted from the classified Ikonos images are considered as ‘presence-data’ or residual patches. Since information pertained to the ‘absence-data’ are lacking, the algorithm developed in (Chapter 3; Figure 3.2) is implemented to algorithmically extract ‘absence-data’ or null-residual patches. The spatial variable importance assessment involves defining explanatory variables, implementing RF, computing the importance values of the predictors, and their marginal effects. A description of each of the techniques and the way in which they are implemented is discussed in §2.2.

5.2.4.3. Implementing spatially explicit predictive model

A model based on presence-absence data was developed where existing residual patches were considered as presence-data while null-residual patches extracted from a simulation algorithm were used as absence-data. The RF based predictive model was initially constructed and calibrated using a hold-out approach; based on 11 fire events in which data records from an individual fire event was held-out for testing while records from the remaining 10 fire events were used for constructing the model. The way in which the model was calibrated and validated is illustrated in §4.2.5 and §4.2.6. In this Chapter, the model is calibrated and constructed using data records from the 11 fire events while independent data (RED084 fire; an event that was not involved in the learning of the tool) was used for evaluating the model’s predictive performance. The predictive power of the model was evaluated using a fixed probability threshold and threshold independent measures (ROC) (Zweig and Campbell 1993). The method of model calibration and validation is illustrated and presented in §4.2.5 and §4.2.6. Finally, the predicted probability map that show the potential areas where residual patches are likely to occur, is produced as *ASCII* format in *R*, and imported into *ArcMap* for visualization and analysis.

5.3. Results

5.3.1. Image classification and accuracy assessment

Post-fire land cover mapping of the RED084 fire event containing 14 land cover classes is presented in Figure 5.6. Land cover classification includes three burn class features (complete, partial, and old burn), three types of forest class (dense conifer, sparse conifer, and deciduous), five types of non-forest (alder shrub, low shrub, marsh, open and treed wetlands), and two non-vegetated class features (bedrock/non-vegetated and water). Although dense and sparse conifer and burned areas dominate the fire footprint, surface water occupies a considerable area of the fire footprint (15% of the fire footprint). Prior to delineating and extracting the fire footprint and residual patches, the classification errors were quantified and evaluated using an error matrix as illustrated in Table 5.3. The matrix provides some statistical and analytical approaches (i.e., in the form of overall, user's and procedure's accuracy) to examine the accuracy of the classification. The classification process resulted in an overall accuracy of 85% while the overall results of the producer's and user's accuracy range from 25% to 100% (Table 5.3).

The low accuracy is associated with the feature classes that are not abundant in the area (i.e., low shrub and deciduous), and hence did not have enough validation data for a thorough accuracy assessment. A more careful look at the error matrix also revealed that there is confusion in discriminating treed wetlands from sparse conifer land cover types. The similarity in tree cover in the treed wetland makes some degree of confusion between treed wetlands and sparse conifer forest inevitable (OMNR 2005). Similarly, there was a problem of discriminating deciduous forest from dense coniferous forest. Dense deciduous forest is not extensive in the northern latitudes, although poplar and birch are quite common in the region. Yet, small pockets of deciduous were evident within densely covered coniferous species to form pockets of mixed forests (OMNR 2005).

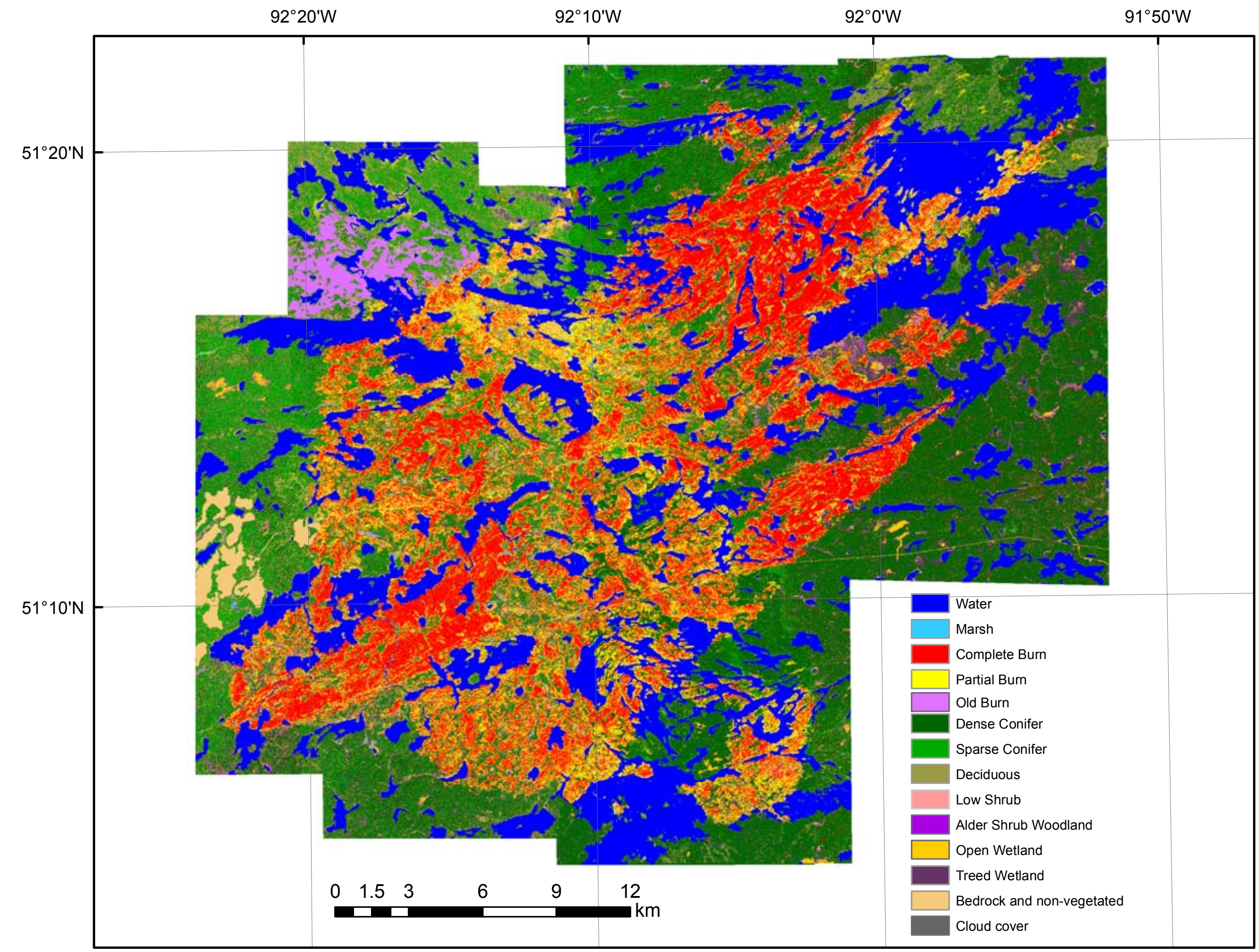


Figure 5.6. Land cover classification of Ikonos image at 4 m spatial resolution containing 14 land cover classes.

Table 5.3. Classification accuracy comparing the classified Ikonos image and the ground sample land cover locations, with overall accuracy, omission error (PA - producer's accuracy), and commission error (UA - user's accuracy).

Reference Data

Classified Data

	BV	**BU***	**DE**	**DC**	**LS**	**MA**	**OW**	**SC**	**TW**	**WA**	**GT***	**UA**
BV	4	2									6	67
BU*	2	127			1		2	4	3		139	91
DE	1	1	2	2	1		1				8	25
DC				49				5			54	91
LS					1						1	100
MA					1						1	
OW							1				1	100
SC		1		4				23	1		29	79
TW		1		2				4	3		10	30
WA										4	4	100
GT*	7	132	2	57	4		4	36	7	4	253	
PA	57	96	100	86	25		25	64	43	100		**85**

*GT - grant total
*BU - burn area which includes both complete (CB) and partial burn (PB)

5.3.2. Fire footprint and residual patches

The fire footprint (defined as the area within the most probable locations of the outer fire boundary) within the classified images, derived at 32 m spatial resolution is shown in Figure 5.7. The footprint is composed of multiple polygons of burned areas, and appeared to be consistent with the fire footprint perimeter mapped by OMNR's as shown in Figure 5.7. This is despite the small discrepancies existed between the derived fire footprints and OMNR's fire maps; this is due to the methodology details of the mapping techniques integrated in the study.

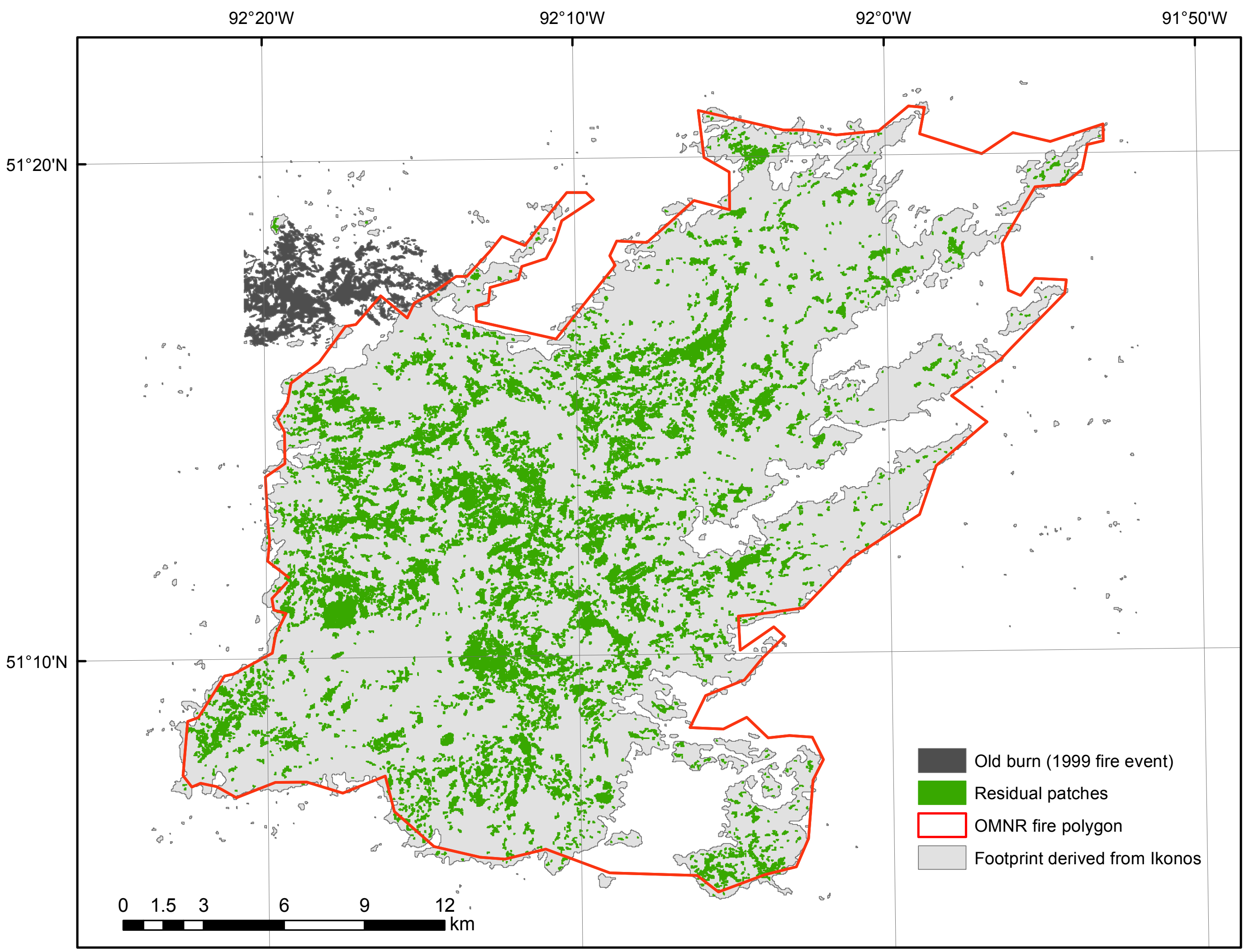


Figure 5.7. The extent of the RED084 fire footprint obtained from the classification Ikonos image and the corresponding OMNR fire region as well the distribution of post-fire residual patches extracted from at 32 m spatial resolutions.

Based on the criteria for defining the residual patches, all the residual patches were covered within the burnable cover types. Within the RED084 fire event, at 4 m spatial resolution, there were 1683 residual patches derived from the classified image, which is almost equivalent to the total number of residual patches from the 11 fire events (1629 residual patches) studied in the previous chapters. The patch sizes range from a minimum of 0.25 ha to a maximum of 224.04 ha, with a mean patch size of 2.97 ha. Most residual patches were small, with 62% of the patches being smaller than 1 ha while only 21% were larger than 2 ha and only 5% were larger than 10 ha (Figure 5.8).

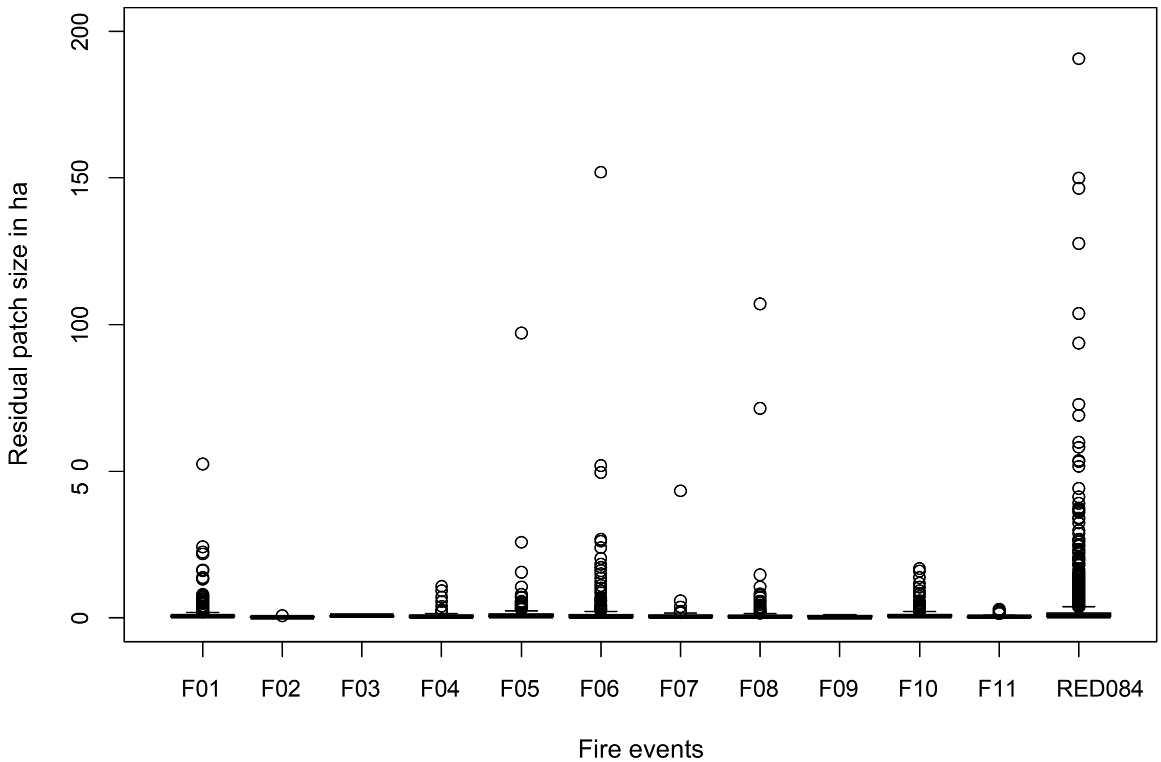


Figure 5.8. A box plot that shows the variability of residual patch size across all the 12 fire events studied in this dissertation at 4 m spatial resolution, each box in the plot is based on the residual patch area computed within fire footprint.

5.3.3. Spatial characteristics of residual patches

5.3.3.1. Spatial pattern analysis: effects of changing grain size

Similar to the study undertaken in Chapter 2, the results of the landscape metric analysis of the RED084 are graphically summarized in the form of landscape metric scalograms, in which pattern indices are plotted against grain sizes. The scalograms are used to assess the scale effect on pattern analysis, and the magnitude and pattern of the response curves in the RED084 were grouped into three general categories: monotonic and predictable response, monotonic change with no simple scaling relationship, and non-monotonic change with erratic responses. The first category (monotonic change and predictable response) includes LPM that decreases with increasing grain sizes in a remarkable consistent power law relationship. Akin to the patterns previously discussed in Chapter 2, three shape related metrics (i.e., LSI, MSI, and FRAC) exhibited a predictable response with a coefficient of determination (R^2) ≥ 97% (Figure 5.9); boxplots are based the LMP values derived from the 11 fire events. The LPM values in the second category increase or decrease with increasing grain sizes, but do not necessarily show a predictable response (i.e., no simple scaling relationship). Landscape metrics such as CA and LPI were expected to increase or decrease monotonically with increasing grain size because of the progressive resampling of smaller patches into larger ones. Accordingly, the two metrics (CA and LPI) tended to show a monotonic and predictable response with increasing grain size (Figure 5.10) with R^2 ≥ 85%. Besides, most of the metrics exhibited an erratic response to scale change, which makes developing a scaling rule difficult.

