

**Assessing Surface Fuel Hazard in Coastal Conifer Forests through the Use of
LiDAR Remote Sensing**

by

Christos Koulas
B.Sc., University of Victoria, 2007

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Geography

© Christos Koulas, 2013
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy
or other means, without the permission of the author.

Supervisory Committee

Assessing Surface Fuel Hazard in Coastal Conifer Forests through the Use of LiDAR Remote Sensing

by

Christos Koulas
B.Sc., University of Victoria, 2007

Supervisory Committee

Dr. K. Olaf Niemann, Department of Geography
Supervisor

Dr. Trisalyn Nelson, Department of Geography
Departmental Member

Abstract

Supervisory Committee

Dr. K. Olaf Niemann, Department of Geography
Supervisor

Dr. Trisalyn Nelson, Department of Geography
Departmental Member

The research problem that this thesis seeks to examine is a method of predicting conventional fire hazards using data drawn from specific regions, namely the Sooke and Goldstream watershed regions in coastal British Columbia. This thesis investigates whether LiDAR data can be used to describe conventional forest stand fire hazard classes. Three objectives guided this thesis: to discuss the variables associated with fire hazard, specifically the distribution and makeup of fuel; to examine the relationship between derived LiDAR biometrics and forest attributes related to hazard assessment factors defined by the Capitol Regional District (CRD); and to assess the viability of the LiDAR biometric decision tree in the CRD based on current frameworks for use. The research method uses quantitative datasets to assess the optimal generalization of these types of fire hazard data through discriminant analysis. Findings illustrate significant LiDAR-derived data limitations, and reflect the literature in that flawed field application of data modelling techniques has led to a disconnect between the ways in which fire hazard models have been intended to be used by scholars and the ways in which they are used by those tasked with prevention of forest fires. It can be concluded that a significant trade-off exists between computational requirements for wildfire simulation models and the algorithms commonly used by field teams to apply these models with remote sensing

data, and that CRD forest management practices would need to change to incorporate a decision tree model in order to decrease risk.

Table of Contents

Supervisory Committee	ii
Abstract	iii
Table of Contents	v
Table of Tables	vii
Table of Figures	viii
Abbreviations	ix
Acknowledgments	x
Dedication	xi
Chapter 1: Introduction	1
1.1 Research Context	1
1.2 Research Problem	4
1.3 Primary Research Objectives	6
1.4 Research Questions	7
1.5 Overview of Research Method	8
1.6 Thesis Organization	9
Chapter 2: Literature Review	10
2.1 Introduction	10
2.2 Quantifying Fuel Variables Associated With Fire Hazards	10
2.3 Classifying Fire Hazard Fuel Types	14
2.4 Mapping Fire Hazard	21
2.5 Viability of LiDAR Biometrics in Relation to Forest Attributes	27
2.6 Summary of Literature Findings	31
Chapter 3: Method	33
3.0 Introduction	33
3.1 Study Area	34
3.2 Sampling Design	35
3.3 Data Collection	36
3.3.1 LiDAR data	36
3.3.2 LiDAR Biometrics	37
3.3.3 LiDAR Data Suitability	39
3.3.4 Field-derived Data	41
3.4 Data Analysis	44
3.4.1 Classification	44
3.4.2 Discriminant Analysis	45
3.4.3 Principal Component Analysis	47
3.5 Summary of Research Method	51
Chapter 4: Findings and Discussion	52
4.0 Introduction	52
4.1 Discriminant Analysis Findings	52
4.1.1 Discriminant Analysis Results	52
4.1.2 Results of Surface Fire Hazard Class vs. Canopy Base Height	54
4.2 Discriminant Analysis Discussion	56
4.2.1 Surface Fire Hazard and Surface Forest Attributes	56

4.2.2	Canopy Fire Hazard and Canopy Forest Attributes.....	58
4.3	Results of Zonal Polygon Statistics of Forest Attributes.....	62
4.4	Zonal Polygon Statistics Discussion.....	64
4.5	Discussion of Factors Limiting Findings.....	66
4.5.1	Field Derived Data Limitations.....	67
4.5.2	LiDAR Derived Data Limitations.....	70
4.5.3	Spatial Resolution and Methodological Limitations.....	73
4.6	Summary of Findings.....	74
Chapter 5: Conclusions.....		77
5.0	Introduction.....	77
5.1	Analysis of Research Question Findings.....	77
5.1.1	Research Question 1.....	77
5.1.1	Research Question 2.....	79
5.1.1	Research Question 3.....	81
5.2	Implications of Findings for CRD Forest Management Practices.....	82
5.3	Implications for Strategic Fire Remote Sensing Research.....	88
5.4	Recommendations for Further Research.....	89
5.4.1	Recommendations for Segmentation Decision Tree Classifier Research....	89
5.4.2	Recommendations for discriminant analysis research.....	89
5.4.3	Recommendations for Spatial Resolution and Sampling Methodologies....	90
5.5	Conclusions.....	90
References.....		92
Appendix A: Methodology Flowchart.....		100
Appendix B: LiDAR Flightlines.....		102
Appendix C: Randomly Sampled Points.....		103
Appendix D: Blackwell Surface Fire Hazard Classes.....		104
Appendix F: Blackwell Canopy Fire Hazard Classes.....		105
Appendix G: Blackwell Canopy Closure Classes.....		106
Appendix H: LiDAR Gap Fraction 75.....		107

Table of Tables

Table 1 - LiDAR data technical specifications	36
Table 2 - LiDAR data bands used in analysis and relevant algorithms	38
Table 3 - Mean and standard deviation of LiDAR biometric between sample and population	39
Table 4 - Distribution of hazard classes for surface fuels in training dataset	40
Table 5 - Distribution of hazard classes for surface fuels in validation dataset	40
Table 6- Distribution of hazard classes for canopy fuels in training dataset	40
Table 7 - Distribution of hazard classes for canopy fuels in validation dataset	40
Table 8 - Estimated relative flammability of selected vegetation types	42
Table 9 -Estimated surface fire hazard in areas with understory vegetation of low relative flammability	42
Table 10 - Estimated surface fire hazard in areas with understory vegetation of moderate relative flammability	43
Table 11 - Estimated surface fire hazard in areas with understory vegetation of high relative flammability	43
Table 12 - Estimated crown fire hazard using percentage crown closure and height to base of crown	44
Table 13 - Estimated total fire hazard using surface fire hazard and crown fire hazard	44
Table 14 - Correlation matrix of input variables to principal component analysis for surface hazard discrimination	48
Table 15 - Correlation matrix of input variables to principal component analysis for canopy discrimination	48
Table 16 - Summary table for discriminant analysis (DA) results using various discriminating variables	53
Table 17 - Summary of discriminant analysis of understory fuel loading.....	54
Table 18 - Summary of discriminant analysis of surface fire hazard	55
Table 19 - Summary of discriminant analysis of canopy base height	56
Table 20 - Summary table for field plots & class thresholds.....	84

Table of Figures

Figure 1 - Flowchart for the generation of above-ground metrics.....	22
Figure 2 - Example of Fuel Fire Index: FFI map of 1.8m AVIRIS derived imagery.....	26
Figure 3 - Study Area located on Southern Vancouver Island	34
Figure 4 - Fire hazard model triangle	41
Figure 5 - Example of input variables for statistical analysis.....	50
Figure 6 - Mean values of skewness, gap fraction 0, and kurtosis for polygons showing surface fire hazard class.....	63
Figure 7 - Mean values of skewness, gap fraction 0, and kurtosis for polygons showing canopy base height class.....	63
Figure 8- Mean values of skewness, gap fraction 50 & 75 for polygons showing canopy closure	64
Figure 9 - Mean values of gap fraction 75 and the mean standard deviation of a polygon separated to show canopy closure class	64
Figure 10 - Visualization of Surface fire hazard class definition.....	69
Figure 11 - Profile of raw discrete LiDAR returns undergoing classification procedure.....	72
Figure 12 - Decision tree classification	85
Figure 13 - Surface fire hazard risk classification in Sooke watershed.....	86
Figure 14 - Adapted Fire Triangle from Blackwell	87
Figure 15 - Segmented approach in DST classification.....	89

Abbreviations

AGB	Above ground biomass
BC	British Columbia
CBD	Canopy bulk density
CRD	Capitol regional district of Greater Victoria region
DEM	Digital elevation model
DST	Decision support tree
FARSITE	Spatially explicit, two-dimensional deterministic fire growth simulation model
FBP	Canadian forest fire behaviour prediction system
FFI	Fuel fire index
IFOV	Instantaneous field of view
LiDAR	Light detection and ranging
MNF	Minimum noise fraction transformation
NPV	Non-photosynthetic vegetation
PCA	Principal component analysis
PV	Photosynthetic vegetation
RADAR	Radio detection and ranging
RMSE	Root mean square error
RS	Remote sensing

Acknowledgments

This thesis took longer than most to complete and that, paired with the sheer magnitude of people whose help was indispensable has caused me to likely forget some names; for that I am sorry, I truly appreciate your efforts.

I'd like to start by thanking Olaf Niemann for giving me the opportunity to pursue this research. I'm certain my antics gave him pause more than once over the course of this research project, so thank you for sticking it out.

They say it takes a village to raise a child, I'd wager it takes a lab to produce a thesis. Thank you to all the members of the Hyperspectral and LiDAR research group for their camaraderie, particularly Ben Arril, Fabio Visintini, Rafael Loos, Geoff Quinn, Roger Stephan, Matt Tomlins, and Diana Parton, who individually warrant a page of thanks but get rifled off on a list instead. Your snippets of code, countless rebounding of thoughts, editing and modifying, encouragement and countless hours of help have allowed this research to be completed and I am forever in your debt.

Thank you to Tim Bogle who squeezed in my rough draft into daily life of PhD, full-time job and full-time dad. You did me a great service and gave me a needed boost.

I'd like to also thank Joel Ussery at the CRD for his assistance and efforts contributing to the research. I would not have been able to collect the necessary data and ask the necessary questions if he didn't make time for this project; thank you.

I'd also like to thank Lisa Thomas-Tench for all her support and encouragement, you were and will always be my thesis angel.

Finally, I'd like to thank my parents and sisters for all their support over the years.

Dedication

This thesis is dedicated to my beautiful and loving wife Carla Koulas. She would have moved to Alcatraz itself to support me in my efforts to pursue this degree and I will never forget the personal sacrifice she made. Thank you for your limitless encouragement, motivation and support over 5-½ years of a 2-year program. I would have never been able to finish this thesis without your help.

I'd also like to extend this dedication to my mother Heather Marshall and my late grandfather Ernie Marshall for inspiring me to pursue grad studies.

Chapter 1: Introduction

1.1 Research Context

Wildland fires are a major environmental issue in a wide range of world ecosystems (Arroyo, 2008). Ensuring quality geospatial information about the location and severity of fuels in forest stands will significantly improve the management of reducing the threat of wildfire. Within an ecosystem perspective, maps of fuels and fire regimes are essential for understanding the ecological relationships between wildland fire and landscape structure, composition, and function (Rollins, Keane, & Parsons, 2004). For forest fire managers, the spatially explicit and comprehensive information on fuels and fire regimes produce key information for managing wildland fire hazard and risk (Rollins, Morsdorf, van der Linden, et al., 2004). The spatial distribution and condition of fuel can be instrumental for forest fire risk and hazard assessments as well as planning practices (Koetz et al., 2008). In other words, they allow forest managers to make more informed decisions when maintaining the forests. Forest fire management in such a wildland urban interface demands a dedicated and comprehensive monitoring of fuel types and their properties at high spatial resolution in order to implement the necessary management policies and practices (Koetz et al., 2008).

Forest fire management assessment has been achieved through the use of remote sensing (RS) techniques. Data from RS platforms provide accurate, spatially explicit information for wildfire mitigation. Forest fire management in terms of remote sensing can be divided into two components: (1) the tactical RS of fire risk; mapping the geo-location of active fires and making that information available to fire suppression managers and behaviour models; and (2) the strategic RS of fire risk; the sampling the

earth's surface and providing a picture of vegetation types and a forest stand's vertical and horizontal distribution (Koulas, 2009). Strategic remote sensing, in particular, can offer quality data for scientific or management inquiries. The use of imaging spectroscopy (hyperspectral) and light detection and ranging (LiDAR) provides the best representation of the vertical and horizontal continuity of fuels as well as explicit geometric information (Koetz et al., 2008; Asner et al., 2007). In this thesis, the term biometrics is used to indicate the biometrics collected by LiDAR.

There are many different types of strategic remote sensing that could be utilized in the prevention of forest fire. The benefit of strategic remote sensing is the ability to collect data over large landscapes for a fraction of the cost and time of conventional field surveys and at higher resolutions. Optical remote sensing is the study of light and other radiation as a function of wavelength that has been emitted, reflected, or scattered from a solid, liquid, or gas (Clark, 1999). In the case of imaging spectroscopy (hyperspectral), it is the simultaneous and continuous acquisition of a large number of narrow spectral bands over the electromagnetic spectrum (Goetz et al., 1985). All optical imaging is based on the interaction of electromagnetic radiation with matter, and provides a precise analytical method for finding the constituents in material having an unknown chemical composition (Clark, 1999). Optical remote sensing can be used to inventory tree species, health, and fuel classes of forest stands. Optical sensors are passive, as they rely on an energy source such as the sun to illuminate an area of interest. LiDAR, alternatively, is an active technology based on the emission of light pulses and interpreting the backscattered light. The measurement principle of LiDAR against forest stands relies on laser pulses

propagating vertically through the canopy, while scattering events within the vegetation are recorded as a function of time (Koetz et al., 2008).

The response obtained by LiDAR depends on the vertical distribution of the canopy elements such as the foliage, branches, and trunks, as well as the underlying terrain (Koetz et al., 2008; Harding et al., 2001). LiDAR has been used to produce biomass estimations in forested areas where ground observations are difficult, forest structure is complex and heterogeneous, and in denser canopies than those that can be accurately analyzed by optical or Radio Detection and Ranging (RADAR) (Lefsky et al., 1999). For example, as with optical systems, reflectance values alone do not provide adequate information for biomass estimation as values correspond to the combination of trees, shadow, understory vegetation and bare ground; the assumption is that there is a predictable relationship between the two-dimensional structural properties of a forest that can be sensed, and the three-dimensional structural properties of a forest that are required for forest volume estimations (Lefsky et al., 1998). LiDAR techniques, in comparison, have been used to accurately estimate fuel load parameters for forests for both fuel mapping and fire behaviour modeling.

Strategic remote sensing techniques have been beneficial to forest fire research. Most notably, it has been used to quantify biomass (Andersen et al., 2005; Riaño et al., 2003; Skowronski, Clark, Nelson, et al., 2007), classify fuel types (Arroyo et al., 2008; Lasaponara et al., 2006; Mutlu, Popescu, Stripling, & Spencer, 2008) and map fire hazard (Varga et al., 2008). These studies are of particular interest because they form the foundation for this scientific inquiry. Additionally, LiDAR has a significant potential to reduce the cost of initial forest inventories and create opportunities for subsequent ones,

because conventional methods are used sparingly (Varga et al., 2008). To this end, the contribution of this research is to evaluate the use of discriminant analysis to predict conventionally obtained fire hazard classes using a LiDAR derived data source in coastal conifer forests in coastal British Columbia. It is thought that if the statistical values of LiDAR can classify conventional forest fire hazard datasets, it would be a small contribution in further developing this idea in the scientific community.

1.2 Research Problem

The research problem that this thesis seeks to examine is a method of predicting conventional fire hazards using data drawn from specific regions, namely the Sooke and Goldstream watershed regions in coastal British Columbia (BC). These regions provide a prime example of the need to reduce the risk of fire, in that a catastrophe in this area could damage Greater Victoria's main drinking water supply and lead to social, physical, and economic challenges. For example, in a nearby region in the same Canadian province, a firestorm in 2003 destroyed over 260 000 hectares of forest and cost the province an estimated 700 million dollars (Filmon, 2004). Following the 2003 record fire season, an independent provincial review was established to evaluate the overall response to the emergency and make recommendations for the future. One of the key recommendations for reducing fuel build up was the assessment of fire prone ecosystems within or adjacent to wildland urban interfaces for fire risk reduction (Filmon, 2004). The fuel reduction initiative had not progressed quickly enough to impact the results of another firestorm in 2004, which burned over 220 000 hectares of forest. After more than half a decade since the recommendations, the fires in BC during the summer of 2009

were some of the worst in history; the direct cost of fighting fires was more than quadruple that of the firestorm in 2004 (BCFS, 2010).

On Vancouver Island, the Sooke and Goldstream Watersheds are the location of the fresh water storage for the Greater Victoria region that service 13 municipalities. The Capitol Regional District (CRD) is the regional government and the watershed protection division is responsible for all the forests within the watershed including the health of the trees and associated hazards that exist there. An extreme forest fire in these watersheds is considered a catastrophic event (J. Ussery, personal communication, May 31, 2011). The first and biggest shock to the catchment would be a drastic debris flow, consisting of ash and sediment, creating a water supply emergency for the local population. Longer-term effects decrease the total capacity of the reservoir, stabilizing only after vegetation had returned to the landscape. Understandably, a considerable amount of energy goes into preventing a fire from occurring in the surrounding forests. Scientific research, investigation into the indicators for insect and disease (Quinn, 2011) and fuel moisture content (Visintini, 2011) in the watershed, are seen by the CRD as additional pillars of defence against a wildfire. The CRD has also mapped fuels and fire hazard extensively over the years. It is thought that the comprehensive monitoring of fuels in areas such as this watershed will significantly optimize fuel thinning and risk reduction practices, increasing efficiency and efficacy.

Despite all of these efforts, there are significant challenges associated with the way that the CRD's framework has been devised, and therefore, there is a risk to the community. Currently, there is no statistically-reliable means of fire hazard mapping in coastal conifer forests like the Sooke Watershed. This is because the CRD has limited its

fire hazard specifications to a framework defined by independent contractor research performed by the company BA Blackwell and Associates (hereafter referred to as Blackwell), employing 433 surveyed plots in the watershed, which may or may not be problematic over the long term. It is evident that this contractor's analyses' classification accuracy has had problems inherent from the visual estimation of the forest attributes in the field dataset. Conventional field-based hazard assessments attempt to inventory conditions over a large landscape, which can be problematic in terms of precision and accuracy. Additionally, issues associated with class definitions, traditional hazard frameworks and spatial resolutions also negatively influenced the overall accuracy of discriminant classifications, in that the coastal conifer forests in the Sooke Watershed do not have a complete classification of fire hazard available for mitigation efforts (J. Ussery, personal communication, May 31, 2011). Given these issues, this thesis examines the viability of the current system, and proposes a new framework to better manage the combination of traditional fire hazard framework and highly robust LiDAR datasets.

1.3 Primary Research Objectives

This thesis investigates whether LiDAR data can be used to describe conventional forest stand fire hazard classes. Strategic wildfire research using LiDAR is a relatively novel research direction. LiDAR remote sensing techniques produce valuable forest stand metrics, but few studies have mapped fire hazard and communicated it effectively to forest managers. This scientific inquiry uses quantitative datasets to assess the optimal generalization of these types of fire hazard data. This will be achieved through the following research objectives:

1. To discuss the variables associated with fire hazard, specifically the distribution and makeup of fuel;
2. To examine the relationship between derived LiDAR biometrics and forest attributes related to hazard assessment factors defined by the CRD;
3. To assess the viability of the LiDAR biometric decision tree in the CRD based on current frameworks for use.

1.4 Research Questions

Given the primary research objectives listed above, the following research questions guide the assessment of data in this thesis.

RQ1. What are the variables associated with fire hazard that should be identified for hazard assessments?

RQ2. What is the relationship between derived LiDAR biometrics and forest attributes related to hazard assessment factors defined by the CRD?

RQ3. Is the LiDAR biometric hazard analysis in the CRD based on current frameworks for use viable or not viable?

Research Question 1 will be answered through an assessment of variables associated with fire hazards which will take place in the literature review. Research Question 2 will be answered through discriminant analysis, where the independent variables are six fire hazard variables defined and measured by the CRD: surface fire hazard, understory vegetation cover, fuel loading, canopy fire hazard, canopy percent coverage and canopy base height, and the dependent variable is accurate mapping, as measured by seven LiDAR metrics: mean, mode, rugosity, skewness, kurtosis, 85th percentile, gap fraction, and variance. Research Question 3 will be answered through an

assessment of the findings from Research Question 2 in combination with the findings from the research literature.