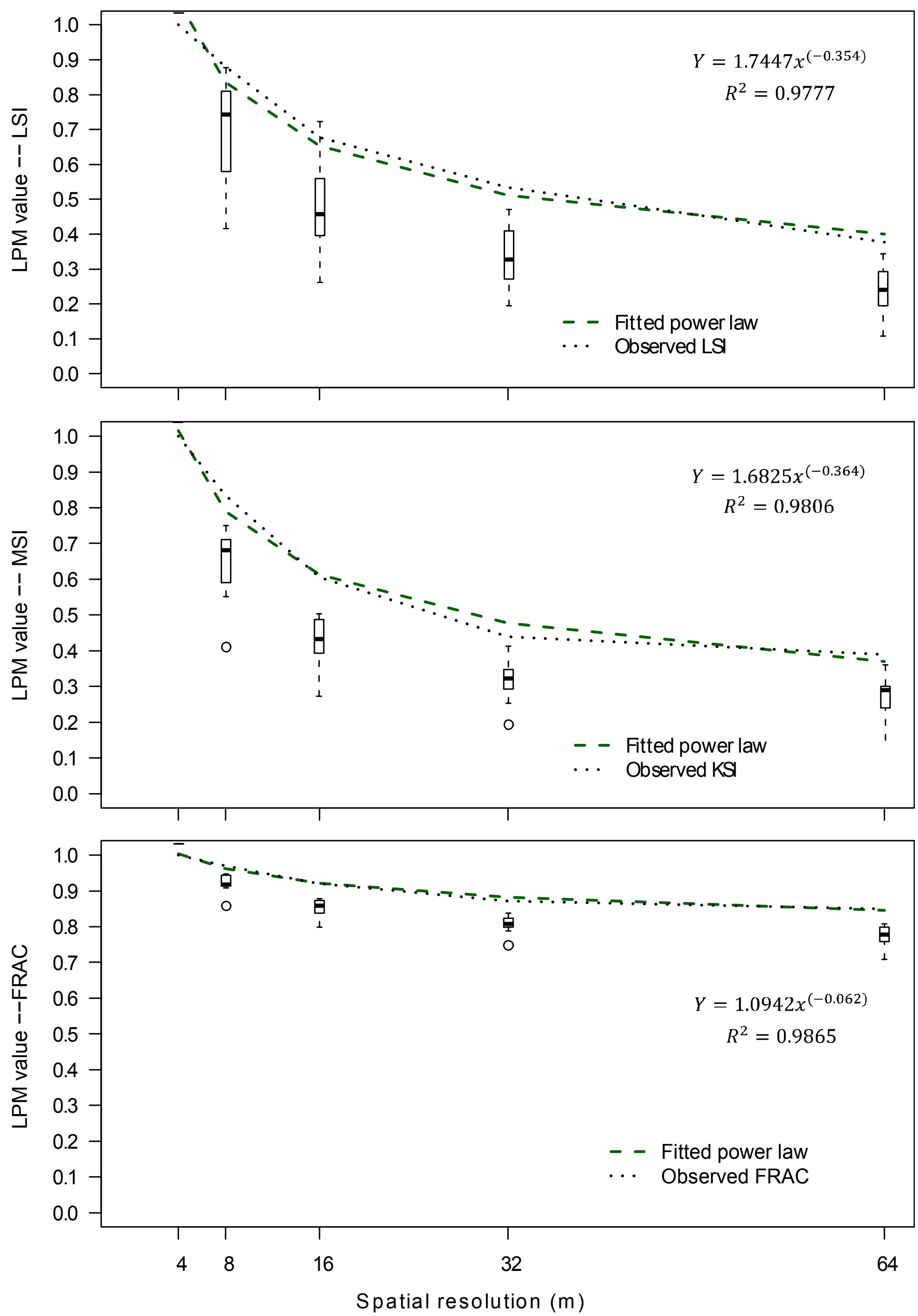


Figure 5.9. Scalograms showing effects of changing grain size on landscape: shape related metrics that exhibit monotonic decreasing function and predictable response with increasing grain sizes. The dot-lines show the observed LPM values for RED084 at different spatial resolutions while the box plots are based on the LPM values obtained from the 11 fire events studied in Chapter 2 (2.3.2; Figure 2.5).

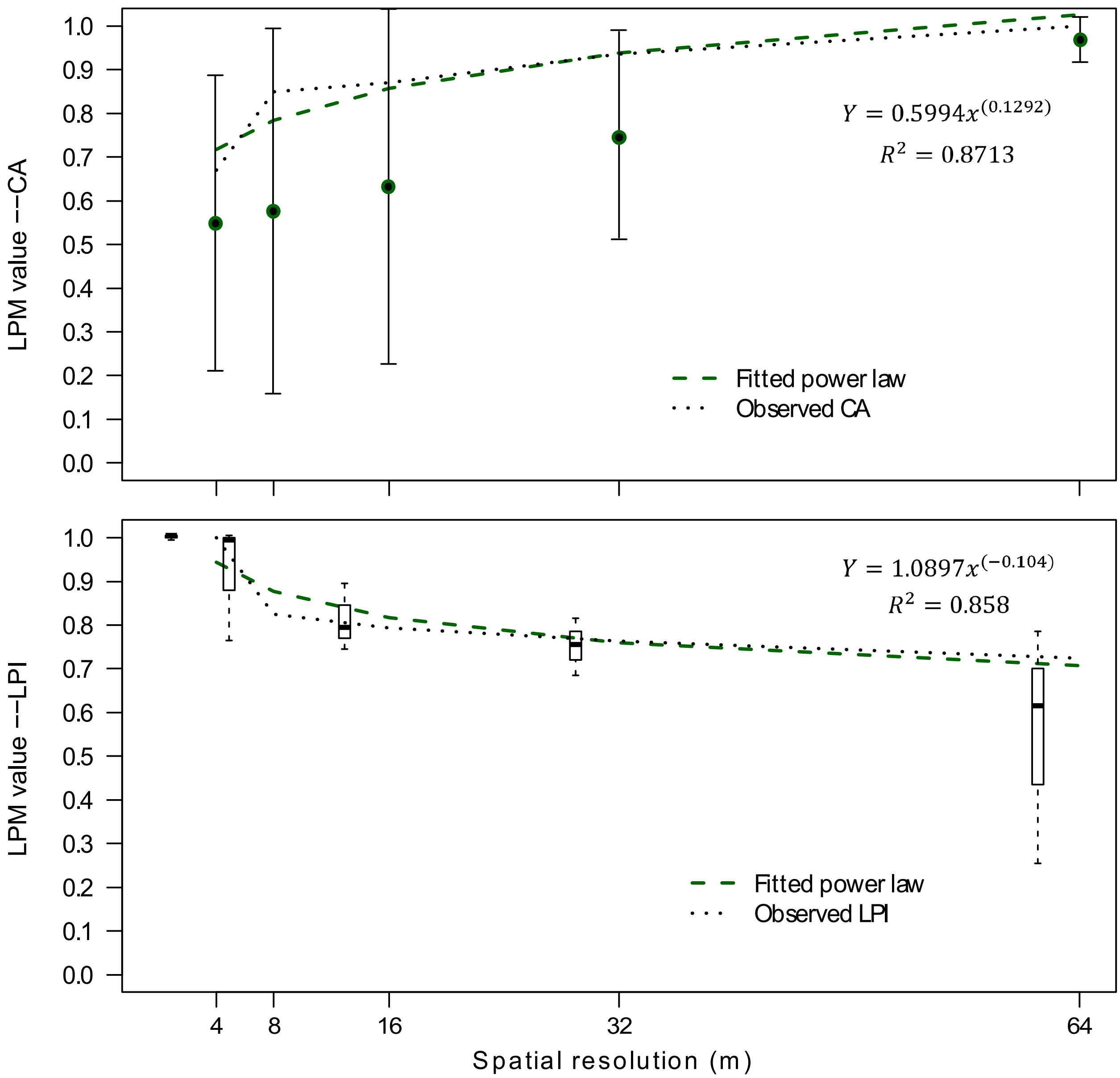


Figure 5.10. Scalograms showing effects of changing grain size on landscape: LPM values exhibiting a monotonic change but do not show a robust scaling rule (i.e., in relation to the metrics studied for the 11 fire events). The dot-lines show the observed LPM values for RED084 at different spatial resolutions while the box plots are based on the LPM values obtained from the 11 fire events.

5.3.3.2. Spatial distribution of residual patches: effects of land cover types

The land cover composition of residual patches, based on selected class-level metrics such as CA and %LAND, were estimated to examine whether particular land cover types were more likely to generate residual patches than other types. Figure 5.11 shows the variability in the proportion of the residual patches occupied by different cover types; box plots indicate the variability across the grain sizes. The results showed that the composition of residual patch area was occupied by different burnable cover types (i.e., dense conifer, sparse conifer, open wetland, treed wetland, deciduous, alder shrub, and low shrub). However, dense conifer, which is the dominant cover type in the fire area, appeared to be prevalent with an over-representation with residual patches, in which more than 48% of the residual patches were occupied by this cover type. Similarly, sparse conifer and treed wetlands were also likely to survive from burning, with around 27% and 8% of the residual patches occupied by these cover types respectively. The remaining cover types (open wetland, deciduous, alder shrub, and low shrub) all together made less than 20% of the residual patches of the RED084.

The fragmentation level of residual patches in relation to land cover types was assessed using two measures of landscape (NP and LPI), which explain landscape heterogeneity. Figure 5.11 shows the variability in NP across 5 spatial resolutions by land cover types. The results indicated that a considerable number of residual patches were occupied by different cover types, and their number is highly variable across the gradient of scales; suggesting the impact of grain size on estimating the LPM values. Additionally, the variability in LPI in relation to the cover types was estimated, and the study indicated that LPI is likely to be associated with the type of land cover types that dominate the residual patches (i.e., higher %LAND values). The trend was similar with the 11 fire events where the largest patch was primarily occupied by the dominant land cover type. Although most of the residual patches were covered by dense and sparse conifer and the largest residual patch is occupied by dense and sparse conifer, the residual patches were distributed through the existing land cover types.

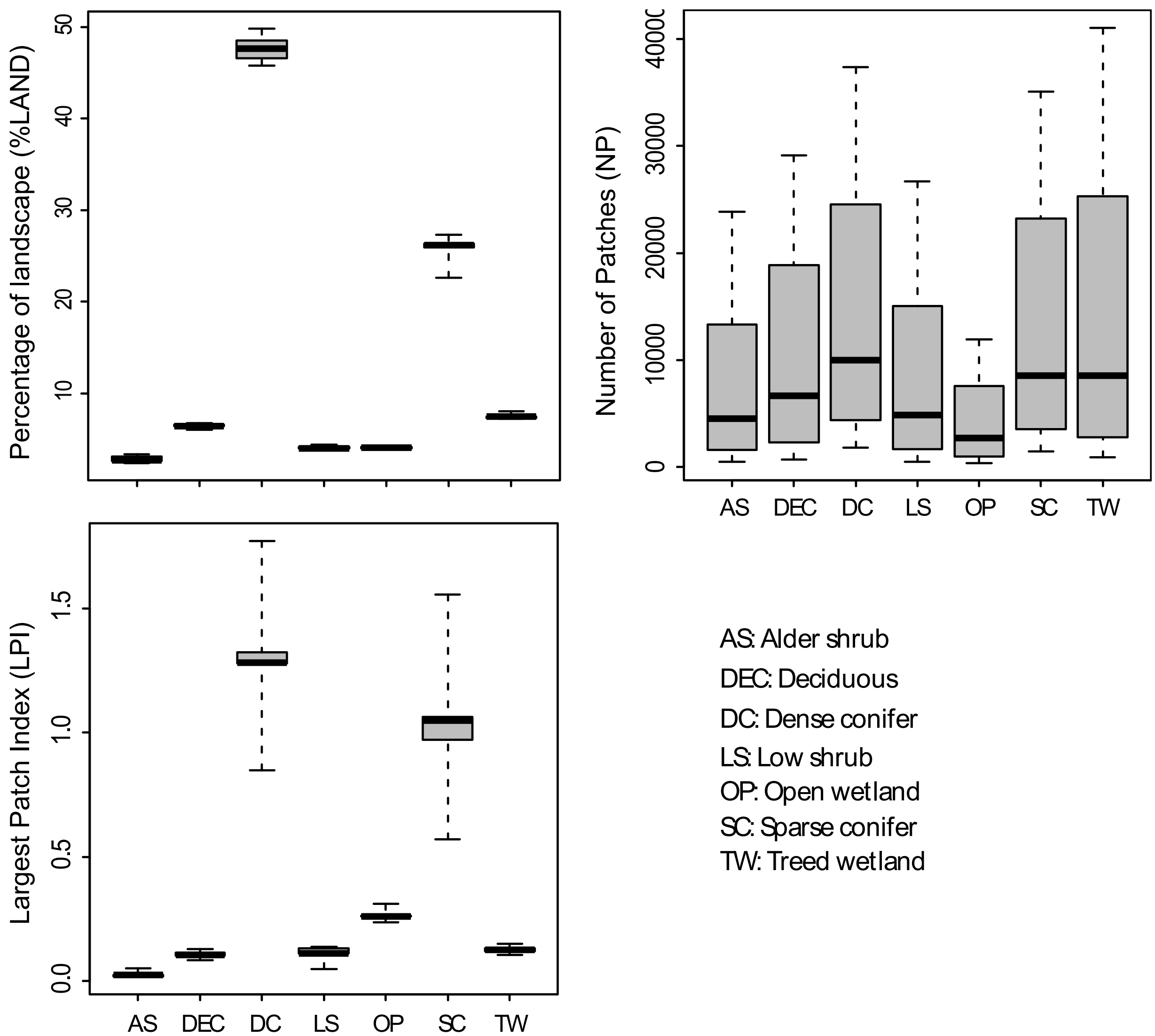


Figure 5.11. Land cover composition of residual patches: 1) the variability in the proportion of landscape (%LAND) occupied by residual patches; the variability in %LAND across the five spatial resolutions by land cover types; 2) the variability in the number of residual patches (NP) occupied by different cover types; each box in the plot is based on the metric values obtained at five spatial resolutions; and 3) the variability in largest patch index (LPI) across the five spatial resolutions by land cover types; each box in the plot is based on the metric values obtained at five spatial resolutions.

5.3.3.3. Spatial association with water and fire footprints

The proximity analysis of residual patch area in relation to distance from surface water and fire perimeter is graphically presented in Figure 5.12 and Figure 5.13. The figures show the relationship between observed mean residual patch area and distance from surface water and fire perimeter. Overall, the residual patch area within each 100-m wide buffer tended to decrease with increasing distance from the water; this is despite some irregularities at certain distance from water (e.g., 400 m). The residual patch extent was higher near the surface water, as expected, and decreased with increasing distance with a linear regression model, with R^2=84%. The same decreasing trend was exhibited for the adjacency analysis with fire footprint. The best-fit model showed that the residual patch area decreased with increasing distance from the edge of the fire perimeter, with a polynomial second order regression and R^2=83% (Figure 5.13).

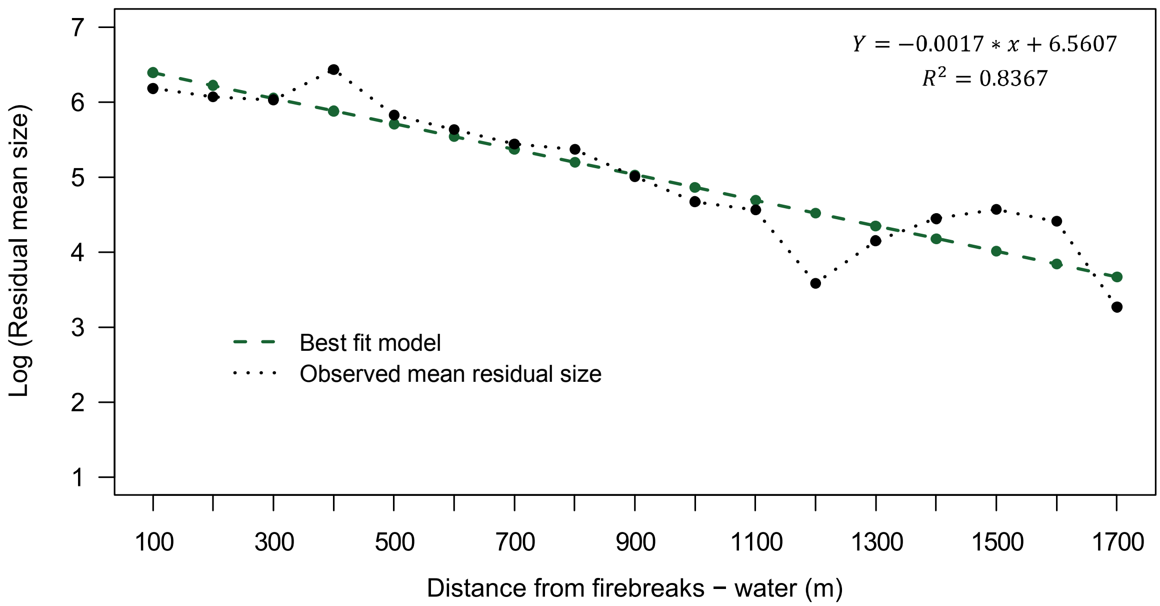


Figure 5.12. The proportionate extent of residual patches in external 100 m wide buffer rings with increasing distance from natural firebreak features (i.e., water). The y-axis shows logarithm of residual patch area.