1.5 Overview of Research Method

The research method used for this study is primarily quantitative in nature. As noted above, the contribution of this research is to evaluate the use of discriminant analysis to predict conventionally obtained fire hazard classes (traditionally derived through photo-interpretation and ground plots), using a LiDAR derived data source, in coastal conifer forests. The main challenge is to find a methodology that will combine a high resolution, statistical and consistent dataset, with one that is heuristic, and a valuable professional estimate. In this study biometrics of discrete 20 metre LiDAR data are used, relating to the shape of a probability distribution and coincident field data that includes expert derived fire hazard classes. The viability of using discriminant analysis to correctly classify LiDAR biometrics and their principal components to conventional values for fire hazard is tested.

The methodology utilizes the extensive field assessment already performed by Blackwell, employing 433 surveyed plots in the Sooke Watershed. The conceptual framework and definitions of fire hazard developed by the CRD, as well as the fundamental forest relationships presented in the literature review are employed. Stratified random sampled points from the LiDAR dataset have been decorrelated using principal component analysis. In turn, discriminant analysis has determined to which fire hazard class the representative pixel belonged. The discriminant analysis is evaluated through discussion of the classification accuracies of hazard classes and its classifying variables. A decision tree classifier approach was created based on the 85th percentile

height and skewness biometrics, which describes surface fire hazard for the entire LiDAR data extent. It is assumed that this technique can be combined with supplementary datasets, such as proximity to drainage basins, to shape future hazard mitigation efforts.

1.6 Thesis Organization

This thesis is organized into five chapters. The research context, objectives, and organization are presented in Chapter 1. Chapter 2 reviews research of scientific literature regarding the remote sensing of forests, fire research, and forest management. The study location, data, and methodology are discussed in Chapter 3. In Chapter 4, results from the analysis are presented. Chapter 5 includes the interpretation of those results, conclusions, and a presentation of key findings as well as alternative approaches. Future research recommendations are also provided in the final chapter.

Chapter 2: Literature Review

2.1 Introduction

This review of the literature aims to assess the current research findings applicable to this study's purpose: to assess whether LiDAR data can be used to describe conventional forest stand fire hazard classes. First, the review will look at research on the variables associated with fire hazards, specifically the distribution and makeup of fuel. Second, the review will examine the relationship between derived LiDAR biometrics and forest attributes related to hazard assessment factors as evaluated in the scholarly literature. Third, the review will assess the viability of the LiDAR biometric decision tree, as well as alternative approaches, based on current frameworks for use. The literature review will conclude with a summary of findings.

2.2 Quantifying Fuel Variables Associated With Fire Hazards

In wildfire research, fuel is often the central focus of the research inquiry. Imprecise use of certain terms regarding forest fuels often causes confusion and misunderstanding (Arroyo, Pascual, & Manzanera, 2008). The disconnect in language leads to new information in research literature not being adopted by practitioners. Either the products or conclusions are not easily to integrated without significant processing or change to management practices, or the prescribed methodology is not practicable. Geographically, the challenge then is to build on completed research and further understanding, while maintaining its usability by forest or suppression managers. It is imperative then that the definition for fuel in any literature is explicit. The fire suppression community, research community and resource management community often

use the words fuel and fire interchangeably, while each constitutes a very different meaning for each of them.

Fuel is defined in terms of the physical characteristics of the live and dead biomass that contribute to the spread, intensity and severity of forest fire (Andrews & Queen, 2001; Arroyo et al., 2008; Burgan, Klaver, & Klaver, 1998). The quantification of fuel has strong research initiative; unfortunately, it is an extremely complex concept. In academic literature, fuel is considered to be the dry weight of the carbon-based biomass, but in reality this is an example of fuel loading. Fuel, consequently, has multiple definitions relating to the different disciplines, and various ways for how it is calculated that need to be discussed.

Although a forest fire can burn below ground, the primary concern is the biomass above the earth's surface. How this is quantified is the focus of many studies, many of which further understanding about the relationships that exist between trees, their densities, and respective biomass. In many ways, the most important component is the above ground biomass (AGB); it is the estimation of standing trees within a given area. A variety of empirical estimation equations produce dry weight values for separated biomass components such as the tree canopy. The estimation of AGB has been successfully completed using remote sensing techniques, especially LiDAR, in a number of studies (Kim, Yang, Cohen, et al., 2009; Naesset & Gobakken, 2008; Riaño, Meier, Allgöwer, et al., 2003).

The methodology associated with this type of estimation generally revolves around two basic approaches. First, tree segmentation methods, such as the application of allometric equations from Canada's national biomass equation database to individual tree

species identified in a LiDAR dataset, produce biomass estimates for whole or parts of individual trees (Ung, Bernier, & Guo, 2008). The equations produce estimates of dry biomass weight for separate or combined components of AGB which are generally accepted as producing meaningful biomass assessments across large regions like Canada. Second, the plot-based approach involves using field-measured biomass regressed against derived statistics from plot-level LiDAR data (Kim et al., 2009). Correlations between LiDAR data and aboveground tree biomass measured in the field were significant with the 80th percentile return explaining 74% of the variability in measurements of forest structure and fuel loads in the Pinelands of New Jersey (Skowronski et al., 2007). A cluster analysis workflow can be used to obtain crown bulk density using foliage biomass equations and estimates of crown volume which is estimated from the crown area times the crown height (Riaño et al., 2003). In particular, LiDAR techniques have been used to estimate understory height, crown bulk density, crown fuel mass as well as the presence of ladder fuels (Skowronski et al., 2007; Andersen, McGaughey, & Reutebuch, 2005; Riaño et al., 2003). In the Capitol State Forest in western Washington State, for example, regression analysis techniques were used to develop predictive models relating a variety of LiDAR based metrics to the canopy fuel parameters. These were estimated from inventory data collected at plots established within stands of varying conditions (Andersen et al., 2005). Strong relationships were found for canopy fuel weight ($r^2=0.86$), crown bulk density ($r^2=0.84$), canopy base height ($r^2=0.77$) and canopy height ($r^2=0.98$) (Andersen et al., 2005).

Understory biomass estimation techniques and the characterization of forest understory is a relatively novel research direction. This is a critical component of fuel

distribution assessments. Normally, fires start in the understory and the presence of ladder fuels govern its spread to a more dangerous crown fire. The accumulation of understory fuel, typically thinned by natural fire cycles that have been suppressed, is a great concern to fire managers, especially in the wildland-urban interface. Some techniques and empirical equations for quantifying this fuel do exist. Sah, Ross, Koptur, and Snyder (2004) have built on prior empirical allometric equations and provide equations for determining dry fuel weight of specific and mixed species, tree or shrub like in nature in the understory of Florida Keys pine forests. Stem basal diameter, crown area, and/or height are typically used as independent variables in shrub biomass equations (Sah et al., 2004). However, understory biomass estimation using LiDAR data is relatively new and just gaining focus.

The use of upward-sensing LiDAR, or ground-based laser profiling, has also been used to calibrate airborne scanning data thus improving conventional canopy bulk density metrics (Skowronski et al., 2011). Both upward and downward scanning LiDAR predict the maximum two-dimensional canopy bulk density (CBD_{max}) and canopy fuel weight well (e.g., Anderson et al., 2005). However, a great amount of detail is omitted by expressing canopy fuel as a single value for each cell and may not capture the effects of prescribed fires or other fuel management activities that affect sub-canopy and understory fuel loading (Skowronski et al., 2011). Through the use of binned canopy bulk density (CBD_{bin}), the vertical nature of fuel loading is preserved for use by fire managers and the development in next-generation fire behaviour modeling (Skowronski et al., 2011).

The above ground biomass estimation of standing trees, the characterization of the understory and the investigation of the ecological relationships that exist between the

overstory and the understory are all important elements of a comprehensive fuel distribution analysis. Using LiDAR data to indirectly quantify total biomass dry weight is an exciting development of current research of AGB estimation. Less attention has focused on the understory; however, new methods of understory fuel distribution analysis may enable researchers to quantify biomass estimations for this component of a forest stand. Ultimately, photo-interpretive and conventional field inventories are costly and prone to large error over large landscapes. LiDAR samples large landscapes consistently, statistically, cost-effectively and with high resolution. For these reasons it is the best methodology in this type of scientific inquiry. The continued research of LiDAR in this context will provide a better, more comprehensive fuel distribution assessment that will help fuel management specialists with mitigation planning.

2.3 Classifying Fire Hazard Fuel Types

It is currently very difficult to describe all physical characteristics for all fuels in an area. Using the prevalent techniques the description of those properties relevant to fire danger estimation and fire propagation studies is based on classification schemes, which summarize large groups of vegetation characteristics (Arroyo et al., 2008). The groups are often defined as fuel types (Arroyo et al., 2008; Pyne, Andrews, & Laven, 1996). Merrill and Alexander (1987) define a fuel type as an identifiable association of fuel elements of distinctive species, form, size arrangement and continuity that will exhibit characteristic fire behaviour under defined burning conditions. Fuel types are generated in different ways and mean different things, country by country.

To accurately model fire behaviour potential, as calculated by the Canadian Forest Behaviour (FBP) system, a fuel type map is required as an input (Nadeau & Englefield,

2006). There are three main ways fuel types are derived by suppression personnel for consumption in the FBP system: Local knowledge, a posteriori, of fuel types in a region by veteran suppression managers, fuel maps generated from a combination of satellite or airborne imagery (Nadeau & Englefield, 2006), and on-site by field personnel. Fuel maps derived from hyperspectral and LiDAR sources have the potential to significantly improve upon the coarse resolution and limited scope of input data of conventional fuel maps. These maps have been generally not suitable for operational fire hazard mitigation and Canadian Forest Service (1992) acknowledges that FBP system fuel types descriptions do not rigorously or quantitatively follow forest inventory patterns and that knowledgeable fire managers will develop methods to classify their land base and vegetation data for their respective fire planning. LiDAR can classify vegetation data accurately, consistently and cost-effectively. A main concern with the development of remote sensing methodologies for creating fuel maps, however, is the integration of a digital system with a largely analogue one. In the FBP system fuel maps are interpreted qualitatively, having elements of stand structure and composition, surface and ladder fuels, and forest and floor cover (Nadeau & Englefield, 2006). Therefore, any quantitative data based on this digital system will have to be classified into more qualitative-like data type. Thus, fuel types are generalized based on the interaction of a multitude of factors, for example the species type distribution.

Species types, additionally, may not be sufficient when describing a local ecosystem, as characteristics of surface fuels and stand age can become more dominant factors in determining the most representative fire behaviour class. The use of a more representative fuel type, constituting perhaps of a differing species than observed on the

ground, is sometimes used to compensate when fire behaviour, under certain environmental conditions for instance, will act more like a differing fuel type. This is a more heuristic approach used by experienced suppression professionals, based on their experience and the methods at their disposal to more precisely model fire propagation behaviour. This methodology shares a lot in common with the concept mapping of fire hazard as it weighs a multitude of variables, when determining how best to describe the risk.

The use of LiDAR sensors is of significant benefit to fuel type classifications because it provides a critical metric needed in the determination of a fuel class, the vertical distribution of the vegetation being classified. Additionally, the use of LiDAR allows researchers an accurate, repeatable, unbiased and efficient estimation of the fuel characteristics over large area of forests (Andersen et al., 2005). To characterize fuels for wildfire risk one must consider factors such as crown bulk density, crown base height, canopy height, percent of canopy cover, surface area-to-volume ratio, vertical and horizontal continuity, dead and live fuel load, and size classes of fuel elements (Riano et al., 2003). LiDAR data are well suited for many of these variables and combined with optical data sources are effective means of generating fuel types for at-risk areas.

A recent study in the Sam Houston National Forest in East Texas integrated LiDAR data for surface fuel type mapping with the intention of using the map as a direct import into the United States' FARSITE fire modeling software, which is a spatially explicit, two-dimensional deterministic fire growth simulation model (Mutlu, Popescu, & Zhao, 2008). Similar to Canada's FBP, FARSITE is a tool used to assess the behaviour of an active fire based on the environmental conditions and fuel types geographically and

provide fire behaviour information to fire managers responsible for suppression. An assessment of FARSITE and its application is apropos for the Mutlu, Popescu, & Zhao (2008) study, due to the fact that it, like the approach in the present study, attempts to merge LiDAR and other data as a means to present a better model for understanding hazard risk. Mutlu, Popescu, and Zhao (2008) combined the LiDAR data with 2.5m spatial resolution QuickBird imagery to create a ten band stack consisting of 4 LiDAR height bins, one band of a canopy cover model, one band of canopy cover variance and the 4 bands of the QuickBird image. The LiDAR height bins were generated by counting the number of LiDAR points within each volume unit and normalizing by the total number of points to negate the effects of variable point density (Mutlu, Popescu, & Zhao, 2008). While bins were created to account for the entire LiDAR point cloud, the use of bins up to 2 meters were only used in the classification process. The authors note that with a leaf-off LiDAR data set, the first four height bins are expected to depict differences in canopy structure and penetration. For instance, fuel models such as grasses, chaparral, and brush will intercept most of the LiDAR hits (>80%) when the canopy is open, or not present at all, and therefore directly measure and characterize these fuel types based on the differences in those height bins (Mutlu, Popescu, & Zhao, 2008). For fuel models such as brush under canopy, closed timber litter, and hardwood litter, the authors do not anticipate the LiDAR hits to characterize the tightly compacted dead foliage occurring under a dense canopy (Mutlu, Popescu, & Zhao, 2008). This is because in areas with a dense overstory, the vertical distribution of a LiDAR point cloud is characterized by the vertical profile of the first four height bins. That is to say, the canopy

condition actually influences the development of a specific fuel model (Mutlu, Popescu, & Zhao, 2008).

The distinction between the brush under canopy fuel model and the timber and hardwood litter fuel models are determined by the profile of the height bins and the canopy condition itself. While this technique worked well for the Sam Houston National Forest, the distinctions of fuel models in other areas may not be as readily distinguishable if LiDAR height bins are 2 meters or more between the different fuel classes. In areas such as coastal conifer forests, more mature forests mean that the canopy dominates the LiDAR signature limiting some direct understory measurements. The challenge is to create a methodology that can be exported to different areas but based on the structural relationship of hazard one can observe in any forest. Statistical pattern recognition methods can exploit and categorize pixels in an image data set to the fuel class it most closely resembles, as calibrated by the fuel models training data set (Mutlu, Popescu, Stripling, & Spencer, 2008). However, training datasets that are not relatively homogeneous may not be available and then not of value when utilizing statistical classifications (Jensen, 2007).

An alternative classification technique that is easily modifiable for distinct regions is the Decision Support Tree (DST) which has been used to create fuel models (Falkowski, Gessler, Morgan, et al., 2005). With respect to the methodology outlined by Mutlu, Popescu, Stripling, and Spencer (2008), a DST may more appropriately distinguish between closed canopy fuel models where separability is weak, or where extra conditions are placed on fuel model assignment decisions. Ultimately, a DST may also allow for better organization and classification of disparate data types such as

LiDAR with optical data. Mutlu, Popescu, Stripling, and Spencer (2008) used surface fuel heights in their height bins, though no data beyond two meters was used while data over 10 meters was available. A supervised classification using the Mahalanobis distance algorithm produced a map with an accuracy of 90% of seven different fuels. The height bin approach was utilized from the method of creating multiband data from scanning data presented by Zhao et al. (2009). This was important because many previously developed models are scale-dependent that need to be fitted and then applied at the same scale or pixel size. Additionally, their methodology included a minimum noise fraction transformation (MNF), retaining six of the ten component bands for use in classification. The MNF transform is a Principal Component Analysis (PCA) data reduction method where one PCA occurs first on the noise in the data and is then incorporated into the second PCA (Green, Berman, Switzer, & Craig, 1998). This was found to be much more effective than just using a PCA transformation (90% vs. 61%), but only 3% more accurate when compared to a classification on the QuickBird-LiDAR stack untransformed (87%). Finally, Mutlu, Popescu, Stripling, and Spencer (2008) evaluated fuel models on a per-plot basis, then applied a majority filter with a 7 x 7 window. The end result is a generalized surface fuel model that was derived through the automatic extraction of forest information that can be used as an input into FARSITE (Finney, 1998).

If the intent is to classify a large area and produce a generalized fuel-type maps this kind of classification may be appropriate. However, fusion of LiDAR data with optical data, with the intent to input the layers into FARSITE, may require a more appropriate integration to ensure good performance. To evaluate the possibility of a more

effective data fusion, a deeper understanding of how FARSITE works and utilizes input layers is required. Prior to FARSITE, distinct fire models were required for surface, crown, spotting, and point source fire modeling (Finney, 1998). The integration of these existing fire behaviour models relied on an assumed sequence of fire activity. Finney (1998) describes this sequence: first, a fire may spread as a surface fire. It burns in the grass, shrubs, or downed woody fuels in contact with the ground surface. If the environmental conditions permit, the fire will accelerate toward some new equilibrium spread condition. Given sufficient fuels, weather, and topography, the fire may make the transition to burning in the aerial fuels of tree crowns (crown fire). If crown fuels are ignited trees are assumed to torch and can loft embers initiating spotting, which is when the fire spreads from crown to crown.

Ultimately, the performance of FARSITE depends on the relationship between the distinct types of fire and their respective propagation algorithms, the spatial resolution of their respective fuels represented in input data, and the quality and resolution of ancillary data. FARSITE uses inputs which LiDAR remote sensing is particularly well suited for providing. Canopy cover, digital elevation models (DEM), slope, aspect, crown stand height, crown base height and crown bulk density are easily obtained from laser backscatter (Andersen et al., 2008; Popescu, Wynne, & Scivani, 2004; Riano et al., 2003). FARSITE requires inputs of fuel types and LiDAR, which is fused with optical data sources, provides the necessary components to develop fuel model maps (Koetz et al., 2008; Mutlu, Popescu, Stripling, & Spencer, 2008; Varga et al., 2008). Additionally, FARSITE requires input layers regarding the environment such as temperature, humidity, precipitation, and wind-speed and direction (Finney, 1998).

As mentioned, the chosen fuel model provides FARSITE with the physical description of the surface fuels complex that is used to determine surface fire behaviour. A total of 13 surface fuel models were initially identified in the United States, each varying in amount, size and arrangement of the fuel. When FARSITE was initially conceived, fuel model generation was generally a highly estimated method using a multitude of spatial data layers and satellite sensors that cannot exactly describe surface fuels (Rollins, Keane, & Parsons, 2004). Additionally, the quality of fuel models varied across the country. Through these recent research efforts, the use of LiDAR data has significantly improved the ability to classify fuel types at a landscape level, accurately and efficiently. This has significantly improved fuel type maps for use in fire behaviour modeling where implemented.

2.4 Mapping Fire Hazard

Prior analysis of fuel distribution in a forest stand is a critical to determining fire and risk in the wildland-urban interfaces. Research combining remote sensing techniques and ecological relationships has successfully mapped fuel distribution and fire regimes for many years, as detailed further in Appendix 1. The final concept of the remote sensing of fuel is its use defining a hazardous fire condition. The identification and mapping of hazardous conditions for use by those responsible in controlling fire risk in wildland-urban interfaces is an important research direction.

To characterize fuels for fire risk, one must consider factors such as crown bulk density, crown base height, canopy height, percent of canopy cover, surface area-to-volume ratio, vertical and horizontal continuity, dead and live fuel load and size classes of fuel elements (Riaño et al., 2003). Figure 1, below, provides an example of the use a

cluster analysis for the generation of tree height and crown base height. Additionally surface canopy height indicates the nature of the vegetation vertically in an intensively managed, homogeneous, Scots Pine with little understory due to thinning (Riaño et al., 2003). This information is traditionally categorized as differing fuel types by classifying LiDAR returns into ground and vegetation groups based on 85th percentile height. The use of this classification and their associated metrics can assist when mapping hazardous fuel distribution, particularly in fire behaviour models. Coincidentally, this type of metric generation combined with other factors relating to hazard, can effectively map fire hazard by using LiDAR remote sensing.