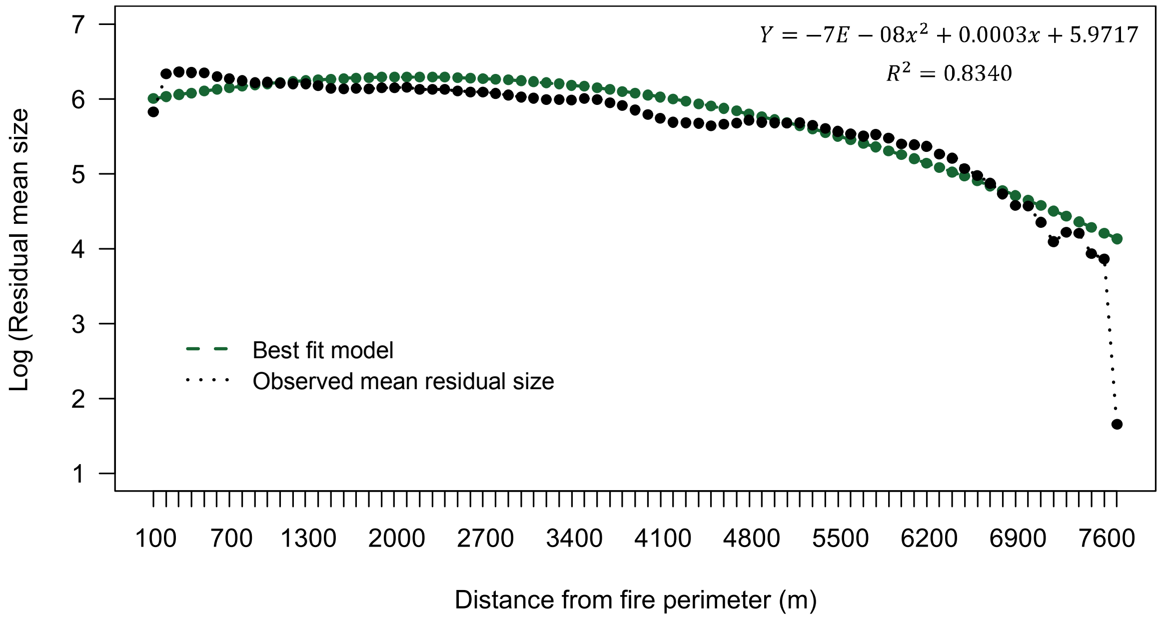


Figure 5.13. The proportionate extent of residual patches in internal 100 m wide buffer rings with increasing distance from the fire footprint perimeter. The y-axis shows logarithm of residual patch area.

5.3.4. Spatial variable importance assessment

The result of the random extraction of null-residual patches for the RED084 fire event along with their corresponding residual patches at 32 m spatial resolution is shown in Figure 5.14. Similar to the 11 fire events, the null-residual patches were supposed to mimic the residual patches in size shape, and orientation but the random simulation was not able to completely mimic all the residual patches. This is attributed to the same reasons described in Chapter four; the presence of substantial (existing) residual patches and unburnable class features (e.g., surface water), which are both not suitable for randomly placing the null-residual patches within the fire footprint. Yet, RF was implemented to determine the most informative variables that explain the existence of residual patches, and assess the marginal effect of the most important predictors.

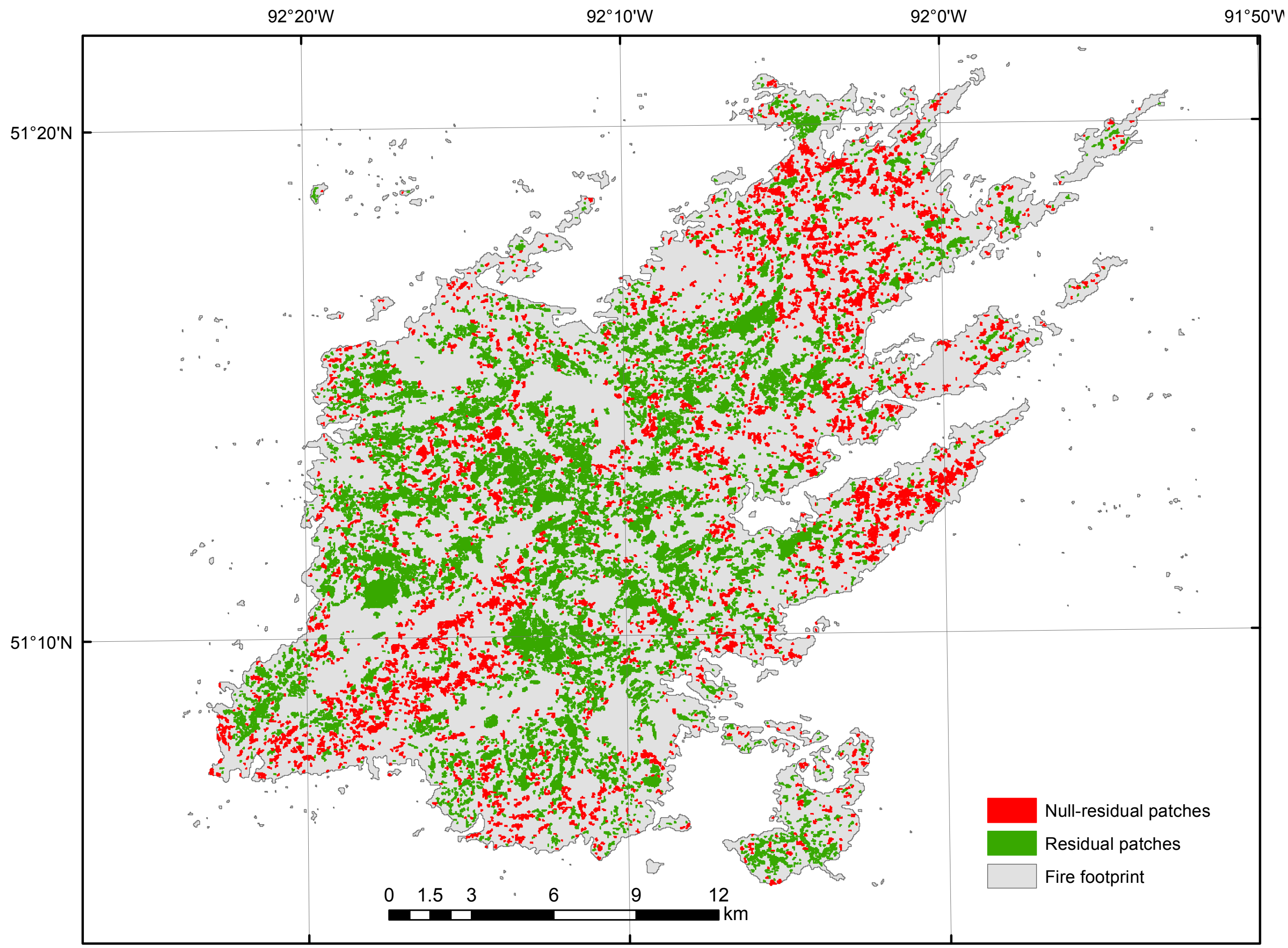


Figure 5.14. The distribution residual patches extracted from classified Ikonos image and the algorithmically simulated null-residual patches for the RED084 fire perimeter at 32 m spatial resolution.

The importance of the predictor variables was computed and ranked in a descending order to determine the predictor(s) that best explain the residual patches. The importance scores of the variables, across R_4, R_8, R_{16}, R_{32}, and R_{64}, are graphically depicted in Figure 5.15. The rule of thumb introduced by Strobl et al. (2009b) was applied to identify the most important predictors that explain the occurrence of residual patches in the RED084 fire event. As stated in Chapter 3, the rule states that a variable is considered informative and important if its conditional importance value is above the absolute value of the lowest negative scoring variable. Additionally, it is stated that the absolute values of the importance scores should not be interpreted or compared over different studies (Strobl et al. 2009a); the assessment of the importance values should be site specific. Nonetheless, the results of our variable importance assessment in the RED084 fire event exhibited a similar trend with the 11 fire events investigated in Chapter 3 where natural firebreak features (specifically wetland) were among the most important predictors.

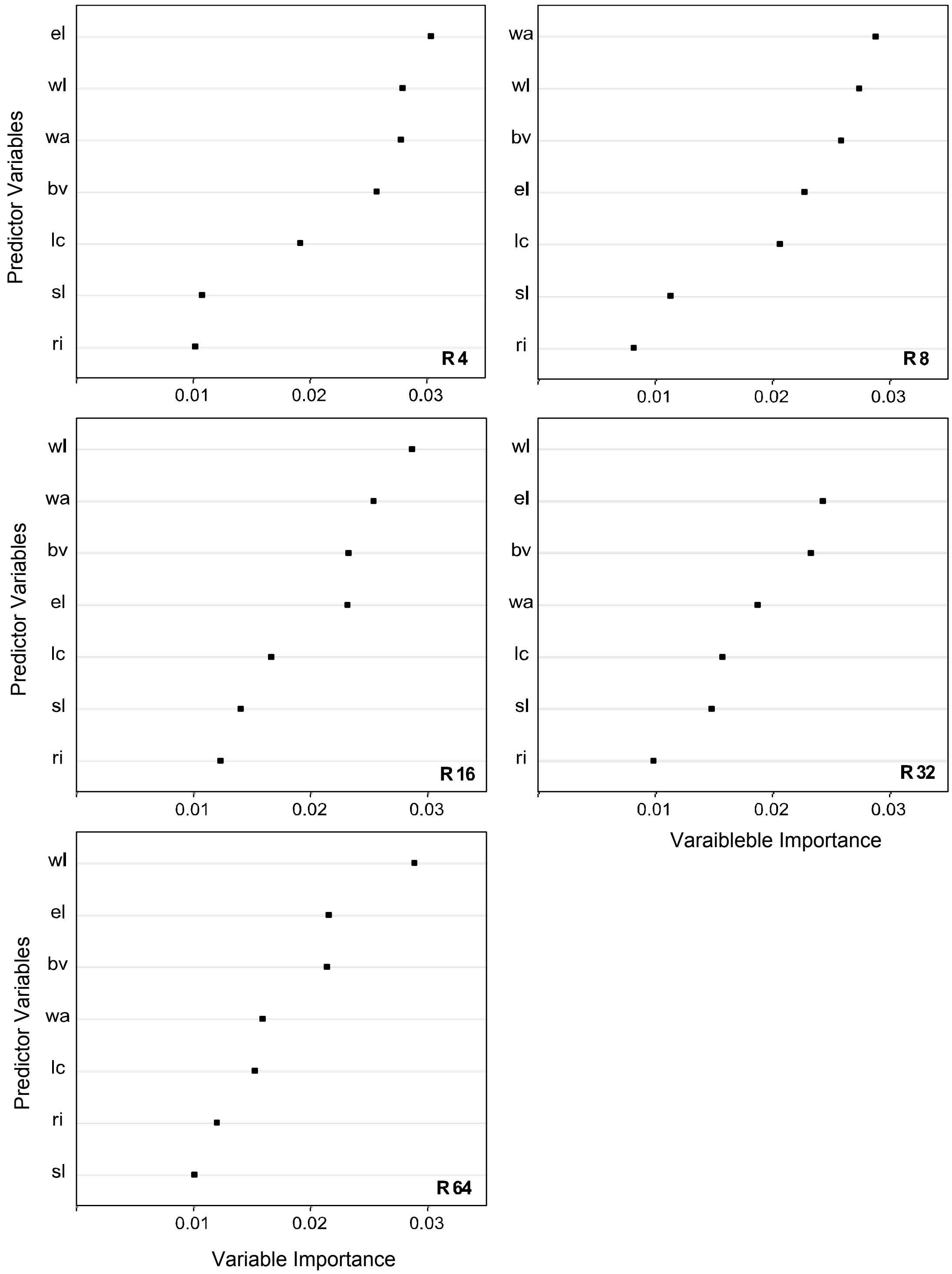


Figure 5.15. Variable importance plots for predictor variables from RF classifications used for predicting presence of residual patches in the RED084. The variable importance score is based on the mean decrease in accuracy. Higher values of mean decrease in accuracy indicate variables that are more important to the classification.

The variability of predictors' importance scores across five spatial resolutions is illustrated in Figure 5.16; boxplots are based on the five spatial resolutions. The slight change in the relative ranking and the variability in predictors' importance values reflect the effect of changing grain size on the importance values; the relative importance measure of the predictor is scale dependent. Despite this, natural firebreak feature (i.e., wetland) remained the most important predictor to explain residual patches followed by water and non-vegetated feature classes. The variable importance measure is interpreted as a relative ranking of the variables.

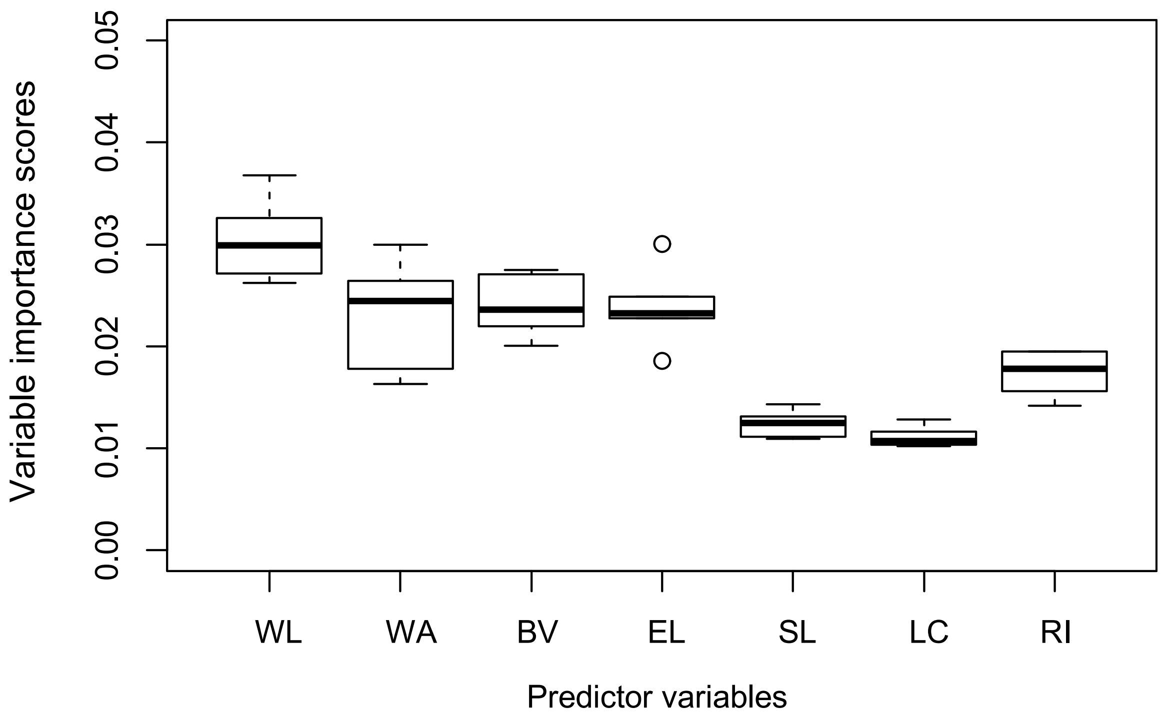


Figure 5.16. Box plots for the relative importance of the predictor variables considered for predicting residual patch occurrence within the RED084; each box in the plots is based on the importance values computed across different spatial resolutions.

5.3.4.1. ANOVA analysis on variable importance

An ANOVA test was also applied to determine if the variable importance scores were significantly different based on the predictors; the test was computed at $\alpha = 0.05$. If there is a significant statistical difference among the means, the null hypothesis is rejected; and hence one of the means is different from the others. However, the test does not tell which groups (variables) are statistically different from one another; it can only tell whether at least one group differs from the others. In this specific scenario, a post-hoc test is required to provide pair-wise tests of mean differences amongst the groups (variables).

The result of ANOVA confirms the view that a significant statistical difference was present amongst the means, allowing us to reject the null hypothesis ($p = 0.05$). Therefore, a Tukey post hoc test was computed to provide pair-wise comparisons of means.

Table 5.4 provides test and *p-values* of the pairs of variables that exhibit a statistical significant difference at 95% confidence level ($p < 0.05$). Despite the contribution each of the variables might have on residual patch occurrence, their importance values are statistically different. This prompted one to look at the marginal effect of each of the variables for residual patch occurrence.

Table 5.4. A post-hoc test based on Tukey test; the table shows the pair-wise variable test that are statistically different. *WL*- wetland, *WA*- water, *BV*-bare/non-vegetated, *EL*- elevation, *SL*- slope, *RI*- ruggedness index, and *LC*- land cover

Pairwise variables	***p-value***
RI-BV	0.0000
SL-BV	0.0003
RI-EL	0.0000
SL-EL	0.0003
WL-LC	0.0000
WA-RI	0.0002
WL-RI	0.0000
WA-SL	0.0010
WL-SL	0.0000
WL-WA	0.0335

5.3.4.2. Marginal effect of predictors on probability of occurrence

The marginal effect of selected predictor variables on class prediction (i.e., residual patches probability of occurrence) is graphically depicted in Figure 5.17. The y-axis of the partial dependence plots is a logit-scale (Chapter 3; Equation 7). The probability of residual patch occurrence in relation to the selected predictors (distance to wetland, water, and non-vegetated areas, and ruggedness index) in the RED084 fire event revealed four different patterns. First, the relationships between distances from natural firebreak features and residual patch occurrence are nonlinear. Yet, high probability of residual patch occurrence is associated with closer proximity to wetlands and surface water; the probability of occurrence tended to decrease monotonically with increasing distance from wetland and water. This supports the hypothesis that high concentration of residual patches occurs closer to natural firebreaks (wetland and surface water). However, some erratic pattern is evident from the plots that show the relationship with distance from surface water (i.e., plots of partial dependency on surface water at R_{16}, R_{32}, and R_{64}; Figure 5.17). Second, the partial dependency plots pertained to bedrock and non-vegetated class indicated that high probability of residual occurrence is likely at closer proximity to non-vegetated areas and seemed to decrease after certain distance (e.g., from 3000 m at R_4), but it is difficult to generalize any kind of trends with increasing distance from the non-vegetated areas. Third, the effect of one the least informative predictor (ruggedness index) show a monotonic decrease with increasing ruggedness index values; indicating that high residual occurrence is associated with plain or nearly level surface, but the area is generally categorized as plain or nearly plain based on the RI classification scheme (§3.4.2; Table 3.8). Finally, the partial dependence of residual patches on natural firebreak features (i.e., wetlands and surface water) in the RED084 exhibit a similar trend with the 11 fire events studied in Chapter 3. Therefore, it can be inferred that the occurrence of residual patches within the disturbed landscapes are largely prevalent in closer proximity (less than 200 m) to natural firebreak features, and decreases with increasing distance from the natural firebreak features.