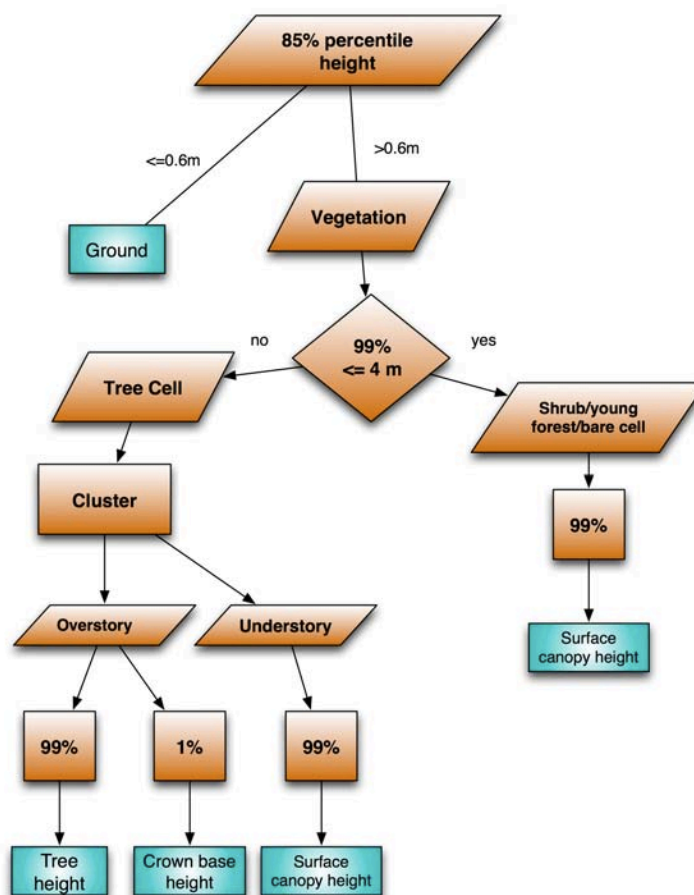


Figure 1 - Flowchart for the generation of above-ground metrics (Adapted from Riaño et al., 2003)

The differences in ecological conditions with high fire risk around the world are too numerous to list. However, the most consistent variable among all regions is that hot, dry and windy environmental conditions exist (Mutlu, Popescu, Stripling, & Spencer, 2008). Hazardous conditions at the Sooke Watershed in British Columbia, Canada, will be used as an example. Fine fuels, generally known as non-photosynthetic (or dead) woody material, on the forest floor are the variable most directly controlling the initiation and spread of fire (Pyne et al., 1996). These fuels accumulate from blow-down, trees killed by wind storms, or from disease and other natural causes which displaces the fuel matter to the surface (J. Ussery, personal communication, May 31, 2011). They also consist of dead standing shrubs and young trees, forest litter and extremely flammable dry needles from coniferous trees (J. Ussery, personal communication, May 31, 2011). These high risk surface fuels need not be dead; in the Sooke Watershed *Cytisus Scoparius*, or Scotch Broom, is a particular hazardous noxious weed with rigid, woody stems growing 1 to 3 metres in height. In fact, Scotch Broom represents a major manual thinning effort annually for the CRD as a way to reduce fire risk (J. Ussery, personal communication, May 31, 2011).

The presence, spatial arrangement, and density of fine fuels constitutes the biggest risk or hazardous fuel condition for a forest stand apart from fuel moisture content (Mutlu, Popescu, Stripling, & Spencer, 2008). Also, the height of the fine fuels is considered an additional scalar to that risk (Varga & Asner, 2008). As with the height above ground of the surface fuels, the canopy base height in relation to those surface fuels represents a further risk multiplier as it created a ladder for the transition from surface fire, to a crown fire (Skowronski et al., 2007). The total fire risk can be thought of

as the product of the surface and canopy risk combined. A high-risk surface fuel condition combined with ladder fuels to high-risk crown fuel condition is considered to be the highest total fire hazard at the CRD. This type of fuel condition, with the topographic conditions of adjacent rising terrain known to accelerate the rate of spread, is of greatest concern to those responsible for preventing fire (Pyne, 1996; Ussery, personal communication, May 31, 2011). Fire risk can be more effectively mitigated if managers know the spatial makeup of the fuel conditions. Optimizing thinning practices with a landscape level assessment of fire hazard will allow managers, like those at the CRD, to more efficiently reduce fire risk (J. Ussery, personal communication, May 31, 2011).

A very notable study by Varga and Asner (2008) maps fire hazard, based on the authors' heuristic definition, specifically applicable to Hawaii. The authors utilized 1.8 metre hyperspectral data focusing on reflectance in the 2078nm to 2278nm wavelength interval to differentiate between surface types (Varga & Asner, 2008). The Probabilistic Spectral Mixture Model (AutoMCU) was created as a clustering technique to map the fuel loads in Hawaiian ecosystems (Varga & Asner, 2008) by determining sub-pixel cover fractions of photosynthetic, non-photosynthetic, and bare substrate and shade (Asner et al., 2000; Asner et al., 2002).

The equation utilized by Varga and Asner (2008) is:

[1]

$$\begin{aligned} \rho(\lambda)_{\text{pixel}} &= \sum [C_e \times \rho(\lambda)_e] + \varepsilon \\ &= [C_{\text{PV}} \times \rho(\lambda)_{\text{PV}} + C_{\text{NPV}} \times \rho(\lambda)_{\text{NPV}} \\ &\quad + C_{\text{B/S}} \times \rho(\lambda)_{\text{B/S}}] + \varepsilon = 1 \\ \varepsilon(\lambda) &= \sum_{e=1}^n [C_e] \end{aligned}$$

where C_e is the fractional land cover of photosynthetic vegetation (PV), non-photosynthetic vegetation (NPV), or bare substrate (B/S). $\rho(\lambda)_e$ is the reflectance of each land cover at wavelength λ and ε is the root mean square error (RMSE) for each pixel. The equation equals to 1.

The purpose of determining fractional land cover is that it can be combined with available LiDAR data to create an index to quantify the three-dimensional fire fuel conditions in the landscape (Varga & Asner, 2008). Varga and Asner (2008) had taken the vegetation height from LiDAR and the lateral extent of live and dead materials from hyperspectral data to create a unit-less metric describing the percentage of volume of dead fuel material in a three-dimensional matrix. The equation is expressed as:

[2]

$$FFI_{(x,y)} = \frac{[C_{NPV}(x,y) \times IFOV \times h_{NPV}]}{[IFOV \times h_{x,y}]}$$

where FFI is the fire-fuel-filled canopy volume at pixel x,y ; $C_{NPV}(x,y)$ is the fractional land cover value of the non-photosynthetic vegetation; IFOV (instantaneous field of view) is the pixel resolution; h_{NPV} is the mean height in metres of non-photosynthetic in the landscape; and $h_{x,y}$ is the height in metres of the uppermost canopy at pixel x,y and acts as a scalar. An example of their map is shown in Figure 2, below.

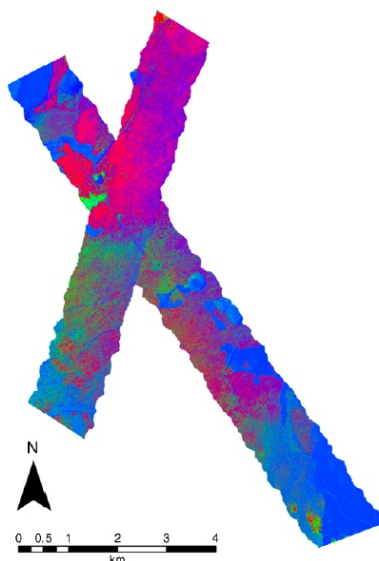


Figure 2 - Example of Fuel Fire Index: FFI map of 1.8m AVIRIS derived imagery. PV, NPV, and BS/S are represented by red, green and blue respectively (Varga et al., 2008)

The non-photosynthetic (dead or senescent) vegetation is the variable of interest as it is likely to be the initial site of fire, whereas photosynthetic vegetation is more likely to serve as a firebreak (Bond et al., 1996; Varga & Asner, 2008). Varga and Asner (2008), through the use of FFI, have described the spatial pattern of hazard in the landscape. FFI represents a geographically applicable hazard metric based relationship between non-photosynthetic vegetation, its height, and the height of the canopy. This simple index effectively translates a quantitative metric of NPV and vertical height, using understood spatial relationships of fuel spread, to create a valuable qualitative metric of hazard. That being said, the fire models or fuel classifications are understood to be specific geographically and applicable for the purpose for which they are developed. The simplicity of the FFI works in the Hawaii grasslands but would quickly saturate a map with high fire hazard if indiscriminantly exported for use in the Sooke Watershed. Instead, the conceptual framework of creating a coastal conifer forest specific, structural-

based, fire hazard metric is the critical building block in fire hazard mapping for this research.

2.5 Viability of LiDAR Biometrics in Relation to Forest Attributes

While a description of fuel structure is required for all models of fire behaviour, the best-practice means for which such a structure or framework for management is created and used is often debated both in the field and in the research literature (Alexander & Cruz, 2013; Costanza, Weiss, & Moody, 2013; Gill & Stephens, 2009; Krivtsov, Vigy, Legg, et al., 2009). Most importantly, the literature points to a gap between the creation of models of fire behaviour and the accuracy of such models linked to the ability of fire hazard teams to use these models effectively (Ager, Vaillant, & Finney, 2011; Alexander & Cruz, 2013; Costanza et al., 2013; Gill & Stephens, 2009), the application and limitations of LiDAR and similar data sources as a means of enhancing traditional models (Fernandes, 2009; Hollingsworth, Kurth, Parresol, et al., 2012; Skowronski, Clark, Duveneck, & Hom, 2011), and, given these overlapping challenges, the overall lack of viability in current modelling and framework-creation techniques in the field (Beguería, 2006; Keane, Drury, Karau et al., 2010; Krivtsov et al., 2009; Thompson & Calkin, 2011).

The flawed practical application of research in the field has led to a disconnect between the ways in which fire hazard models have been intended to be used by academic model developers and the ways in which they are used by those tasked with prevention of forest fires, according to the literature (Ager et al., 2011; Alexander & Cruz, 2013; Costanza et al., 2013; Gill & Stephens, 2009). There is evidence that many academic models do not take into account the challenges faced by field workers, but also

that field workers may try and use academic models or data in a different way than intended by researchers (Costanza et al., 2013), which can lead to increased risk over time (McKenzie & Kennedy, 2011). In addition, because of this gap between model development and practice, most existing frameworks, namely, the combination of models with data analysis management and practical assessment, have undergone limited testing against observations garnered from planned or accidental wildland fires (Alexander & Cruz, 2013). This has resulted in “at least a decade of model misapplication in fire and fuel management simulation modelling stemming from a lack of model evaluation” in western North America (Alexander & Cruz, 2013, p. 65).

The reason that this knowledge-practice gap exists stems, in part, from either an over-reliance on data modelling techniques, or simply a lack of understanding of the limitations of data modelling techniques in the field (Alexander & Cruz, 2013; McKenzie & Kennedy, 2011). One of the ways in which modelling can fail, for example, is when the scale of the landscape disturbances changes, which shifts the ways in which fuel operates and interacts with other landscape factors. As noted in the literature, the physical mechanisms of heat transfer remain the same across different forest and fuel scales, and fire spread does depend on local connectivity of fuels, but at the same time, the estimates of connectivity across landscapes are sensitive to spatial resolution and will change the ways in which a fire will spread (McKenzie & Kennedy, 2011). Power laws for scaling may or may not have an impact on the efficacy of a model to predict burn, depending on landscape factors, such as fire frequency (McKenzie & Kennedy, 2011) or pine beetle infestation (Alexander & Cruz, 2013). These are factors that either change quickly, or are not regularly measured, or are not entered into data models. In other words, there are only

so many variables that can be algorithmically linked to create a valid model, especially in certain forest contexts that may have mitigating factors that can shift in a very short time period, such as those in western Canada (Alexander & Cruz, 2013). This results in data-based models that have a limited use in assessing the flammability of natural forests and the effectiveness of fuel treatments in reducing fire potential, even when the data used is altogether accurate (Alexander & Cruz, 2013).

The literature demonstrates that LiDAR data assessment, and assessment of similar data sources, is particularly vulnerable to misuse (Fernandes, 2009; Hollingsworth, Kurth, Parresol, et al., 2012; Skowronski et al., 2011). In an assessment of LiDAR efficacy in measuring canopy fuel distribution, Skowronski et al. (2011) found that while LiDAR could be effective at some levels in the canopy, differences between sensors and their positions in the forest affected the ability to detect canopy fuels estimated from biometric measurements. While the technology was able to aptly provide large-spatial scale estimates, the Skowronski et al. (2011) measured gaps in LiDAR sensors' ability to detect changes in canopy fuel parameters and height profiles, even when Gaussian distributions to LiDAR were applied. In essence, this study found that data can become overly smoothed and therefore valuable information about the vertical distribution of the canopy fuel, as well as minimum canopy bulk density, may be lost because the data is too computationally intensive. Skowronski et al. (2011) discovered that the interpretation of LiDAR data can therefore be problematic not only because of computational intensity, but also because of the difficulty in displaying these data visually without resorting to some type of classification scheme that can skew the ways in which the data is read and assessed.

Challenges linked to computational intensity and classification of data are also found in non-LiDAR data sources and evaluation methods (Hollingsworth et al., 2012). In a review of different models for remote sensing and evaluation, Krivtsov et al. (2009) note that in order for data to be used accurately, and therefore models to be viable, spatially explicit representations are necessary. Mapping fuel models over large areas would only be possible by integrating georeferenced databases such as remote sensing and geographic information systems data, but only within the parameters of addressing the heterogeneity and texture of individual strata in a given area. The findings of Krivtsov et al. (2009) reflect those of Thompson and Calkin (2011), who reviewed data factors taken into account for creating viable models. As they write, the evidence from the literature demonstrates that a significant tradeoff exists between computational requirements for wildfire simulation models and the algorithms commonly used by field teams to apply these models with remote sensing data. To this end, the creation of models that place more restraints on optimization are able to model exposure analysis and effects analysis more robustly over time (Thompson & Calkin, 2011).

What is clear from the literature is that there is no single data source or data evaluation method that is currently able to provide the needed level of utility or accuracy for a diverse and viable fire management program over the long term (Hollingsworth et al., 2012). This may be due to a limited ability for data integration among fire behaviour models, as well as constrained linkages to geographic information systems, management-level data and factor definitions, and the ability of a team to compute and use data easily (Ager et al., 2011). Four key factors have been identified as being especially problematic: a) computational resources available to fire management organizations; b) high quality,

spatially consistent, management-oriented spatial data layers at the appropriate scale and resolution; c) lack of error and uncertainty estimates for the spatial data layers; and d) improper spatial analysis techniques (Keane et al., 2010).

To address this lack of viability in current frameworks for data use and application in the field, the literature suggests that multiple data treatment alternatives need to be created (Ager et al., 2011), but that for these to be viable these alternatives need to be updated with relative frequency (Thompson & Calkin, 2011), and that the alternatives need to also align with multiple public interest objectives (Ager et al., 2011; Hollingsworth et al., 2012). In other words, there is no indication in the literature that a single data source, drawn from LiDAR or any remote sensing process, can be assessed through any data analysis system and used in a viable manner. Overlapping data sources, tools, and interpretation processes that are repeated over time are the only means by which to obtain a clear picture of forest attributes and avoid the pitfalls of knowledge-practice gaps in order to create viable probabilistic and predictive models.

2.6 Summary of Literature Findings

This literature review evaluated research on the variables associated with fire hazards, specifically the distribution and makeup of fuel, the relationship between derived LiDAR biometrics and forest attributes related to hazard assessment factors as evaluated in the scholarly literature, and the viability of the LiDAR biometric decision tree, as well as alternative approaches, based on current frameworks for use. Findings demonstrated that properties relevant to fire danger estimation and fire propagation studies are based on classification schemes of fuel types and vegetation characteristics, but that fuel types are generated in different ways and mean different things depending on

the specific location of a fire management team which can potentially become arbitrary when describing a local ecosystem. Nonetheless, analysis of fuel distribution in a forest stand is a critical to determining fire and risk in wildland-urban interfaces. Although research combining remote sensing techniques and ecological relationships has successfully mapped fuel distribution and fire regimes for many years, there are significant limitations in these processes both from a data management and a field application point of view. While a description of fuel structure is required for all models of fire behaviour, the best-practice means for which such a structure or framework for management is created and used is often debated both in the field and in the research literature. What is clear from the literature is that there is no single data source or data evaluation method that is currently able to provide the needed utility or accuracy for a diverse and viable fire management program; data systems and types must be combined for an accurate reading of fire hazard risk and management.

Chapter 3: Method

3.0 Introduction

Research questions 3 asks whether classification using LiDAR biometrics with the current frameworks and data available from the CRD is viable. Therefore, the methodology for this study utilized the extensive field assessment already performed by Blackwell (Unpublished report, 2006), employing 433 surveyed plots in the Sooke Watershed. The conceptual framework and definitions of fire hazard developed by the CRD, as well as the fundamental forest relationships presented in the literature review were employed. Stratified random sampled points from the LiDAR dataset have been decorrelated using principal component analysis. In turn, discriminant analysis has determined to which fire hazard class the representative pixel belonged. The discriminant analysis was evaluated through discussion of the classification accuracies of hazard classes and its classifying variables. A decision tree classifier approach was created based on the 85th percentile height and skewness biometrics, which describes surface fire hazard for the entire LiDAR data extend. It is assumed that this technique can be combined with supplementary datasets, such as proximity to drainage basins, to shape future hazard mitigation efforts. The rationale behind the chosen methodology, which is primarily quantitative, was the need to effectively map fire hazard in western coniferous forest stands.

In this chapter, the methodology for this study is presented in detail. The location for the study is presented, and the sampling process as well as a description of the data collected is detailed. The method by which the data is manipulated to answer the research questions is outlined. A summary of the methodology is presented at the end of the

chapter. A flowchart depicting the processing of the inputs and outputs of the methodology is included in Appendix A.

3.1 Study Area

The study area, known as the Sooke Watershed, is located in southern Vancouver Island. It is comprised of typical coastal vegetation with 80-90% Douglas Fir (*Pseudotsuga menziesii*) and Western Hemlock (*Tsuga heterophylla*), several stands of Red Cedar (*Thuja plicata*) and very sparse Lodgepole Pine (*Pinus contorta*). Typical landcover is a Douglas Fir and Salal (*Gaultheria Shallon*). The presence of Scotch Broom (*Cytisus Scoparious*) is of particular concern for its high flammability, making it the target of intensive manual thinning efforts. Overall the forest stands are generally young and intermediate age plantations with extreme tree heights suggesting advanced aged regenerated forests. There are human structures such as power lines and roads, and many natural features such as rivers, lakes, and exposed earth.

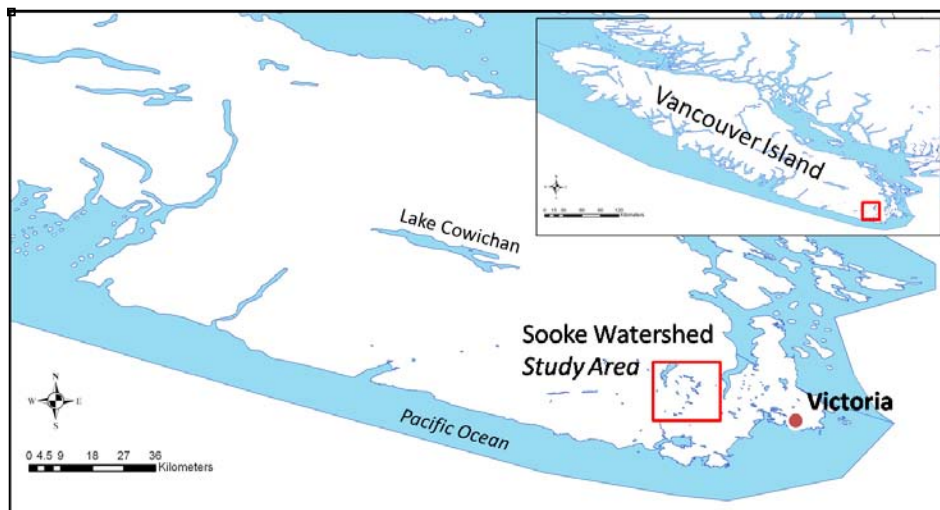


Figure 3 - Study Area located on Southern Vancouver Island

3.2 Sampling Design

A stratified random sampling of 2000 20m² pixels (also referred to as points) was used from the total number of 66695 pixels that overlapped with the spatial extent of the field-derived, forest attributes dataset. 2000 samples were selected because they could be split into training and validation subsets, be sufficient for the use of inferential statistics as indicated through the use of a chi-square test. Additionally it was used because it represents a sufficient number of cases for testing normality using a Kolmogorov-Smirnov test. Stratified random sampling differs from systematic random sampling in that it has the additional feature that a population (or area) is divided into strata with known proportions of the whole (McCune & Grace, 2002). Sample units are selected at random from within the strata. A total of 14 points were discarded from 2000 random samples drawn because they had no LiDAR data associated with them. This occurs when using geographic information systems and overlaying square pixels of raster layers with vector delineated shape files. The 1986 remaining points were then used to compile a merged dataset that consisted of the value for each LiDAR biometric and the field-derived forest attributes for that representative 20m² area. The merged dataset was then equally divided using a random subset tool in ArcGis© to represent the training and validation datasets. All analyses and visualization was based on the training dataset. The method of validation utilized is known as bootstrapping and is an effective method for validating remote sensing classifications when physical validation is not possible. Discriminant scores from the training analyses were consistent with those from the validation dataset; this suggests there is no statistically significant difference in the classification results between the two subsets. Due to the nature of the classification results and the agreement between in the training and validation datasets with regards to

the classification accuracies, subsequent testing to determine if the discriminant analysis was performing well because the discriminant function was borne from the training points was not necessary.