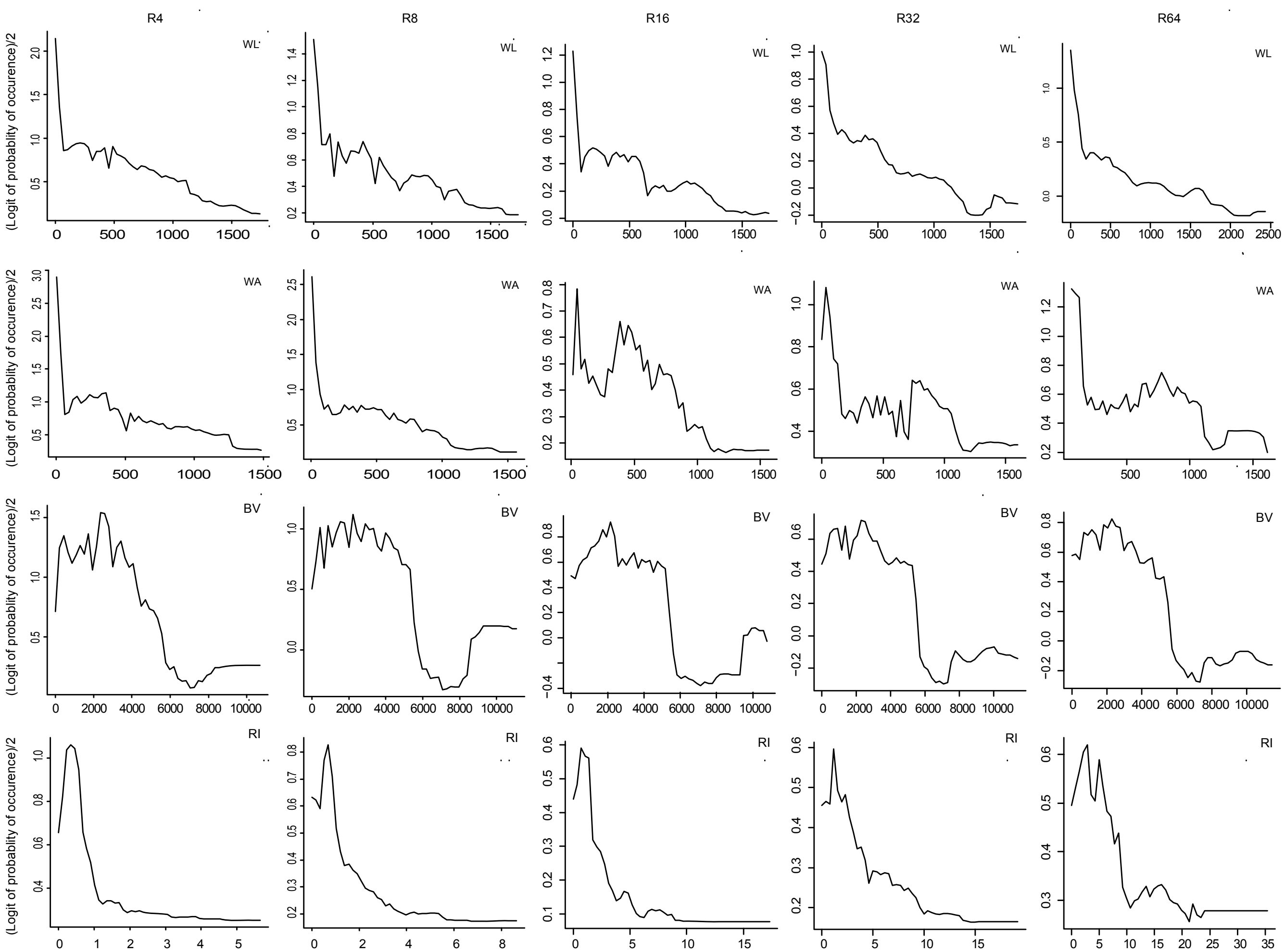


Figure 5.17. Partial dependence plots for selected predictor variables for random forest predictions of the presence of residual patches for RED084 at R_4, R_8, R_{16}, R_{32}, and R_{64}. Partial dependency is the dependence of the probability of presence on one predictor variable after averaging out the effects of the other predictor variable in the model. The x-axis of each plot indicates the explanatory variables (WL – distance to wetland, WA – distance to surface water, BV – distance to bedrock and non-vegetated areas, and RI – ruggedness index) while the y-axis is a half of the probability of occurrence (Equation 7).

5.3.5. Spatial prediction model of residual patch occurrence

5.3.5.1. Evaluating model's predictive performance

Prior to implementing the independent measure of the model's performance, the model was assessed using a fixed-probability threshold where three measures of accuracy: overall accuracy (*PCC*), sensitivity (S_n), and specificity (S_p) were incorporated (Table 5.5). The overall accuracy of discriminating residual patches from null residual patches was generally low across the gradient of scales, with *PCC* values less than 65%. This is contrary to the previous implementation of the model (Chapter 4) where the overall accuracy always exceeds 65% for all fire events across multiple scales. Unlike the previous study, the accuracy measure does not also show any trend with increasing grain size, suggesting that it is difficult to establish any scaling effect with increasing grain sizes. The results from the alternate measures of accuracy also showed that 1) the model had difficulties in discriminating residual patches from null-residual patches; yet a reasonable accuracy (S_n) was attained at R_4, and 2) the values were scale dependent, but the pattern of the scale effect is unpredictable. In the previous chapter, the model had done a reasonable job for discriminating residual from null-residual patches with high overall and sensitivity accuracies values as high as 96% and 99% respectively for one of the largest fire events (i.e., F01; R_4), indicating the model's sensitivity not only to grain sizes but also to geographic settings (i.e., ecoregions).

Table 5.5. Accuracy measures for classification of residual patches based on a fixed-probability threshold.

		R_4	R_8	R_{16}	R_{32}	R_{64}
RED084	PCC	36.71	63.08	52.12	58.33	62.05
	S_n	70.94	46.39	56.50	56.85	52.50
	S_p	31.85	65.95	51.37	58.60	63.88

A threshold-independent measure of model performance ROC was also used to assess the predictive power of the model. The ROC curve has the sensitivity plotted vertically and the reversed scale of the specificity plotted horizontally. The predictive performance of the model implemented for residual patch occurrence in the RED084 is graphically summarized in Figure 5.18, in which five ROC curves representing the predictive power of the model at R_4, R_8, R_{16}, R_{32}, and R_{64} are presented. The closer the curve follows the upper-left border of the ROC space, the higher the overall accuracy of the test while the closer the curve comes to the 45^o diagonal of the

ROC space, the lower the accuracy of the test. Generally, ROC curves lie between these two extremes, and provide a comprehensive and visually attractive ways to summarize the accuracy of predictions. The results of this study reflected such kind of trend where all the ROC curves lie between the extremes of perfect and random predictions. However, it is not easy to compare the prediction accuracy of the model directly from the ROC curves. While the ROC curves contain most of the information about the accuracy of a continuous variable, a quantitative summary measure of the ROC curve was also desirable. Therefore, the area under the ROC curve - AUC, which is a quantitative summary measure of a predictive model, was computed and presented along with the ROC curves in Figure 5.18 to provide a better quantitative estimate of the model's predictive performance in reference to the independent dataset.

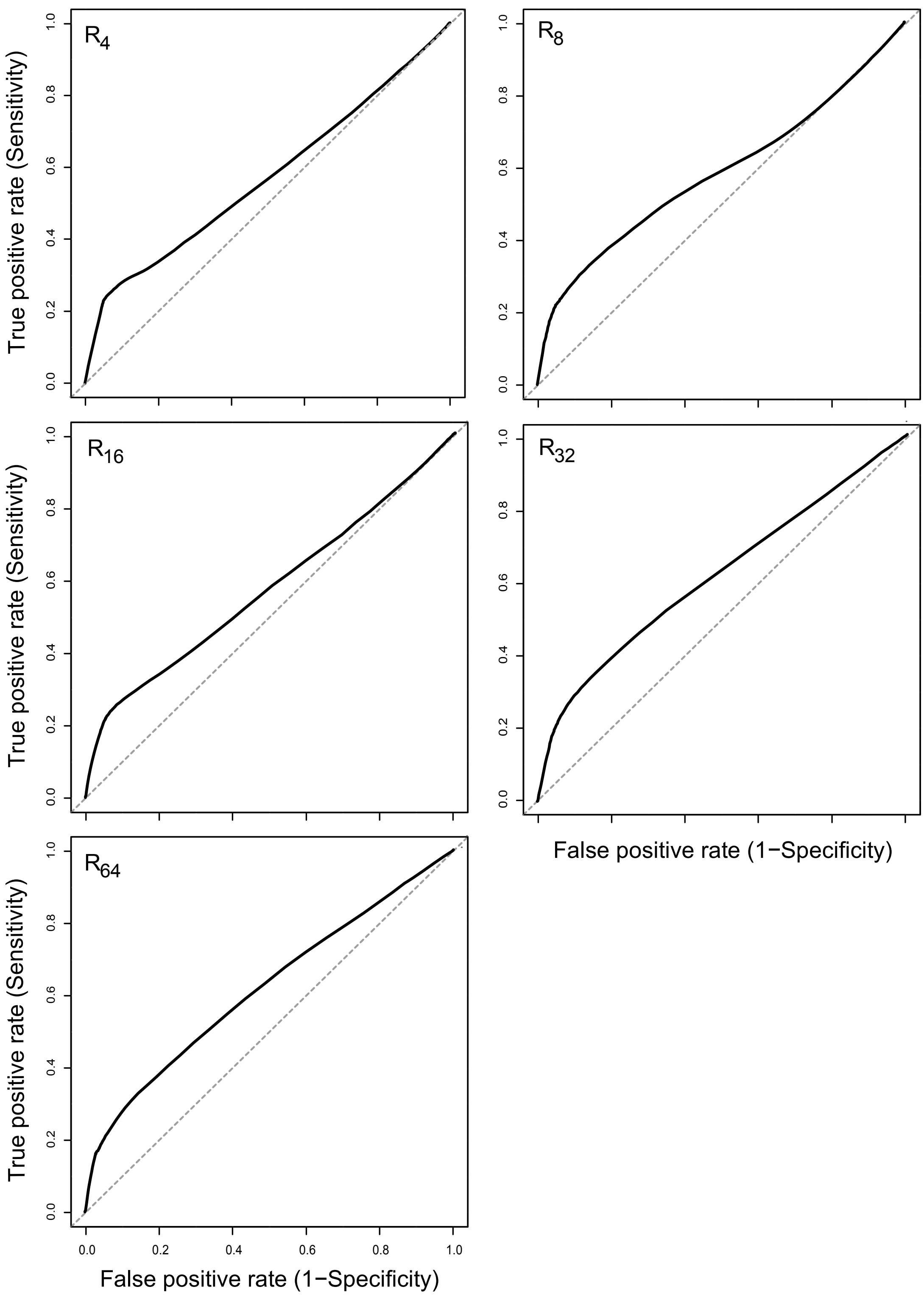


Figure 5.18. Graphical depiction of the model’s predictive performance.

The AUC is a better way of assessing a model's performance because it reflects the test performance at all possible cut-off levels. The AUC values are often between 0.5 and 1.0; the larger area, the better performance. Based on Swet's classification of AUC values (Chapter four; Table 4.2), the model has low discrimination ability with an index value that lies within the range of 0.5 and 0.7 across all the grain sizes (Table 5.6) (Swet 1988). Yet, the AUC values found to be significantly better than that expected from a random model ($\rho < 0.05$). The result suggested that the robustness of the method for residual patch occurrence across different ecoregion is uncertain. Moreover, a comparison of the predictive power of the model across gradient of scales indicated that the change in grain size does not affect the model substantially where low performance was attained across all the grain sizes.

Table 5.6. Prediction accuracy and statistical significance of RF model across five spatial resolutions.

Spatial resolution	**AUC**	***p-value***
R_4	0.571	0.0000
R_8	0.583	0.0000
R_{16}	0.572	0.0000
R_{32}	0.615	0.0000
R_{64}	0.617	0.0000

5.3.5.2. Spatial prediction of residual patches

The accuracy and statistical validity of the model was assessed, and the graphical representation of the model output was also derived to present spatially explicit predicted probability maps of residual patch occurrence. In this section, the results of spatial prediction at R_4 and R_{32} are presented in Figure 5.19 and Figure 5.20 respectively. The result from binomial response models is a probability value scaled from 0 to 1 for each grid cell, with predictions closer to 1 indicating greater chance of residual patch occurrence.

Based on the AUC values, the predictive performance of the model was poor; yet visual interpretation of the predicted maps showed that the model was able to identify some potential areas where residual patches are likely to occur. In the variable importance assessment, wetland was considered to be the most informative predictor for explaining residual patch occurrence. This is reflected in the predicted probability maps where the light shading (i.e. high probability of residual occurrence), green areas (existing residual patches), and cross-hatched areas (wetland) of the map tended to coincide in some parts of the area. This indicates the capability of the model to identify potential unburnable areas. In summary, visual inspection of the probabilities underlying the prediction indicated that 1) some potential areas were identified despite the poor performance of the model, and 2) residual patches cannot be retained within surface water, this has been reflected in the predicted maps where surface water are associated with low (zero) probability values. Also, given the predictors considered in the study, the low prediction accuracy of the model suggests that the model was not robust enough to predict residual patch occurrence across different ecoregions.

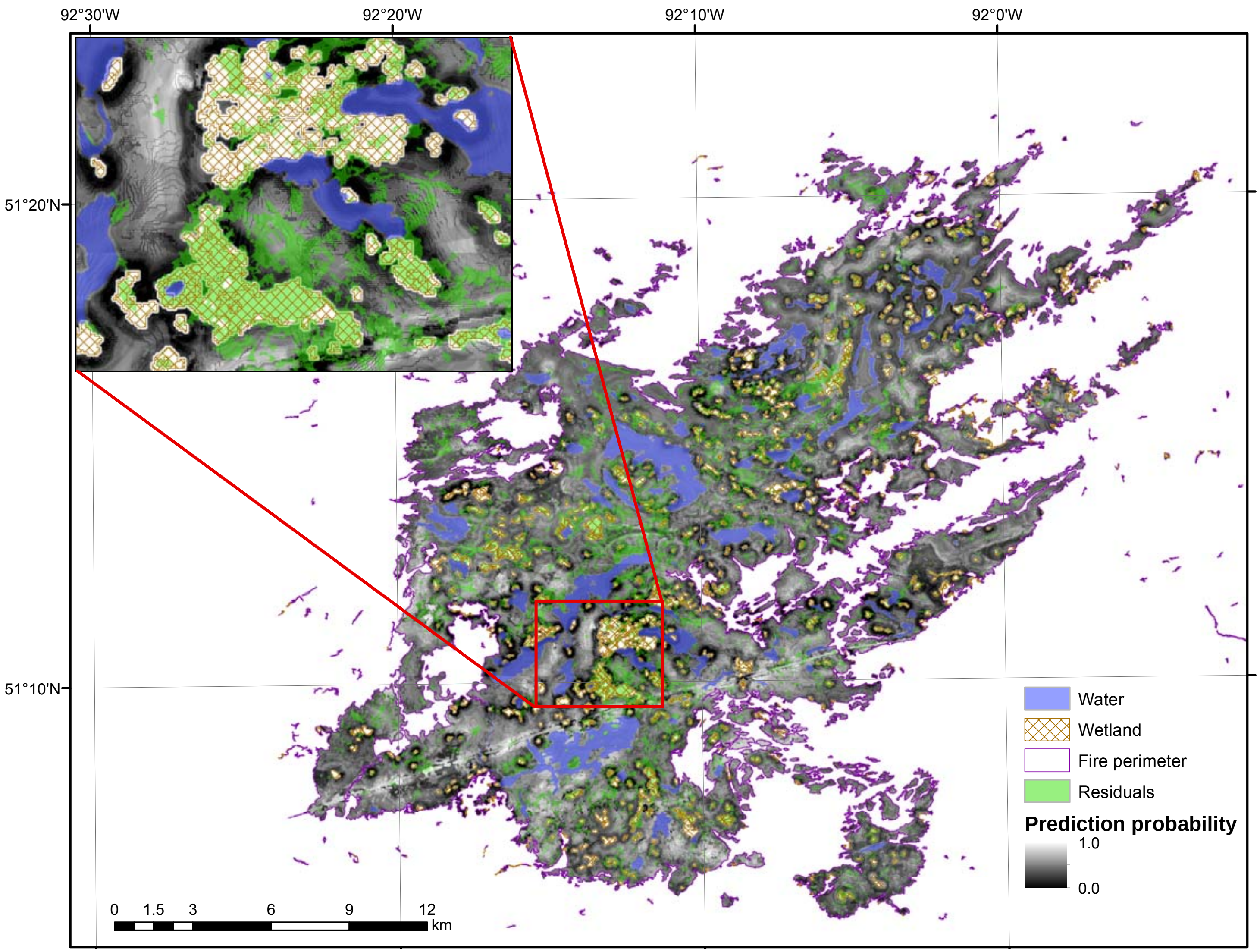


Figure 5.19. Predicted probability map of residual patch occurrence in the RED084 at R_4.