3.3 Data Collection

3.3.1 LiDAR data

The airborne acquisition occurred on July 24, 2004. The technical specifications of the sensor are outlined in Table 1. The LiDAR data is in format 1, an ASCII text file based on the ASPRS LiDAR data exchange format (LAS) (<http://www.lasformat.org>). The fields detail the spatial position of a return in x, y, z, the intensity, return number, number of returns (given pulse), scan direction flag, edge of flight line, classification, scan angle, and GPS time.

Table 1 - LiDAR data technical specifications

Parameter	Specification
Laser	Nd- YAG
Wavelength	1089 nm
Beam divergence	0.45 milliradians
Pulse duration (length)	8 nanoseconds
Pulse rate	40 kHz
Scan frequency	30 Hz
Scan angle	24

(Langford, 2006)

The size of the entire dataset is 3.06 gigabytes and represents over 130 hectares, the maximum amount of data available. The raw data is divided into 28 files, each representing 1000m² sub-area. The largest of these files is 250 megabytes in size and that file contains 3.7 million LiDAR points, which gives a scope of the quantity of points subsequent processing is based on. Terra Remote Sensing acquired the data and initial

processing performed by Terra and the Hyperspectral and LiDAR research group at UVic. The flightlines are outlined in the appendix.

3.3.2 LiDAR Biometrics

A 20m grid as a spatial template to extract a number of biometrics from the normalized LiDAR point cloud composed of heights above the classified ground surface was applied. A 2m-fixed threshold was used to eliminate near-ground discrete LiDAR returns because they are often associated with uncertainty and noise in the dataset (inability to reconcile terrain returns from vegetation or other surface features) (Niemann & Frazer, 2009; Nilsson, 1996). Derived biometrics in combination with the terrain-normalized height data constitute the bands used in the analysis and are outlined in Table 2, below. The mean band is the average height of LiDAR laser returns for a given pixel and is useful for classifying areas of vegetation from the ground as well as making inferences on forest stand maturity. The mode band can be used to highlight the most frequent height of vegetation and is particularly useful when there exists high variability in height of vegetation to which mean may not be suitable for making inferences on stand age. The rugosity band characterizes the amount of variation in LiDAR height returns in the horizontal dimension for a given pixel and can indicate the smoothness of either a forest canopy or of ground vegetation depending on the mean height of LiDAR returns. The skewness band indicates the vertical distribution of LiDAR height returns and can be used to infer the vertical continuity of fuels for a forest stand. The kurtosis band is used to describe how extreme the displacement is in the values over the mean and the rest of the distribution of the LiDAR height returns. The 85th percentile band is used to describe the co-dominant tree height for the pixel. Gap fractions are separated into 4 percentiles and

the band describes the density of vegetation in the vertical dimension by characterizing the amount of openness, or amount of LiDAR returns that were able to hit the forest floor.

Table 2 - LiDAR data bands used in analysis and relevant algorithms

Band	Description	Algorithm
Mean	The mean height in metres of trees in a 20m by 20m pixel	$\bar{x} = \frac{\sum x}{N}$
Mode	The mode of the height measure for a 20m pixel	Mode(s)
Rugosity	The standard deviation of height	$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$ $\bar{x} = \frac{\sum x}{N}$
Skewness	Skewness describes the vertical distribution of discrete returns in a 20m pixel	$\tau_3 = \lambda_3 / \lambda_2$ where: $\lambda_3 = \frac{1}{3} E[X_{3,3} - 2X_{2,3} + X_{1,3}]$ $\lambda_2 = \frac{1}{2} E[X_{2,2} - X_{1,2}]$
Kurtosis	Relates to the 'peakedness' of the distribution curve	$\tau_4 = \lambda_4 / \lambda_2$ where: $\lambda_4 = \frac{1}{4} E[X_{4,4} - 3X_{3,4} + 3X_{2,4} - X_{1,4}]$ $\lambda_2 = \frac{1}{2} E[X_{2,2} - X_{1,2}]$
85 th Percentile	The 85 th percentile is the co-dominant canopy height in a 20m pixel	$X = \mu + Z\sigma$ where μ is the mean, σ is the standard deviation of the variable X, and Z is the value from the standard normal distribution for the desired percentile
Gap Fraction (0,25,50,75)	Gap fraction measures the % of gaps in a 20m forest stand	Gap Fraction X = [first returns / all returns] * 100
Variance	The variance around the mean height of the vertical canopy for a 20m pixel	$Var(X) = E[(X - \mu)^2]$

The method of implementing *L*-moments, as developed by Hosking (1990), were used to provide a measure for location, dispersion and scale as well as skewness and kurtosis (the canopy height distribution within each cell). This method is more robust than conventional method of moments because they show less bias, variance and ambiguity in the presence of outliers, skewness and small sample sizes ($n < 100$); they also are able to accurately identify and distinguish between a much broader range of distribution shapes (Niemann & Frazer, 2009; Hosking, 1992; Vogel & Fennessey, 1993). Finally, canopy density or cover was computed at 20 equal intervals of relative height following the methods of Gobakken and Næsset (2008). These intervals are combined into percentiles (0,25,50,75) and are then referred to as canopy gap fraction.

3.3.3 LiDAR Data Suitability

One of the assumptions of discriminant analysis is that the sampled points need to be representative of the population. Table 3 shows some measures of sample suitability between the randomly sampled points and the population they were drawn from and their suitability for statistical analysis.

Table 3 - Mean and standard deviation of LiDAR biometric between sample and population

	Sample Mean	Sample SD	Population Mean	Population SD	Kolmogorov-Smirnov Score	Significance
GF 0	0.1715	0.2249	0.1638	0.2128	6.847	0.000
GF 25	0.3058	0.2675	0.2996	0.2625	4.386	0.000
GF 50	0.8199	0.1566	0.4993	0.2692	4.076	0.000
GF 75	0.5035	0.2732	0.8213	0.1465	1.903	0.001
85th % (m)	20.6204	11.8833	21.0221	11.8453	1.423	0.035
Mode (m)	11.4685	12.4373	11.6797	12.3816	7.795	0.000
Kurtosis	0.1529	0.7657	0.1657	1.0615	7.376	0.000
Variance	52.9355	57.9797	54.2433	60.3530	5.973	0.000
Mean (m)	14.2753	9.0745	14.5494	9.0057	1.984	0.001
Skewness	-0.0375	0.2653	-0.0376	0.2689	2.506	0.000
Rugosity	0.5706	0.3110	0.5611	0.3068	3.562	0.000

A bootstrapping validation technique was used to verify performance of the classification model using completely distinct sampled points. Tables 3 and 4 below show the distribution of hazard classes amongst the training and validation sampled points for surface fire hazard conditions.

Tables 4 and 5 show the distribution of hazard classes amongst the training and validation points sampled for canopy hazard conditions. The tables indicate that the proportion of pixels belonging to a given group in both sample subsets is representative of the population.

Table 4 - Distribution of hazard classes for surface fuels in training dataset

Hazard class	N per group	% of total N	Total class % in population
Low	560	56	55
Moderate	387	39	40
High	46	5	5

Table 5 - Distribution of hazard classes for surface fuels in validation dataset

Hazard class	N per group	% of total N	Total class % in population
Low	552	56	55
Moderate	387	40	40
High	46	4	5

Table 6- Distribution of hazard classes for canopy fuels in training dataset

Hazard class	N per group	% of total N	Total class % in population
Low	417	42	43
Moderate	342	34	38
High	234	24	19

Table 7 - Distribution of hazard classes for canopy fuels in validation dataset

Hazard class	N per group	% of total N	Total class % in population
Low	432	43	43
Moderate	365	37	38
High	196	20	19

3.3.4 Field-derived Data

The fire hazard assessment was completed using a CRD water services fuel hazard assessment form developed by Blackwell and modified by the CRD to include parameters of particular interest. The fire hazard model that was created is shown in Figure 4, below. The estimated surface fire hazard, for example, is based on three visually estimated variables: the vegetation flammability, the estimated understory percent of coverage, and woody fuel loading (kg/m^2). Table 8 outlines the relative flammability of selected vegetation types.

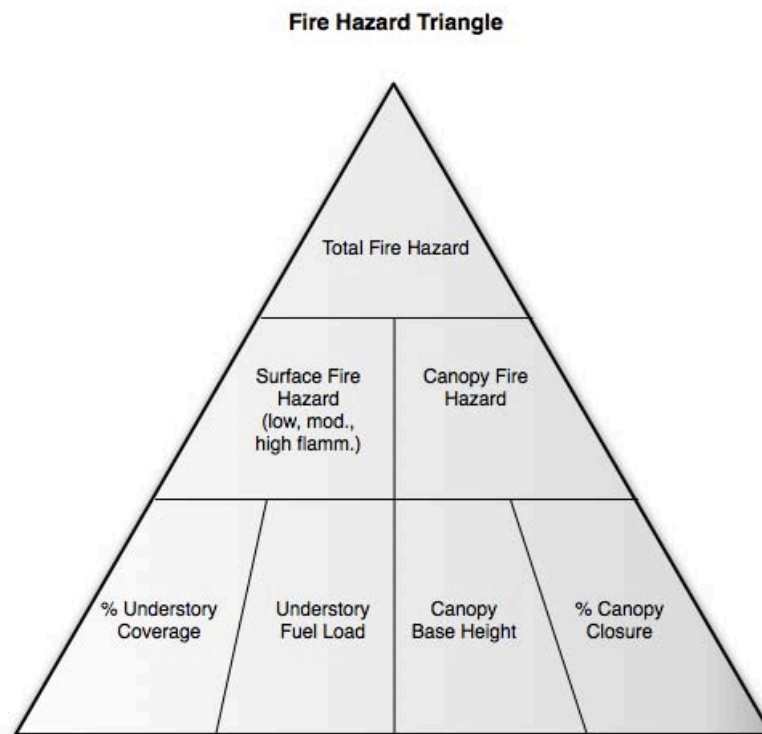


Figure 4 - Fire hazard model triangle (Source: Blackwell, 2006)

Table 8 - Estimated relative flammability of selected vegetation types

Low	Low-Moderate	High
< 10cm green grass	> 10cm green grass <10 cm cured grass	cured grasses
<i>Sambucus racemosa</i> (red elderberry)	<i>Mahonia nervosa</i> (dull Oregon grape)	<i>Pseudotsuga menziesii</i> (Douglas-fir) regen.
<i>Rubus spectabilis</i> (salmonberry)	<i>Gaultheria shallon</i> (salal)	<i>Tsuga heterophylla</i> (western hemlock) regen.
<i>Oplopanax horridus</i> (Devil's club)	<i>Vaccinium Parvifolium</i> (red huckleberry)	<i>Thuja plicata</i> (western red cedar) regen.
<i>Polystichum munitum</i> (sword fern)	<i>Menziesia ferruginea</i> (false azalea)	<i>Pinus contorta</i> (lodgepole pine) regen.
<i>Athyrium filix-femina</i> (lady fern)	<i>Holodiscus discolor</i> (ocean spray)	<i>Cytisus scoparius</i> (Scotch broom)
<i>Lysichitum americanum</i> (skunk cabbage)	<i>Amelanchier alnifolia</i>	Thinning or pruning slash
<i>Alnus rubra</i> (red alder) regen		

When combined with understory percent cover and fuel load, a fire risk hazard class is assigned through the matrices of Table 9 for low flammability, Table 10 for moderate flammability, and Table 11 for high flammability.