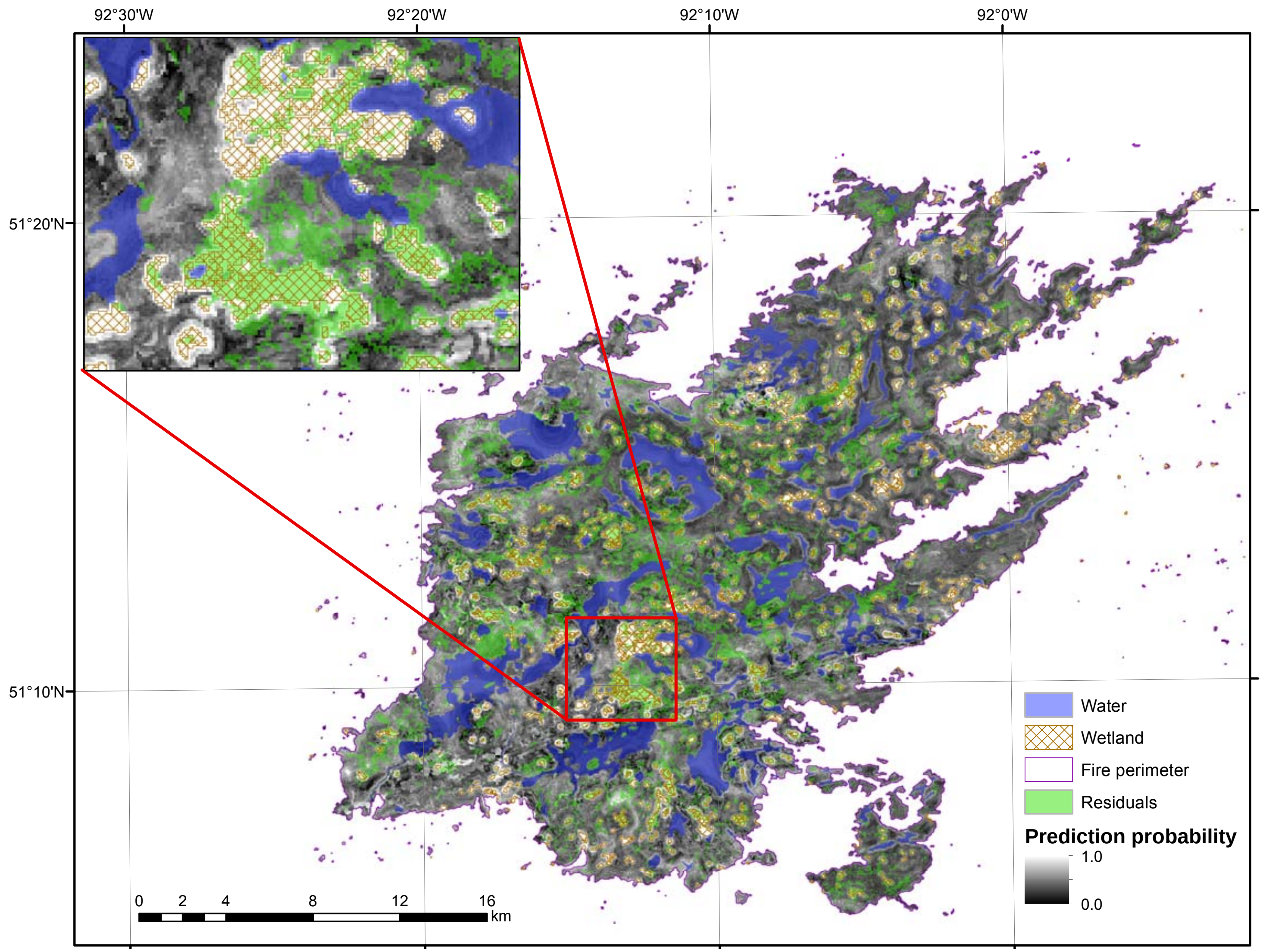


Figure 5.20. Predicted probability map of residual patch occurrence in the RED04 t at R_{32}.

5.4. Discussion

5.4.1. Spatial characteristics of residual patches

With respect to the spatial patterns of residual patches, the results showed that a substantial (around 10%) area of the fire footprint has escaped burning. This is similar to the average residual patch area computed from the 11 fire events. Yet, most of the residual patches were small, with 62% of the patches smaller than 1 ha. Being one of the largest fire events in northwestern Ontario, the RED084 consumed a considerable area; but a substantial area escaped burning, indicating the presence of considerable natural firebreak features. Many of the residual patches were also dispersed throughout the footprint with residual patches that tended to show a linear and elongated shape structure (MSI of 3.27) at 4 m spatial resolution. This can also be supported with the fractal dimension (FRAC = 1.23) computed for the same landscape. The findings are relatively consistent with the 11 fire events where MSI values ranges from 3.55 to 5.91 and FRAC was between 1.27 and 1.37. As a diagnostic, the finding is different from a study undertaken by Dragotescu and Kneeshaw (2008) where circular shaped residual stands (MSI = 1.5 – 1.6) were generated. Andison (2004) also found MSI values ranging from 1.3 to 2.9 depending on fire size. Such discrepancies in landscape metrics could be attributed not only to the landscape, but also to the spatial resolution of the dataset, the way in which the minimum patch area is defined, and the resampling (aggregation) method applied.

In landscape studies, the presentation of surface properties and ecological processes (e.g., fire disturbances) is inherently linked to the scale of analyses (Moody and Woodcock 1995). The scale multiplicity in spatial heterogeneity indicates the need to incorporate scaling effects in landscape research. The patterns of residual patches were analysed at multiple scales and the patterns of residual patches within the RED084 were sensitive to changes in spatial resolution. This confirms the findings of earlier studies on landscape metrics, with the maximum structural detail of landscape objects is obtained when using the highest possible spatial resolution of the input data. Using the Wu et al. (2000) classification of metrics, the patterns of landscape metrics related to shape (LSI, MSI, and FRAC) can be categorized into a group where metrics exhibit a (consistent) monotonic change and predictable patterns. The same trend was observed in the previous study (Chapter 2) where the three shape related metrics exhibited a predictable pattern; hence a robust scaling relation in the form of a power law over a range of scales can be established for these metrics across different fire events observed in different ecoregions.

However, such a predictable pattern was not reflected for most of the landscape metrics considered in this study. Looking at the best fit model and observed LPM values (Figure 5.9), one would argue that a scaling rule can be developed across multiple scales, but this is true only for the specific fire event (RED084). The metrics did not show consistent responses with the other fire events; rather the same metrics (CA and LPI) showed an erratic behaviour for the largest fire events studied in Chapter 2 (i.e., F01, F06, F08, and F10). This hinders one to develop a robust scaling rule that predicts patterns across multiple grain sizes for fire events from different geographic settings. Moreover, other landscape metrics (PD, NP, MPS, PSSD, PSCV, and MNN) that are expected to decrease or increase monotonically with increasing grain sizes exhibited an erratic response. The non-monotonic or erratic response curves demonstrate the unpredictable patterns of the LPM; hence it is not easy to develop any scaling law to predict pattern at different scales.

Large fires often leave some unburned islands due to the abundance of natural fire break features and diversity in fuel characteristics or land cover type within the extent of the fire perimeter. However, one needs to understand the degree to which, and the reasons why, certain land cover types tend to avoid burning. Large fires (e.g., RED084) are undoubtedly the results of extreme fire behaviour and would consume a considerable area of almost all cover types (Burton et al. 2008). The results of this study indicated that residual patches are mainly dominated by dense conifer, which is the dominant cover type in the area. Sparse conifer and treed wetland also constitute a considerable area of residual patches. This is contrary to the previous study where both sparse conifer and treed wetlands comprised majority of the residual patches for most of the fire events studied in Chapter 2. In a similar study by (Burton et al. 2008), deciduous forests were less likely to burn in boreal mixedwood region; hence dominant the post-fire residual patches. Kafka et al. (2001) also found that unburned areas were positively associated with deciduous or mixedwood and negatively associated with conifer species.

Furthermore, the distribution of residual patches could also be explained in relation to their proximity to surface water and edge of the fire perimeter. Perera et al. (2009) indicated that proximity to water has no influence on residual patches; yet a study by Madoui et al. (2010) found a significant association with the proximity to surface water. Despite the distribution of residual patches throughout the fire footprint, there was a decreasing trend in the distribution of residual patches with increasing distance from surface water (Figure 5.12). A similar decreasing trend was also apparent with proximity to the edge of the fire perimeter. This is plausible because there is a tendency for forest features to escape burning as the intensity and fire spread rate start to decrease, and eventually leave unburned areas as it reaches the fire edge. For example, Smith and Hendry (1998) discovered that almost all residual islands were located adjacent to a fire perimeter. Owing to wind direction, however, there is a potential for residual patches to be concentrated on certain edges of the fire footprint.

5.4.2. Spatial variable importance assessment

The occurrence of residual patches can be explained in relation to several factors, but the study considered seven explanatory variables that can be grouped into three categories: topography, natural firebreak features, and land cover type. Topographic variables of slope and ruggedness index play an important role for determining residual patch occurrence (van Wagtendonk 2004; Cuesta et al. 2009; Madoui et al. 2010); this is particularly true in rugged topography. In spite of this, my results revealed that topographic variables of slope and ruggedness index were among the least informative predictors in the RED084 fire event. In a similar study, on the 11 fire events, these variables were among the least important predicators. One potential reason is pertained to the ruggedness of the area; based on the RI computed at the given spatial resolution (Riley et al. 1999), the RED084 is located in a relatively flat plains or nearly level (i.e., land with relatively low relief). However, it has to be noted that the spatial resolution of the source data was coarse, and might hinder one to estimate the local variability in topography effectively.

The second category is related to natural firebreak features, including wetland, surface water, and non-vegetated areas. Various studies have addressed the impact of natural firebreak features for determining residual patch existences (Turner et al. 1997; Perera et al. 2007; Cuesta et al. 2009). Some studies have specifically addressed that the occurrence and distribution of residual patches is attributed to the abundance of surface water (Madoui et al. 2010; Dragotescu and Kneeshaw 2012). Despite the change in the relative ranking of the predictors, the study suggested that the three natural firebreak features (surface water, wetlands, and non-vegetated areas) were among the most important variables that explain the residual patches. Compared with the 11 fire events, the relative ranking of surface water and non-vegetated variable is more prevalent in the RED084 fire event. This is attributed to the relative abundance and distribution of surface water and non-vegetated feature classes in the area. Yet, the relative ranking of the importance scores indicated that natural firebreak features (i.e., wetland) is the most important predictor variable across all the grain sizes considered. This was in line with the finding on the 11 fire events where wetland was found to be the most important predictor.

The abundance and distribution of residual patches is also related to land cover types, and residual patches are often associated with the dominant cover types in the landscape. However, certain cover types (e.g., deciduous forest) tend to dominate the existing residual patches in burned landscape despite their low abundance (Kafka et al. 2001; Burton et al. 2008). In a study undertaken by (Madoui et al. 2010) for example, sparse conifer was dominant cover type among the residual patches. Compared with dense conifer, sparse conifer has a low tree density and may not be prone to high-intensity fires because of a lack of fuel. A study on the 11 fire events indicated that the abundance of residual patches is attributed to land cover types with high abundance in the landscape and cover types that are less prone to high fire severity (e.g., treed wetlands). In this specific scenario, majority of the residual patches were dominated by dense and sparse conifer. Also, the boreal region is characterized by low plant species diversity (Engelmark 1999), and land cover variable was less informative to explain the residuals compared to other predictors.

The theory behind random forests provides a very natural way to rate the relative importance of variables. However, RF implementation of variable importance assessment is not designed for traditional statistical inference, and the importance scores have been used to subjectively identify important predictors for interpretation (Cutler et al. 2007). Therefore, digging deeper than just which variable is important would allow one to examine the effect of different values of a given variable on the class prediction. Based on Strobl's rule of thumb (Strobl et al. 2009b), most of the variables deemed to interactively affect the occurrence of residual patches and the relative ranking of the predictors changed slightly with increasing grain sizes. The results of this study showed different patterns, as mentioned above, but the occurrence of residual patches was largely prevalent in closer proximity to natural firebreak features.

5.4.3. Model predictive performance and spatial prediction

A computer-based predictive model was implemented to generate probability maps for residual patch existence within a burned landscape. Such kind of model has been used to predict and extrapolate species distribution in various geographical settings or across large regions (Pearce et al. 2001). This provides the required spatial information for conservation planning and resource management (Ferrier et al. 2002; Mellor et al. 2013). At geographical extent beyond the area of experimentation, predictive models provide a means to design and test hypotheses pertained to the factors affecting the characterization and distribution of spatial elements (Manel et al. 2001). Given the importance of statistical models, continuous and progressive evaluation of the models is necessary; effective and correct model assessment has real significance to landscape ecological studies (Fielding and Bell 1997; Pearce and Ferrier 2000; Pearce et al. 2001; Manel et al. 2001; Austin 2007). This provides information on the uncertainties associated with the predictions, and hence the desired precaution can be made when implementing the model for specific applications (Ferrier et al. 2002). Ideally such models should be tested with independent data. However, a review on published literatures indicated that many users of predictive models make no evaluation at all (Manel et al. 2001). This limits the applicability and suitability of the model for the desired applications. Knowledge of the predictive performance of methods and their domain of application becomes an important issue in developing a project aimed at mapping species distribution (Brotons et al. 2004).

Magness et al. (2009) argued that for effective monitoring programs the confidence in reliability of a model should be placed on the predictive ability of the map; not on the relationship of each of the predictor to the response variable. Therefore, in this study, the predictive accuracy of the model was assessed by comparing the actual observation with predicted probabilities of occurrence generated by the model. The model was evaluated in relation to independent data (RED084 fire event) other than those used to develop the model. This involves computing the proportion of locations at which presence (residual patches) or absence (null-residual patches) is correctly predicted, although the study emphasized on the prediction success of residual patches. The predictions of residual occurrence derived from RF model can be used in two ways. First, predictions can be used as a relative suitability of a location for residual occurrence, where higher values indicate locations that are more likely to be occupied by residual patches. Second, predicted probabilities can be converted into a binary set (presence and absence) using a cut-off value to the predicted probability range. Some of the applications include maps of relative suitability of locations for residual occurrence or the ranking of priority areas for retaining residual patches during harvesting for emulating forest disturbances. The discrimination index based on AUC values of the ROC curves provides a summary measure of these capabilities.

Given the environmental variables considered, the results showed that RF model predicted the occurrence and distribution of residual patches with low accuracy (weak model outcome with AUC < 0.7) than the previous implementation of the model. The most obvious question is why the model did not perform well when it was applied in a different setting. Burton et al. (2008) noted that depending on fuel availability and source of ignition, every fire represents a unique fire combination of fire skips that affect forest species and habitat features in the canopy, understory, and in the forest floor. The variation would result in a mosaic of fire size distribution and post-fire landscape structure. Similarly, it was argued that there are interregional or inter-landscape differences within the boreal forest as a function of climate and topographic effects (Burton et al. 2008). Thus, the differences within the landscape caused by different environmental components and abundance of natural firebreak features could contribute to the low prediction accuracy.

5.5. Summary and conclusions

In this chapter, the utility of RF for building distribution maps for a fire-disturbed landscape using a set of predictor variables used in the previous chapter was explored. The purpose was to evaluate the predictive power of the model implemented in the previous chapter using independent dataset from the extensive RED084 fire event. Based on the findings, the following conclusions are drawn. First, given the seven predictor variables incorporated in the study, the predictive model was not robust enough to predict the occurrence and distribution of residual patches in the independent fire event (i.e., RED084). The RF algorithm produced weak predictive model with prediction accuracy (AUC < 0.7). This reflects the idea that modelling approaches developed and applied even in relatively data-rich regions may not necessarily work effectively in a different site (Ferrier et al. 2002). Second, despite the low predictive performance of the model, the model was better at predicting residual patch occurrence than random guessing ($p < 0.001$). The model was able to identify potential areas where residual patches are likely to occur. Specifically, high prediction probability was associated with the existence and abundance of natural firebreaks, particularly wetlands. This reflects the potentiality of the predictor variables and the model itself to predict residual patches to a certain extent.

The other goal of this chapter was to explain the patterns of residual patches, impact of land cover composition, and proximal analysis of residual patches using various spatial metrics. Moreover, knowledge about the factors influencing the distribution of spatial elements (e.g., residual patches) is among the most important aspects in landscape ecology (Manel et al. 2001). Therefore, RF was implemented to estimate the variables that govern the existence of residual patches within the RED084 burned landscape and investigate the marginal effects of each of the predictors on probability of residual patch occurrence. Based on these, the following conclusions are also drawn. First, the landscape pattern indices used to quantify residual patterns were scale dependent. The sensitivity among the metrics varied greatly, and the sensitivity of the metrics can be explained in relation to three major categories: 1) metrics with monotonic change and predictable patterns, 2) metrics that exhibit monotonic change, but unpredictable pattern, and 3) metrics that show erratic behaviour. Comparing RED084 with the 11 fire events studied in Chapter 2, the sensitivity also varied spatially across the study areas. The variation in the sensitivity among the indices and over space may indicate that the responses of indices should not be evaluated the same way for all indices across different sites. Second, the residual patches were mainly occupied by the dominant cover types (i.e., dense conifer) in the landscape, but cover types that are less abundance in the landscape (e.g., treed wetland and sparse conifer) occupies a considerable amount of residual patches. Therefore, the occurrence of residual patches is associated with dominant land cover types or features types that are less prone to burning (e.g., treed wetland). Third, similar to the previous study on the 11 fire events; natural firebreak features such as surface water and wetlands were determinant for retaining residual patches. Additionally, residuals patches are often retained in closer proximity to natural firebreak features (water and wetlands), specifically wetlands. Finally, the overall findings confirm that ensemble classifiers can be used to learn complex and non-linear relationship between spatial elements and the environmental variables, which is in line with previous studies undertaken based on RF models (Edwards et al. 2007; Evans and Cushman 2009; Mellor 2013).

6. Conclusions and future directions

6.1. General summary and conclusions

The Canadian boreal forest is characterized by a short growing season, low mean temperatures, long summer daylight hours, low biological productivity, and variable annual rainfall. This set of factors create an environment susceptible to different types of disturbance, of which wildfire is the most common type of natural disturbances that occurs in boreal ecosystems and produces a spectrum of effects on wildlife species (e.g., loss or decrease in wildlife population), soil, vegetation, and wetland components of a landscape. The ignition and occurrence of wildfires in the boreal forests is generally triggered either by human activities (e.g., recreation and forestry activities) or natural factors, but fire naturally caused by lighting is the most common and frequent type of disturbances in the region. Owing to the various forms of environmental factors (such as weather conditions, vegetation, topography, natural firebreak features, and fuel characteristics), wildfire acts as a major shaping force for the spatial heterogeneity in forest ecosystems, including the occurrence of residual patches following a fire.

The patterns and distributions of residual patches vary spatially and temporally as a function of various geo-environmental factors, including topography, local weather variables, fuel distribution, land cover, and abundance of natural firebreak features. The composition and spatial variability of residual patches should then be mapped and measured to provide insight into the cover patterns following a fire, understand forested landscape, provide baseline data for wildlife studies, examine fire behaviour, and implement disturbance-based forest management practices. To this end, an integrated and rigorous measurement framework has been designed and implemented for: 1) characterizing the spatial composition, structure, and variability of residual patches, 2) assessing the factors responsible for the composition and variability of residuals, and 3) developing a spatially explicit predictive model to generate probability maps and assessing its predictive performance in a given landscape. This study showed that there is a consistent and robust measurement framework that enable us to study the patterns, characteristics and distribution of residual patches in a given fire-disturbed landscapes.

In this study, the process of measuring and quantifying the spatial composition and configuration of residual patches was assessed using geospatial tools, including GIS, remote sensing, and spatial statistics, coupled with different landscape metrics. This study found that there was a variation in terms of residual patch composition, configuration, and fragmentation across the 12 fire events studied throughout this dissertation, reflecting the uneven burn severity or fire intensity across the fire events. The variation was explained in relation to disturbance size where the proportion of the fire footprint occupied by residual patches was relatively substantial for the large sized fire events while it was comparatively small for the small sized fire events. The landscape metrics (such as NP, PD, and LPI) used to quantify the spatial variability and heterogeneity of residual patches also revealed that the residual patches were irregular and complex, and are highly fragmented with more than 75% of the residual patches less than 1 ha. The spatial variability in composition and fragmentation across the fire events pertains to the abundance of natural firebreak features and variation in fuel distribution. The occurrence of residual patches and their variabilities can also be associated not only with the dominant land cover types, but also to cover types that are less abundant in a landscape and less susceptible to burning (e.g., treed wetlands and sparse conifer). Besides, the study showed that the relative abundance and distribution of residual patches within the fire perimeter in relation to surface water and edge of fire perimeter would mostly be uneven.

Knowledge about the factors that influence the patterns and variabilities of residual patches in a landscape is among the most important aspects of ecological theories. In response, this study implemented a replicable data mining technique (i.e., RF) to unravel the complex relationship between residual patches and environmental variables, using measurements that could be made locally, not relying on interpolated weather or fuel moisture content data. The algorithm was implemented to measure the relative importance of obtainable environmental predictor variables, and their marginal effect for explaining the mechanism and casual factors of residual patch structure. In data mining applications, it is rare to see that input environmental variables are equally relevant for explaining the patterns of spatial objects; there is a tendency for certain predictors to become more informative than others. This was reflected in this study where the variable importance measure in RF was used to subjectively identify important variables for interpretation. Despite the variation in the composition and configuration of residual patches across the 12 fire events, I discovered that certain predictor variables, specifically variables pertained to natural firebreak features such as wetlands, were among the most important predictors while topographic and land cover related variables were among the least informative predictors throughout the study. The relative importance measures of the predictors provided a ground for assessing the marginal effects of the most important variables. Accordingly, the results of the partial dependency plots revealed that the relationship between the predictor variables and residual patch occurrence was non-linear and the occurrence of residual patches was prevalent within 100 m from the wetlands.

In order to further examine the impact of the predictors on residual patch occurrence, the overall effects of the available environmental variables, rather than individual variables, were assessed using a predictive modelling approach. While a wide range of methods have been developed to build predictive models based on presence/absence, but models based on ensemble methods such as RF have been extensively tested and shown to be robust in a number of independent situations (Brotons et al. 2004), as RF is a package of fully nonparametric statistical methods designed for data analysis. In response, I presented a predictive model based on RF to develop learning rules that determine areas where residual patches are likely to occur, generate timely and spatially explicit probability maps, and test the predictive performance of the model on larger and independent fire event.

Given all the desired parameters for prediction, the RF model was determined to be a robust modelling approach for predicting residual patch distribution from presence/absence data. I also discovered that the RF model was an effective tool to distinguish between presence and absence data, and identify areas where residual patches are likely to occur within a burned landscape with different prediction accuracy. The RF algorithm, coupled with the geoinformatics tools, also enabled me to generate inexpensive and spatially explicit maps of post-fire forest characteristics over a gradient of scales. Parallel to the previous findings, where natural firebreak features were found to be more informative, I found that high prediction probability of residual patches was associated within or in closer proximity to wetlands. In addition to this, prediction based on RF makes no assumption about the predictor or response variables, and can handle situations in which the number of predictor variables exceeds the number of observations. Given the unique characteristic features and experiences from this research, I concluded that RF model a robust (ensemble) and replicable approach for learning complex and non-linear relationship, predicting residual patch distribution from presence-absence data which is in agreement with previous studies undertaken based on RF models (e.g., Edwards et al. 2007; Evans and Cushman 2009; Dahinden 2011).

The patterns and characteristics of residual patches vary as a function of measurement, spatial resolutions, and fire event sizes. Thus, I examined the effects of changing grain size not only on the spatial composition and configuration residual patches, but also on the relative importance of the predictors and the predictive performance of the RF model. The study reflected that scale multiplicity is inherent in spatial heterogeneity and a multi-scale analysis is imperative for detecting and understanding the multi-scale structure of spatial heterogeneity. This study showed that all the landscape metrics used to characterize and quantify the spatial composition and variability of residual patches were sensitive to spatial resolutions, but the overall response curves from the 12 fire events can be summarized into three general categories: metrics that decreased monotonically and showed a predictable pattern with increasing grain sizes, metrics that increased or decreased monotonically but with no simple scaling rule, and those that generate erratic response. The predictable response curves generated in this study were pertained to shape related metrics (MSI, LSI, and FRAC) while the remaining metrics fall within the last two categories. Based on this, I concluded that a robust scaling rule can be developed for determining certain metrics (shape related metrics) while most of the metrics cannot easily be dictated using a scaling rule. Therefore, characterizing the patterns and variability of residual patches should be undertaken at specific grain sizes.

Similarly, the effects of changing grain sizes on the relative importance of the predictors and spatial prediction were assessed. I explored that the relative importance values of the predictors varied across the five spatial resolutions, and there has also been a change in the relative ranking of the predictors, which affects the variable selection, and on the marginal effects of predictor variables across the five spatial resolutions; suggesting the need to implement RF model at specific grain size. I also demonstrated that the predictive performance of the model was sensitive to changing grain sizes; with finer grain sizes (e.g., R_4 and R_8) attained high prediction accuracy than the coarser grain sizes (R_{32} and R_{64}). Additionally, this study found that the predictive performance of the model was sensitive not only to grain sizes, but also to sample sizes; for a model to predict residual patches with the desired prediction accuracy, sufficient sample is required as an insufficient size sample (i.e., residual patch data records) affects the performance.

The overall methods developed and implemented in this dissertation revealed that there is a repeatable, robust measurement framework for characterizing residual patches and understanding their variability across different landscapes and spatial resolutions. I incorporated various approaches to characterize residual patches and predict their likely occurrence. The methods designed and implemented were replicable and broadly applicable; they are not limited to the fire events and spatial resolutions considered in this dissertation. The replicable methodologies integrated in this dissertation can consistently and broadly be applied to other fire events or landscapes (i.e., fire events in different ecoregions or ecozones) where 1) post-fire forest characteristics can be mapped using the geospatial tools in a similar way; 2) the extent of fire footprint and residual patches can be mapped and extracted using the procedures developed by (Remmel and Perera 2009) and implemented in this study; 3) spatially and environmentally relevant variables can be derived from remote sensing and DEM data for developing a predictive modelling for prediction; 4) the absence-data can algorithmically be extracted using the techniques developed and implemented; 5) learning rules can be established to dictate where residual patches are likely to occur; and 6) the predictive performance of any model developed can be evaluated in an identical way implemented in this study.

For any fire-disturbed landscape, the methods integrated in this dissertation can be applied to measure and understand the characteristics of residual patches, and improve our understanding of the mechanism and causal factors of residual structure and probabilities of residual patch occurrence. This would subsequently help to review and evaluate the existing management guidelines for natural disturbance pattern emulation (OMNR 2001), and for boreal landscapes and conservation biodiversity (OMNR 2014). The management guidelines have been revised and modified after evaluating and reviewing their effectives, efficiency, and effects (OMNR 2014). The guidelines are reviewed based on the applicability of new scientific approach and advancements in analytical and operational technology (OMNR 2014). Although it is beyond the scope of this study, the methods implemented in this dissertation may be beneficial to reviewing and evaluating the existing management guidelines.

6.2. Future directions

This dissertation formulated a comprehensive framework for examining the patterns, characteristics, and geographic distributions of post-fire residual vegetation patches within a fire-disturbed landscape. The study showed that there is a reliable and rigorous method in the study of post-fire residual patches and has laid a foundation for exploring 1) disturbance as necessary agents of change, not as element to be excluded entirely, 2) the potential RF for understanding the mechanism and causal factors of residual structure, and 3) the use of RF based predictive model for investigating the potential areas where residual patches are likely to occur. The robust method implemented in this study is broadly applicable and opens up avenue for future work in implementing the approaches using different scenarios to further improve our understanding of residual patch characteristics.

The results of this study provided insights on the parameters that explain residual patches, marginal effects of the predictors, and spatially explicit predictive maps that show the potential distribution of residual patches within a fire-disturbed landscape. The methods designed and implemented were based on the obtainable environmental variables that are related to topography, land cover, and natural firebreak features. It would be interesting to conduct future research work to incorporate additional and accessible predictors to study the patterns of residual patches; this includes predictors such as forest age composition (the composition of pre-burn forest age class), fuel type composition (composition of pre-burn fuel type), and fine resolution weather data. It may also be interesting to test whether a finer spatial resolution DEM, with finer levels of detail has an effect on the relative importance of the predictors and the predictive qualities of the RF model. Therefore, these are areas of research and development where the rigorous measurement framework implemented in this study can be replicated to further understand the patterns of wildfire residuals. Besides, the multi-scale approach implemented in this study has yielded a useful outcome for characterizing the patterns of residual patches and develop spatially explicit predictive model to dictate residual patch occurrence over a gradient of scales. Another potential area of future research, pertained to the multi-scale analysis, would be to consider different spatial aggregation methods (i.e., independent as opposed to 'iterative' spatial aggregation and majority rule as opposed to random rule-based spatial aggregation) in the process of measuring and characterizing residual patches over multiple grain sizes.

7. References

Agee, J.K. 1998. The landscape ecology of Western forest fire regimes. Northwest Science **72**: 24-34.

Ahlgren, I.F., and C.E. Ahlgren. 1960. Ecological effects of forest fires. Botanical Review **26**: 483-533.

Ainsworth, J. and T.A. Doss. 1995. Natural history of fire and flood cycles. California Coastal Commission. http://www.coastal.ca.gov/fire/ucsbfire.html

Anderson, R.P., D. Lew, A.T. Peterson. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models **162:** 211-235.

Andison, D.W. 2004. Island remnants on Foothills and Mountain Landscapes of Alberta: Part II on residuals. Alberta Foothills Disturbance Ecology Research Series Report No. 6.

Andison, D.W. 2012. The influence of wildfire boundary delineation on our understanding of burning patterns in the Alberta foothills. Canadian Journal of Forest Research **42**: 1253 – 1263.

Andison, D.W. 2013. Wildfire patterns in western Boreal Canada. Healthy Landscapes Research Series Report No. 8.

Allen, T. F. H. and Hoekstra, T. W. 1992. Toward a unified ecology./ Columbia Univ. Press.

Archer, K.J. and R.V. Kimes .2008. Empirical characterization of random forest variable importance measures. Computational Statistics and Data Analysis **52** (4): 2249-2260.

Austin, M.P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling **157**: 101-118.

Austin, M. 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecological Modelling **200**: 1-19.

Baldwin, D.J., J.R. Desloges, and L.E. Band. 2000. Physical geography of Ontario et al. 2000: *In*: Perera, A.H., D.E. Euler and I.D. Thompson. (Editors). Ecology of a managed terrestrial landscape: patters and processes of forest landscapes in Ontario, University of British Columbia Press, Vancouver, BC. 141-162 p.

Baysal, I., M. Quellette, and J. Antoszek. 2011. Red Lake 084 of 2011: a reconnaissance survey of a large Boreal wildfire. Ontario Ministry of Natural Resources, Ontario Forest Research Institute, Sault Ste. Marie, Ontario. Forest Research Report No. 177.

Beauvais, G.P., D.A. Keinath, P. Hernandez, L. Master, and R. Thurston. 2006. Element distribution Modelling: a primer (Version 2), Nature Serve, Arlington Virginia.

Benson, B.J. and M.D. MacKenizie. 1995. Effects sensor spatial resolution on landscape structure parameters. Landscape Ecology **10**(2): 113-120.

Bergeron, Y., O. Engerlmark, B. Harvey, H. Morin, and L. Sirois. 1998. Key issues in disturbance dynamic in boreal forests. Journal of Vegetation Science **9**(4): 464-468.

Bergeron, Y., P. Drapeau, S. Gautheir, and N. Lecomte. 2007. Using knowledge of natural disturbances to support sustainable forest management in the northern Clay Belt. The Forestry Chronicle **83**(3): 326-337.

Blaschke, T., M. Conradr, S. and S. Lang. 2002. Multi-scale analysis for ecological monitoring of heterogeneous, small structured landscapes. Proceedings of SPIE, Toulouse.

Bonan, G.B., and H.H. Shugart. 1989. Environmental factors and ecological processes in boreal forests. Annu. Rev. Ecol. System **20**: 1-28.

Breiman, L. 2001. Random Forests. Journal of Machine Learning **45**(1): 5-31.

Brenning, A. 2005. Spatial prediction models of landslide hazards: review, comparison, and evaluation. Natural hazards Earth System Science **5**: 853-862.

Brotons, L., W. Thuiller, M. Araujo, and A. Hirzel. 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography **27**: 437-448.

Burton, P.J., M. Parisien, J. Hicke, R. Hall, and J. Freeburn. 2008. Large fires as agents of ecological diversity in the North American boreal forest. International Journal of Wildfire **17**: 754-767

Butler, G. 2007. City of Golden fire department draft community wildfire protection plan. WALSH project number: 7404-030.

Cain, D.H., K. Riitters, and K. Orvis. 1997. A multi-scale analysis of landscape statistics. Landscape Ecology **12**: 199-212.

[CBI] Canadian Boreal Initiative. 2005. State of Canada's forests. URL: http://www.borealcanada.ca/boreal-did-you-know-e.php

Chafter, C.J., M. Noonan, E. Macnaught. 2004. The post-fire measurement of fire severity and intensity in the Christmas 2001 Sydney wildfires. International Journal of Wildland Fire **13**: 227-240.

Chehata N., L. Guo, and C. Mallet. 2009. Airborne LiDAR feature selection for urban classification using Random Forests. ISPRS Laser scanning workshop.

Chu, T., and X. Guo. 2014. Remote sensing techniques in monitoring post-fire effects and patterns of forest recovery in boreal forest regions: a review. Remote Sensing *ISSN 2072-4292* **6**:470-520.

Cifaldi, R.L., J.D. Allan, J.D. Duh, and D.G. Brown. 2004. Spatial pattern in land cover of exurbanzing watersheds in southeastern Michigan. Landscape and Urban Planning **66**: 107-123.

Clark, J., and T. Bobbe. 2007. Using remote sensing to map and monitor fire damage in forest ecosystems. In: Wulder, M.A., and S.E. Franklin. (Editors). Understanding forest disturbance and spatial patterns: remote sensing and GIS approaches. Broken Sound Parkway, NW: Taylor and Francis Group. 113-131 p.

Collins, B.M., M. Kelly, J.W. van Wagtendonk, S.L. Stephens. 2007. Spatial patterns of large natural fires in Sierra Nevada wilderness areas. Landscape Ecology **22**: 545-557.

Coops, N.C., M.A. Wulder, J.C. White. 2007. Describing forest disturbance and spatial pattern: data selection issues and methodological implications. *In*: Wulder. M.A. and S.E. Franklin. (Editors). Understanding forest disturbance and spatial pattern. Broken Sound Parkway, NW:Taylor and Francis Group. 31-62 p.

Congalton, R.G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment **37**: 35-46.

Corry, R.C. and R. Lafortezza. 2007. Sensitivity of landscape measurements to changing grain size for fine-scale design and management. Landscape Ecology **3**: 47-53.

Cuesta, R. M., M. Garcia, and J. Retana. 2009. Factors influencing the formation of unburned forest islands within the perimeter of a large forest fire. Forest Ecology and Management **258**:71-80.

Cui, W., M.R. Ouelette, A.H. Perera, and M. Gluck. 2009. Using BFOLDS to characterize fire regimes: a case study from a boreal forest landscape. Ontario Ministry of Natural Resources, Ontario Forest Research Institute, Sault Ste. Marie, ON. Forest Research Information Paper No. 173.

Cutler, D. R., T.C. Edwards, K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, and J. Lawyer. 2007. Random Forest for classification in Ecology. Ecology **88**(11): 2783-2792.

Dahinden, C. 2011. An improved Random Forests approach with application to the performance prediction challenge datasets. I. Guyon, G. Crawley, G. Dror, A. Saffari (Eds.), Hands-on Pattern Recognition, Challenges in Machine Learning, vol. 1, Microtome, MA (2011), pp. 223–230.

Dale, V.H., L.A. Joyce, S. McNulty, R.P. Nelson, M.P. Ayres, M.D. Flannigan, P.J. Hanson, L.C.Irland, A.E. Lugo, C.J. Peterson, D.S. Simberloff, F.J. Swanson, B.J. Stocks, and B.M. Wotton. 2001. Climate change and forest disturbances. Bioscience **51**(9): 723-734.

De'ath, G. and K. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Journal of Ecology **81**(11): 3178-3192.

Delong, S.C., and D. Tanner. 1996. Managing the pattern of forest harvest: lessons from wildfire. Biodiversity and Conservations **5**: 1191-1205.

Diaz-Delgado, R., F. Lloret, and X. Pons. 2004. Spatial patterns of fire occurrence in Catalonia, NE, Spain. Landscape Ecology **19**:731-745.

Dragotescu, L. and D.D. Kneeshaw. 2012. A comparison of residual forest following fires and harvesting in boreal forests in Quebec, Canada. Silva Fennica **46**(3): 365-376.

Drever, C.R., G. Peterson, C. Messier, Y. Bergeron, and M. Flannigan. 2006. Can forest management based on natural disturbances maintain ecological resilience? Canadian Journal of Forest Research **36**:2285-2299.

Druhjell, D. 2004. Fire effects on rangeland factsheet series. http://www.agf.gov.bc.ca/range/publications/documents/fire2.htm

Duning, X., and L. Xiuzhen. 1999. Core concepts of landscape ecology. Journal of Environmental Science **11**(2):131-135.

Eberhart, K.E. and P.M. Woodard. 1987. Distribution of residual vegetation associated with large fires in Alberta. Canadian Journal of Forest Research **17**: 1207-1212.

[ESA] Ecological Society of America. 2002. Fire ecology. URL: http:/www.esa.org

Edwards, T.C., D.R. Cutler, K.H. Beard, and J. Gibson. 2007. Predicting invasive plant species occurrences in national parks: a process for prioritizing prevention. Final project report Np. 2007-1, USGS Utah cooperative Fish and Wildlife Research Unit, Utah State University, Logan, UT 84322-5290 USA.

Engelmark, O. 1999. Boreal forest disturbances. *In*: Walker, L.R. (Editor). Ecosystems of disturbed ground. ELSEVIER. 161- 186 p.

Epting, J. and D. Verbyla. 2005. Landscape-level interactions of pre-fire vegetation, burn severity, and post-fire vegetation over 16-year period in interior Alaska. Canadian Journal of Forest Research **35**: 1367- 1377.

Evans, J. 2004. Topographic ruggedness index. URL: http://arcscripts.esri.com/details.asp?dbid=12435

Evans, J. E., and S.A. Cushman. 2009. Gradient Modelling of conifer species using Random Forests. Landscape Ecology **24**: 673-683.

Flannigan, M.D., and B.M. Wotton. 2001. Climate weather and area burned. *In*: Johnson, E.A. and K. Miyanishi (Editors). Forest fires: behaviourand ecological effects. San Diego, California: Academic Press. 258-301 p.

Eunsik, P., and Y. Wenbao. 2011. Marker selection by a threshold independent measure associated with partial AUC. Proceeding of the 58th World statistical congress, Dublin, Ireland.

[FAO] Food and Agricultural Organization. 2011. Global forest resources assessment 2010: main report. FAO Forestry Paper http://www.fao.org/docrep/013/i1757e/i1757e.pdf

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

Forman, R.T.T. and M. Godron. 1986. Landscape ecology. John Wiley & Sons. 4 – 31 p.

Ferrier, S., G. Watson, J. Pearce, and M. Drielsma. 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in north-east New South Wales. I. Species-level modelling. Biodiversity and Conservation **11:** 2275–2307

Fisher, J.T., F.N. Erasmus, T.F, Witkowsi, J. van Aardt, K.J. Wessels, and G.P. Asner. 2013. Savanna woody vegetation classification – now in 3D. Applied Vegetation Science http://dx.doi.org/10.1111/avsc.12048

Foody, G.M. 2002. Status of land cover classification accuracy assessment. Remote Sensing and Environment **80**: 185-201.

Foster, D. R., D. H. Knight, and J. F. Franklin. 1998. Landscape patterns and legacies resulting from large, infrequent forest disturbances. Ecosystems. **1**: 497-510.

Friedman, J. 2009. Tutorial: Getting started with MART in R. http://www-stat.stanford.edu/~jhf/r-mart/tutorial/tutorial.pdf

Frescino,T.S, Edwards, T.C., G.G. Moisen. 2001. Modelling spatially explicit forest structural attributes using Generalized Additive Models. Journal of Vegetation Science **12**: 15-26.

Genuer, R., J. Poggi, and C. Tuleau-Malot. 2010. Variable selection using Random Forests. Pattern Recognition **31**(14): 2225-2236

Gislason, P.O., J.A. Benedictsson, and K.R. Sveinsson. 2006. Random Forests for land cover classification. Pattern Recognition Letters **27**(4): 294-300.

[GOC] Government of Canada. 2007. Canada's initial report under the Kyoto protocol. http://unfccc.int/files/national_reports/initial_reports_under_the_kyoto_protocol/application/pdf/initial_report_of_canada.pdf

Gorte, R.W. 2006. Forest fire/wildfire protection. CRS report for congress. Congressional Research Service – the Library of congress.]

Griffith, J.A. 2004. The role of landscape pattern analysis in understanding concepts of land cover change. Journal of Geographical Sciences **14**(1): 3-27.

Guisan, A. and N.E. Zimmerman. 2000. Predictive habitat distribution models in ecology. Ecological Modelling **135**:147-186.

Gustafson, E.J. 1998. Quantifying landscape spatial pattern: what is the state of the art?. Ecosystems **1**: 143-156.

Haire, S.L. and K. McGarigal. 2009. Changes in fire severity across gradients of climate, fire size, and topography: a landscape ecological perspective. Fire Ecology **5**(2): 86-103.

Haire, S.L. and K. McGarigal. 2010. Effects of landscape patterns of fire severity on regenerating ponderosa pine forests (*Pinus ponderosa*) in New Mexico and Arizona, USA. Landscape Ecology **25**: 1055-1069.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. The Elements of statistical learning: data mining, inference, and prediction. Springer Series in Statistics.

He, H.S. and D.J. Mlandenoff. 1999. Spatially explicit and stochastic simulation of forest landscape fire disturbance and succession. Ecology **80**(1): 81-99.

He, H.S., S.J. Ventura, and D.J. Mlandenoff. 2002. Effects of spatial aggregation approaches on classified satellite imagery. International Journal of Geographic Information Science **16**(1): 93-109. S.J.

Hely, C., C.M. Fortin, K.R. Anderson, and Y. Bergeron. 2010. Landscape composition influence local pattern of fire size in the eastern Canadian boreal forest: role of weather and landscape mosaic on fire size distribution in mixedwood boreal forest using the Prescribed Fire Analysis system. International Journal of Wildland Fire **19**: 1099-1109.

Hills, G.A. 1961. The Ecological Basis for Land-use Planning. Research Report 46, Research Branch, Ontario Department of Lands and Forests, Maple, Ontario. 204pp.

Hooper, E.R., P. Legendre, and R. Condit. 2004. Factors affecting community composition of forest regeneration in deforested, abandoned land in Panama. Ecology **18**(12): 3313-3326.

Hothorn, T., T. Hornik, K. Zeileis. 2008. Party: A Laboratory for Recursive Part(y)itioning.R package version 0.9-96, URL http://CRAN.R-project.org/package=party.

Huang, C., E.L. Geiger, J.A. Kupfer. 2006. Sensitivity of landscape metrics to classification scheme. International Journal of Remote Sensing **27**(14): 2927-2948.

Huff, M.H. and J.K. Smith. 2000. Fire effects on animal populations. *In:* Smith, J.K. (Editor), Wildland fire in ecosystems: effects of fire on fauna. General Technical report RMRS-GTR-42- **1**: 25-32.

[IFFN] International Forest Fire News. 2004.Forests fires in the boreal zone: climate change and carbon implications. http://www.fire.uni-freiburg.de/iffn/iffn_31/19-IFFN-31-Boreal-Climate-2.pdf

Iverson, L.R., A.M. Prasad., and A. Liaw. 2004. New machine learning tools for predictive vegetation mapping after climate change: Bagging and Random Forest perform better than Regression Tree Analysis. In: Proceedings of UK International Association for Landscape Ecology, Cirencester, UK. p.317-320.

Jepsen, J.U. 2004. Spatially explicit models in landscape and species management. National Environmental Research Institute, Ministry of the Environment, Denmark.

Johnson E. A. 1995. Fire and vegetation dynamics: studies from the North American boreal forest. Cambridge: Cambridge University Press. 125 p.

Johnson, E.A., K. Miyanishi, and J.H. Weir. 1998. Wildfires in the western Canadian boreal forest: landscape patterns and ecosystem management. Journal of vegetation science

9: 603- 610.

Jun. s. 2013. Boosted regression trees and random forests. Statistical consulting report, Department of Statistics, UBC.

Kachmara, M. and Sanchez-Azofeifa, G.A. 2003. Detection and analysis of post fire residual using medium and high resolution satellite imagery. Sustainable Forest Management Network Project report.

Kafka, V., S. Gauthier, and Y. Bergeron. 2001. Fire impacts and crowing in the boreal forest: study a large wildfire in western Quebec. International Journal of Wildland Fire **10**: 119-127.

Kerby, J.D. and S.D. Fuhlendorf. 2007. Landscape heterogeneity and fire behaviour: scale-dependent feedback between fire and grazing processes. Landscape Ecology **22**: 507-516.

King, A.W. and A.H. Perera. 2006. Transfer and extension of forest landscape ecology: a matter of models and scale. *In*: Perera, A.H., L.J. Buse and T.R. Crow. Forest landscape ecology. Springer. 19-42 p.

Klenk, N., G. Bull, and D. Cohen. 2008. What is the "END" (emulation of natural disturbance) in forest ecosystem management? An open question. Canadian Journal of Forest Research 38: 2159-2168.

Kok, K. and A. Veldkamp. 2001. Evaluating impact of spatial scales on land use pattern analysis in Central America. Agriculture, Ecosystems and Environment **85**: 205-221.

Kramkowski, V. 2012. END-games: the material and political interests of emulating natural disturbances in the Canadian boreal forest. Environments **38**(1): 81-98.

Krummel, J.R., R.H. Gardner, G. Sugihara, R.V. O'Neill, and P.R. Coleman. 1987. Landscape pattern in a disturbed environment. OIKOS **48**: 321-324.

Kuusela, K. 1992. Boreal forestry in Finland: a fire ecology without fire. Unasyla **43**(170): 22-25.

Leduc, A., Y. Bergeron, and S.Gauthier. 2007. Relationship between pre-fire composition, fire impact and post-fire legacies in the boreal forest of Eastern Canada. USDA Forest Service Proceedings RMRS-P-46CD 187-193.

Liaw, A. and A. Wiener. 2002. Classification and regression by Random Forests. The Newsletter of the R Project **2**(3):18-22.

Lin, Y.P., H.J. Chu, and C.F. Wu. 2010. Spatial pattern analysis of landslide using landscape metrics and logistic regression: a case study in Central Taiwan. Hydrological Earth System Science **7**: 3423-3451.

Linke, J., M.G. Betts, M.B. Lavigne, and S.E. Franklin. 2007. Introduction: structure, function and change of forest landscape. *In*: Wulder. M.A. and S.E. Franklin. (Editors). Understanding forest disturbance and spatial pattern. Broken Sound Parkway, NW: Taylor and Francis Group. 1-29 p.

Linn, R., J. Winterkamp, C. Edminster, J.J. Colman, W.S. Smith. 2007. Coupled influences of topography and wind on wildland fire behaviour. Int. Journal of Wildland fire **16**: 183-345.

Liu, C., M. White, and G. Newell. 2009. Measuring the accuracy of species distribution models: a review. 18th World IMACS/MODSIM Congress, Cairns, Australia.

Lund, H.G. 2011. Definitions of forest, deforestation, afforestation, and reforestation. URL: http://home.comcast.net/~gyde/DEFpaper.htm

Long, J.N. 2009. Emulating natural disturbance regimes as a basis for forest management: a North American view. Forest Ecology and Management **257**: 1868-1873.

Madoui, A., A. Leduc, S. Gauthier, and Y. Bergeron. 2010. Spatial pattern analyses of post-fire residual stands in the black spruce boreal forest of western Quebec. International Journal of Wildland Fire **19**: 1110-1126.

Magness, D.R., F. Huettmann, and J. Morton. 2009. Using Random Forest to provide predicted species distribution maps as a metric for ecological inventory and monitoring programs. Studies in Computational Intelligence (SCI) **122**: 209-229.

Malmberg, U. And M. Miljoanalys. 2001. BALANS land cover and land use classification methodology. Environment and climate programme area 3.3: Centre for Earth Observation Contract no. ENV4-CT98-0748.

Major, D.J. 2005. Effects of fire disturbance on terrestrial salamanders in mixed-coniferous forests of the Klamath/Siskiyou region of the pacific northwest. PhD Dissertation.

Manel, S., H.C. Williams, and S.J.Ormerod. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. Journal of Applied Ecology **38**: 921-931.

Marceau, D., and G. Hay. 1999. Remote contribution to the scale issue. Canadian Journal of Remote Sensing **25**(4): 357-366.

Marzano, R., E. Lingua, and M. Garbarino. 2012. Post-fire effects and short-term regeneration dynamics following high severity crown fires in a Mediterranean forest. iForest-Biogeosciences and Forestry **5**: 93-100.

Max, B., S.A. Fitzgerald, B. Parker, M.L. Main, A. Perleber, C. Schnepf, R.L. Mahoney. 2010. Reducing fire risk on your forest property. Extension service – Pacific Northwest Extension publishing.

McGarigal, K. and B.J. Marks. 1995. Fragstats: spatial pattern analysis program for quantifying landscape structure. USDA for Serv. Gen. Technical report. PNW-GTR-351.

McGarigal, K.; Cushman, M.; Neel, C.; Ene, E. FRAGSTATS; Spatial pattern analysis program for categorical maps. University of Massachusetts: Amherst, MA, USA, 2002. Available online: www.umass.edu/landeco/research/fragstats/fragstats.html ???

McKenney, D.W., J.H. Pedlar, K. Lawrence, P.A. Gray, S.J. Colombo, and W.J. Crins. 2010. Current and projected future climatic conditions for ecoregions and selected natural heritage areas in Ontario. Climate change research report CCRR-16.

McKenzie, D., C. Miller, and D.A. Falk. 2011. Toward a theory of landscape fire. *In*: McKenzie, D., C. Miller, and D.A. Falk (Editors). The landscape ecology of fire. Springer. 3-26 p.

Meddens, A.J.H., A.T. Hudak, F.S. Evans, W.A. Gould, and G. Gonzalez. 2008. Characterizing forest fragments in Boreal, Temperate and Tropical ecosystems. Ambio **37**(7): 569-576.

Mellor A., A. Haywood, C. Stone, and S. Jones. 2013. The performance of Random Forests in an operational setting for large area sclerophyll forest classification Remote Sensing **5**: 2838-2856.

Mermoz, M., T. Kitzberger, and T.T. Veblen. 2005. Landscape influences on occurrence and spread of wildfires in Patagonian forests and shrublands. Ecology **86**(10): 2705-2715.

Moisen, G.G., and T.S. Frescino. 2002. Comparing five Modelling techniques for predicting forest characteristics. Ecological Modelling **157**: 209-225.

Moody, A. and C.E. Woodcock. 1995. The influence of scale and the spatial characteristics of landscapes on land-cover mapping using remote sensing. Landscape Ecology **10**(6): 363-379.

Moritz, M.A., P.F. Hessburg, and N.A. Povak. 2011. Native fire regimes and landscape resilience. *In*: McKenzie, D., C. Miller, and D.A. Falk (Editors). The landscape ecology of fire. Springer. 51-86 p.

Morlon, H., G. Chuyong, R. Condit, S. Hubbell, D. Kenfack, D. Thomas, R. Valencia, and J.L. Green. 2008. A general framework for the distance-decay of similarity in ecological communities. Ecology Letters **11**(9): 904-917.

Munoz, J., and A.M. Felicisimo. 2004. Comparison of statistical methods commonly used in predictive Modelling. Journal of Vegetation Science **15**: 285–292.

Nelson, R.M. 2001. Water relations of forest fuels. *In*: Johnson, E.A. and K. Miyanishi (Editors). Forest fires: behaviourand ecological effects. San Diego, California: Academic Press. 79-143 p.

Nielsen, S.E., C. Johnson, D. Heard, and M. Boyce. 2005. Can models of presence-absence be used to scale abundance? Two case studies considering extremes in life history. Ecography **28**: 197-208.

Nikora, V.I., C.P. Pearson, and U. Shankar. 1999. Scaling properties in landscape patterns: New Zealand experience. Landscape Ecology **14**:17-33.

North, M., and W. Keeton. 2008. Emulating natural disturbance regimes: an emerging approach for sustainable forest management. *In*: Lafortezza, R., J. Chen, G. Sanesi and T. Crow (Editors), Landscape ecology: Sustainable management of forest landscapes. Springer-Verlag Press, The Netherlands, 341-372 p.

[NRC] Natural Resource Canada. 2014. Forests in Canada. http://www.nrcan.gc.ca/forests/canada/13161

Olden. J.D., D.A. Jackson, P.R. Peres. 2002. Predictive models of fish species distributions: a note on proper validation and chance predictions. Transactions of the American Fisheries Society **131:** 329–336.

Olsson, R. 2009. Boreal forest and climate change. Air pollution and climate series 23.

O'Neill, E.G., R.V. O'Neill, and R.J. Norby. 1991. Hierarchy theory as a guide to Mycorrhizal research on large-scale problems. Environmental pollution **73**: 271-284.

O'Neill, R.V., J.R. Krummel, R.H. Gardner, G.Sugihara, B. Kackson, D.L. DeAngelis, B.T. Milne, M.G. Turner, B. Zygmunt, S.W. Christensen, V.H. Dale, and R.L. Graham. 1988 Indices of landscape pattern. Landscape Ecology **1**(3):153-162.

[OMNR] Ontario Ministry of Natural Resources. 2001. Forest management guide for natural disturbance pattern emulation (version 3.1). Ontario Ministry of Natural Resources, Forest Management Branch, Sault Ste. Marie, ON.

[OMNR] Ontario Ministry of Natural Resources. 2005. Ontario Spatial Landcover Database, 2000, second ed. Ontario Ministry of Natural Resources, Peterborough, ON.

[OMNR] Ontario Ministry of Natural Resources. 2009. Ontario's forest: the boreal forest. http://www.mnr.gov.on.ca/stdprodconsume/groups/lr/@mnr/@forests/documents/document/241209.pdf

[OMNR] Ontario Ministry of Natural Resources. 2010. Forest management guide for conservation biodiversity at the stand and site scales: background and rationale for direction. Toronto: Queen's printer for Ontario.

[OMNR] Ontario Ministry of Natural Resources. 2011. Large fire mapping Guide, V1.1. Ontario Ministry of Natural Resources, Peterborough, ON.

[OMNR] Ontario Ministry of Natural Resources. 2014. Forest management guide for boreal landscapes. Toronto: Queen's printer for Ontario.

Openshaw, S. 1984: Ecological fallacies and the analysis of areal census data. Environment and Planning **16**: 17–31.

Ostapowicz, K., P. Vogt, K.H. Riitters, J. Kozak, C. Estreguil. 2008. Impact of scale on morphological spatial pattern of forest. Landscape Ecology **23**:1107-117.

Pal, M. 2005. Random Forest classifier for remote sensing classification. International Journal of Remote Sensing **26**(2): 217-222.

Parisien, M., S.A. Parks, M. A. Kramchuk, M. D. Flannigan, L.M. Bowman, and M.A. Moritz. 2011. Scale-dependent controls on the area burned in the boreal forest of Canada, 1980- 2005. Ecological Applications **21**(3): 789-805.

Pearce, J., S. Ferrier, and D. Scotts. 2001. An evaluation of the predictive performance of distributional models for flora and fauna in north-east New South Wales. Journal of Environmental Management **62**: 171-184.

Pearce, J. and S. Ferrier. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. Ecological Modelling **133**: 225-245.

Perera, A.H., and D.L. Euler. 2000. Landscape ecology in forest management: an introduction: *In*: Perera, A.H., D.E. Euler and I.D. Thompson. (Editors). Ecology of a managed terrestrial landscape: patters and processes of forest landscapes in Ontario, University of British Columbia Press, Vancouver, BC. 3-11p.

Perera, A.H. and L.J. Buse. 2004. Emulating natural disturbance in forest management: an overview. *In:* A.H. Perera, L.J. Buse, and M.G. Weber. (Editors). Emulating Natural Forest Landscape Disturbance: Concepts and Applications. Columbia University Press, New York, NY. 3-7 p. [1]

Perera, A.H., L.J. Buse, and R.G. Routledge. 2007. A review of published knowledge on post-fire residuals of published knowledge on post-fire residuals relevant to Ontario's policy directions for emulating natural disturbance. Ontario Ministry of Natural Resources, Ontario Forest Research Institute, Sault Ste. Marie, Ontario. Forest Research Report No. 168.

Perera, A.H. T.K. Remmel, L.J. Buse, and M.R. Ouellette. 2009a. An assessment of residual patches in boreal fires, in relation to Ontario's policy directions for emulating natural forest disturbance. Ontario Ministry of Natural Resources, Ontario Forest Research Institute, Sault Ste. Marie, Ontario. Forest Research Report No. 169.

Perera, A.H., B.D. Dalziel, L.J. Buse, and R.G. Routledge. 2009b. Spatial variability of stand-scale residuals in Ontario's boreal forest fires. Canadian Journal of Forest Research **39**: 945-961.

Perveen, S., and L,A. James. 2010. Multi-scale effects on spatial variability metrics in global water resources data. Water Resources Management. doi:10.1007/s11269-009-9530-2.

Peters, J., B.D. Baets, N. E. Verhoest, R. Samson, S. Degroeve, P.D. Becker, and W. Huybrechts. 2007. Random Forests as a tool for ecohydrological distribution modelling. Ecological Modelling **207**: 904-318.

Pickett, S.T.A., J. Wu, and M,L. Cadenasso. 1999. Patch dynamics and the ecology of disturbed ground: a framework for synthesis. *In*: Walker, L.R. (Editor). Ecosystems of disturbed ground. ELSEVIER. 689-707 p.

Prasad, A.M., L.R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques:bagging and Random Forests for ecological prediction. Journal of Ecosystems **9**:181-199.

Pretorius, E. 2013. Image classification towards mapping of vegetation structure: a practical approach. International Cartographic Conference, Dresden, Germany.

R Development Core Team. 2013. R: a language and environment for statistical computing. Vienna: R foundation for Statistical Computing. Available at http://www.r-project.org

Refaeilzadeh, P., L. Tang, and H. Liu. 2008. Cross-validation. Online lecture notes. Arizona State University.

Remmel, T.K. and F. Csillag. 2003. When are two landscape pattern indices significantly different? Journal of Geographical Systems **5**(4): 331-351.

Remmel, T.K. and A.H. Perera. 2009. Mapping natural phenomena: boreal forest fires with nondiscrete boundaries. Cartographica **44**(4): 274-288.

Riitters, K.H., R.V. O'Neill, C.T. Hunsaker, J.D. Wickham, D.H. Yankee, S.P. Timmins, K.B. Jones, and B.L. Jackson. 1995. A factor analysis of landscape pattern and structure metrics. Landscape Ecology **10**(1): 23-39.

Riley, S., S.D. DeGloria, and R. Elliot. 1999. A terrain ruggedness index that quantifies topographic heterogeneity. International Journal of Sciences **1**(4): 23-27.

Rochadi, A., A.S. Adhikerana, R. Ubaidillah, and R. Suharna. 1999. Preliminary study of the ecological impact of forest fires in G. Massigit, G. Gede-Pangrango National park, West Java. Korean Journal of Ecology **23**(2):125-129.

Robert-Granie, C., K.A. Cao, and M. SanCristobal. 2009. Predicting qualitative phenotypes from microarray data-the Eadgene pig data set. In proceedings of BioMed Central, EADGEN and SABRE post analyses, Lelystad, The Netherlands pages 1-53.

Rodriguez, J.J. and L.I., Kuncheva. 2006. Rotation Forest: a new classifier ensemble method. *IEEE* Transactions on Pattern Analysis and Machine Intelligence **28**(10): 1619-1630.

Routledge, R.G. 2007. Sampling methodology for studying boreal post-fire residual stand structure with high resolution aerial photography and field plots. Ontario Ministry of Natural Resources, Ontario Forest Research Institute, Sault Ste. Marie, Ontario. Forest Research Report No. 169.

Rowe, J.S., and G.W. Scotter. 1973. Fire in the boreal forest. Quaternary Research **3**:444-464.

Runesson, T. 2011. Boreal forests: Ontario's northwest forest. borealforest.org.

Ryan, K.C. 2002. Dynamic interactions between forest structure and fire behaviour in boreal ecosystems.Silva Fennica **36**(1):13-39.

Saura, S. 2004. Effects of remote sensor spatial resolution and data aggregation on selected fragmentation indices. Landscape Ecology **19**: 197-209.

Schmiegelow, F. K. A., D. P. Stepnisky, C. A. Stambaugh, and M. Koivula. 2006. Reconciling salvage logging of boreal forests with a natural-disturbance management model. Conservation Biology 20:971-983.

Schroeder D. and A. H. Perera. 2002. A comparison of large-scale spatial vegetation patterns following clearcuts and fires in Ontario's boreal forests. Forest Ecology and Management **159**: 217-230.

Schroff, F., A. Criminisi, and A. Zisserman. 2008. Object class segmentation using Random Forests. Proceeding of the British Machine Vision Conference (BMVC).

Schwartz, M. W., L.R. Iverson, A.M. Prasad, S.N. Matthews, and r.J. O'Connor. 2006. Predicting extinctions as a result of climate change. Ecology **87**: 16111615

Smith, K.S., and R. Hendry. 1998. Residual tree left by fire: final report. Enhanced Forest Management Pilot Project Report.

Spectranalysis. 2005. IKONOS land cover classification for the assessment of forest burns. Internal report prepared for Forest Landscape Ecology Program, Ontario Forest Research Institute, Ministry of Natural Resources.

Spectranalysis. 2004. Introduction to the Ontario land cover Data Base. Internal report prepared for the inventory, monitoring, and assessment section Science and Information Branch, Ontario Ministry of Natural Resources.

Strobl, C. T. A. Boulesteix, T. Kneib, t. Augustin, and A. Zeileis. 2008. Conditional variable importance for Random Forests. BMC Bioinformatics **9**:307.

Strobl, C., J. Malley, and G. Tutz. 2009a. An introduction to Recursive Partioning, rationale, application and characteristics of classification and regression trees, bagging and random forests. Psychological methods. **14**(4): 323-348.

Strobl, C., T. Hothorn, and A. Zeileis. 2009b. Party on! A new, conditional variable importance measure for random forests available in the party package. The R Journal. **1**(2): 14-17.

Swets, J. 1988. Measuring the accuracy of diagnostic systems. Science **240**: 1285-1293.

Swystun, M.B., J.M. Psyllakis, and R.M. Brigham. 2001. The influence of residual tree patch isolation on habitat use by bats in central British Columbia. Acta Chiropteroligca **3**(2): 1997-201.

Thompson, I.D. 2000. Forest vegetation of Ontario: factors influencing landscape change: *In*: Perera, A.H., D.E. Euler and I.D. Thompson. (Editors). Ecology of a managed terrestrial landscape: patters and processes of forest landscapes in Ontario, University of British Columbia Press, Vancouver, BC. 30-51p.

Turner, M.G. 1989. Spatial and temporal analysis of landscape patterns. Landscape Ecology **4**(1): 21-30.

Turner, M. G., W.H. Romme, R.H. Gardner, and W. W. Hargrove. 1997. Effects of fire size and pattern on early succession in Yellowstone National Park. Ecological Monographs **67**(4): 411-433.

Turner, M. G. 2005. Landscape ecology: what a state of the science? Landscape Ecology **36**:319-344.

[UNEP] United Nations Environment Programme. 2009. Global ecological forest classification and forest protected area gap analysis. Freiburg University Press

Van Wagtendonk, J.W. 2004. Fire and landscapes: patterns and processes. USDA Forest Service Gen. Technical Report. PSW-GTR-193.

Vera, P. 2001. Fire history and ecology of remnant forest patches in the Sub-Boreal Pine-Spruce Zone of Central British Columbia. M.S. Thesis, University of British Columbia, Vancouver, British Columbia.

Viegas, D.X. 1993. Fire behaviour and fireline safety. Annual Mediterranean Burns Club **6**(3): 179-185.

Vinatier, F., P. Tixier, P.F. Duyck, and F. Lescourret. 2010. Factors and mechanisms explaining spatial heterogeneity: a review of methods for insect populations. Methods in Ecology and Evolution doi: 19.1111/j.2041-210X.2010.00059.x.

Walz, U. 2011. Landscape structure, landscape metrics and biodiversity. Living reviews in landscape research **5**: 3:35.

Watts, J. D. and R. L. Lawrence. 2008. Merging Random Forest classification with an object-oriented approach for analysis of agricultural lands. The international archives of the Photogrammetry, Remote sensing and spatial information sciences, Vol. XXXVII, Beijing 579-582.

Weber, M. G., and B.J. Stocks. 1998. Forest fires and sustainability in the boreal forests of Canada. Royal Swedish Academy of Science **27**(7): 545-550.

Weber, M.G. and M.D. Flannigan. 1997. Canadian boreal forest ecosystem structure and function in a changing climate: impact of fire regimes. Environment reviews **5**: 145-166.

Wells, J., d. P.L Roberts, R. Cheng and M. Darveau. 2010. A forest of blue—Canada's boreal forest: the world's waterkeeper. International Boreal Conservation Campaign, Seattle. 74 pp.

Whelan, R.J. 1995. The ecology of fire. Cambridge University Press, Cambridge, UK. 1-56 p.

White. P.S. 1979. Pattern, process and natural disturbance in vegetation. Botanical Review **45**(3): 229-299.

Wichham, J.D., R.V. O'Neill, K. H. Riitters, T.G. Wade, and K.B. Jones. 1995. Sensitivity of selected landscape pattern metrics to land cover misclassification and differences in land cover composition. Photogrammetric Engineering and Remote Sensing **63**(4): 397-402.

Wiens, J.A. 1989. Spatial scaling in ecology. Functional Ecology **3**: 385-397.

Wiken, B., D. Gauthier, I. Marshall, K. Lawton, and H. Hirvonen. 1996. A perspective on Canada's ecosystems: an overview of the terrestrial and marine ecozones. The Canadian Council on Ecological Areas, Ottawa, Ontario.

William, J.C, P.A. Gray, P.W. Uhling, and M.C. Wester. 2009. The ecosystems of Ontario, part 1: Ecozones and Ecoregions. Science and Information Branch: inventory, monitoring, and assessment section, Technical Report SIB TER IMA TR-01.

Wilfried, T., A. Miguel, and L. Sandra. 2003. Generalized models vs. classification tree analysis: predicting spatial distributions of plant species at different scales. Journal of Vegetation Science **14**:669-680.

Wong, D.W. 2009. The Modifiable Area Unit Problem (MAUP) In: the SAGE Handbook of Spatial Analysis, London: Sage Publications.

Wooddcock, C.E., and A.H. Strahler. 1987. The factor of scale in remote sensing. Remote Sensing of Environment **21**: 311-332.

Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing **25**(4): 367-380.

Wu, J., D.E. Jelinski, M, Luck, and P.T. Tueller. 2000. Multi-scale analysis of landscape heterogeneity: scale variance and pattern metrics. Geographic information sciences **6**(1): 6-19.

Wu, J. W. Shen, W. Sun, and P. T. Tueller. 2002. Empirical patterns of the effects of changing scale on landscape metrics. Landscape Ecology **17**: 761-782.

Wu, J. and Y. Qi. 2000. Dealing with scale in landscape analysis: an overview. Geographic information sciences **6**(1): 1-5.

Wu, J. 2004. Effects of changing scale on landscape pattern analysis: scaling relations. Landscape Ecology 19: 125-138.

Wu, J. and H. Li. 2006. Perspective and methods of scaling. *In*: Wu, J., K.B. Jones, H. Li, and O.L. Loucks. Scaling and uncertainty analysis in ecology. Springer. 3-15 p.

Zaniewski, A.E., A. Lehmann, M.C. Overton. 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. Ecological Modelling **157**: 261-280.

Zhu, M., N. Jiang, J. Li, J. Xu, and Y. Fan. 2006. The effects of sensor spatial resolution and changing grain size on fragmentation indices in urban landscape. International Journal of Remote Sensing **27**(21): 4791-4805.

Zweig, M.H. and G. Campbell. 1993. Receiver-Operating Characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical Chemistry **39**(4): 561-577.