Table 9 -Estimated surface fire hazard in areas with understory vegetation of low relative flammability

Surface Fire Hazard - Low Flammability				
		Woody Fuel Loading (kg/m ²)		
		0-10	10-15	15
Understory % Cover	0-30	L	L	M
	30-60	L	L	M
	>60	L	M	M

(Source: Blackwell, 2006)

Table 10 - Estimated surface fire hazard in areas with understory vegetation of moderate relative flammability

Surface Fire Hazard - Moderate Flammability				
Woody Fuel Loading (kg/m ²)				
Understory % Cover	0-10		10-15	15
	0-30	L	M	M
	30-60	L	M	H
	>60	M	H	H

(Source: Blackwell, 2006)

Similarly, canopy fire hazard is based on the combination of two variables, percent of crown closure and the height of understory vegetation to the canopy base height. The canopy base height is visually estimated as one of three categories: less than 2m, 2-6m, and greater than 6m.

Table 11 - Estimated surface fire hazard in areas with understory vegetation of high relative flammability

Surface Fire Hazard - High Flammability				
Woody Fuel Loading (kg/m ²)				
Understory % Cover	0-10		10-15	15
	0-30	M	M	H
	30-60	H	H	H
	>60	H	H	H

(Source: Blackwell, 2006)

A visual estimation of the openness of the canopy produced the percentage of crown closure value was recorded as less than 50%, 50-80 %, or greater than 80%. Table 12 shows how the two variables were combined into the matrix to determine canopy fire hazard. The surface and canopy fire hazard are then used to derive a total fire hazard rating using Table 13.

Table 12 - Estimated crown fire hazard using percentage crown closure and height to base of crown

		Crown Fire Hazard		
		Canopy Base Hgt.		
		>6m	2-6m	<2m
Crown Closure	<50%	L	L	L
	50-80%	L	M	H
	>80%	M	H	H

(Source: Blackwell, 2006)

The CRD uses this information to prioritize forest management practices within those areas sampled for wildfire mitigation. Surface fuel removal in high hazard areas reduces the threat of surface fires transitioning to crown fires.

Table 13 - Estimated total fire hazard using surface fire hazard and crown fire hazard

		Total Fire Hazard		
		Surface Fuels Hazard		
Crown Fire Hazard	Low	L	L-M	M
	Moderate	L	M-H	H
	High	M	H	H

(Source: Blackwell, 2006)

Using mapped areas of low fire hazard, and support from physical topography, fuel breaks are created by the CRD. These fuel thinning efforts are applied to at risk areas, such as critical forested parts of the drainage basin, were applicable. Mapped forest hazard classes are an example of how they have incorporated this data for wildfire mitigation.

3.4 Data Analysis

3.4.1 Classification

The classification of objects into meaningful groups is an important procedure in social sciences. With respect to this research, the classification of LiDAR biometrics into fire hazard classes is an evolutionary step in strategic fire remote sensing. Cluster

analysis is a generic term used to describe a myriad of methods and procedures that can be used to create a classification. Clustering methods are designed to create homogenous groups of cases or entities within clusters. Most cluster analysis methods are relatively simple procedures (Aldenderfer & Blashfield, 1984). Generally, they have the advantage of quickly sorting large amounts of data and identifying groups, or classes, not readily apparent by visual inspection. The strategy of cluster analysis is structure seeking, although its operation is structure imposing (Aldenderfer & Blashfield, 1984). Clustering methods place objects into the groups the method is designed to fit; those groups change in composition with the different methodologies used. This strategy differs from discriminant analysis, for instance, which is more properly described as an identification procedure (Aldenderfer & Blashfield, 1984). Discriminant analysis is used to examine the relationship between a categorical variable and a set of interrelated variables where the categories are known *a priori* (McLachlan, 1992).

This research attempts to classify randomly selected pixels into predefined groups of fire hazard where the classes and partitions are known. Blackwell created the definition of fire hazard and the plots in the field-derived data were given a class designation for fire hazard. The object is to use statistical analysis to predict what hazard class a pixel belongs to based on the LiDAR data. Therefore, the classification technique of discriminant analysis was utilized over the exploratory classification procedures of cluster analysis.

3.4.2 Discriminant Analysis

Discriminant analysis is a statistical technique used to study the difference between two or more groups of objects with respect to several variables simultaneously

(Klecka, 1980). It is particularly useful in three ways; (1) it allows a researcher to assess whether or not variables are valuable in predicting, in the case of this research, ownership of a pixel to a fire hazard class. It allows us to study and interpret the ways the groups differ based on a set of characteristics and additionally which characteristics are the best discriminators. (2) How the variables might be combined into a mathematical equation to predict the most likely outcome. These equations are called discriminant functions and serve to classify which group the pixel belongs to. (3) It is useful in assessing the accuracy of the derived equation (Klecka, 1980). It is only appropriate to use discriminant analysis when the following assumptions are met. First of all, the data cases should be members of two or more mutually exclusive groups, for example, moderate and high surface fire hazard. Second, the discriminating variables need to be measured at the interval or ratio level. Finally, that each group is drawn from a population which has a multivariate normal distribution. A Kolmogorov-Smirnov test was performed using IBM SPSS 18©. The results satisfy the assumption that the input variables be normally-distributed and more importantly, that parametric statistical inferences can be made on this LiDAR dataset.

Given a set of independent variables, discriminant analysis attempts to find linear combinations of those variables that best separate the groups of fire hazard. These combinations are called canonical discriminant functions and have the form of:

[3]

$$d_{ik} = b_{0k} + b_{1k}x_{i1} + \dots + b_{pk}x_{ip}$$

where d_{ik} is the value of the k^{th} discriminant function for the i^{th} case; and p is the number of predictors; and b_{jk} is the value of the j^{th} coefficient of the k^{th} function; and x_{ij} is the value of the i^{th} case of the j^{th} predictor (SPSS, 2009).

The first function attempts to separate the groups as much as possible. A second function is uncorrelated from the first function and provides additional separation in the groups. The procedure continues until a maximum number of functions are created, determined by the number of predictors and categories in the dependant variable. Let us consider a spatial interpretation of this in p -dimensional space. If each data case is a point in space and the groups differ in their behaviour with respect to the dependent variables, then one can observe a group as being a swarm of points concentrated in some portion of this space. It is possible to summarize a group's position by calculating its centroid, an imaginary point which has its coordinates defined by the groups mean of each of the variables. The centroid represents a typical position of the group, which one can then study to obtain an understanding of how the groups differ (Klecka, 1980).

3.4.3 Principal Component Analysis

By their nature, some of the LiDAR biometrics are inter-correlated and have some degree of redundant information. This is because the different bands are biometrics describing the same underlying population. Therefore, they are not appropriate as direct inputs into discriminant analysis. This has necessitated the biometrics to have a principal component analysis (PCA) performed to both reduce the original set of variables into a smaller set and to un-correlate the variables while still preserving most of the information

inherent in them. PCA has no underlying statistical model of the observed variables and focuses on explaining the total variation in the observed variables on the basis of the maximum variance of principal components (Dunteman, 1989). Tables 14 and 15 show the inter-correlation between the LiDAR biometrics used in the classification (only biometrics related to the differing surface and canopy classifications were used). The table for each classification illustrates the necessity for the principal component analysis to be performed before they could be used in discriminant analysis.

Table 14 - Correlation matrix of input variables to principal component analysis for surface hazard discrimination

	Gap fraction 0	Kurtosis	Skewness
Gap fraction 0	1	0.19	0.51
Kurtosis	0.19	1	0.09
Skewness	0.51	0.09	1

Table 15 - Correlation matrix of input variables to principal component analysis for canopy discrimination

	Gap Fraction 50	Gap Fraction 75	Kurtosis	Variance	Mean	Skewness	Rugosity
Gap Fraction 50	1	0.80	-0.05	-0.34	-0.71	0.38	0.52
Gap Fraction 75	0.80	1	-0.01	-0.22	-0.76	0.53	0.80
Kurtosis	-0.05	-0.01	1	-0.07	-0.01	0.09	0.25
Variance	-0.34	-0.22	-0.07	1	0.64	-0.28	-0.09
Mean	-0.71	-0.76	-0.01	0.64	1	-0.47	-0.64
Skewness	0.38	0.53	0.09	-0.28	-0.47	1	0.54
Rugosity	0.52	0.80	0.25	-0.09	-0.64	0.54	1

Aabel©, a statistical exploration and analysis package, was used to perform correlation matrix data reduction while standardizing (z-score) all the variables to take into account their different units. Subtracting the mean and dividing by standard deviation standardized the variables. Following this, the values will have a mean of zero

and standard deviation of 1. This procedure ensures that one variable isn't unfairly weighted against another when calculating the variance in the PCA because of their units. Different combinations of the input variables for the principal component analysis were evaluated and subsequently chosen because of their ability to maximize performance in the classification. A stepwise discriminant analysis technique was also used to determine if other variables had valuable discriminating abilities and to test all possible combinations of variables. The variables ultimately used in the discriminant analysis are in fact the metrics that logically relate to the characteristics of the chosen forest stand. Logarithmic transformations were evaluated; however, they decreased the overall accuracy of the classification slightly so were discarded.

Figure 5 is an example of how variables from the airborne and field surveys were compiled for discriminant analyses associated with the surface.

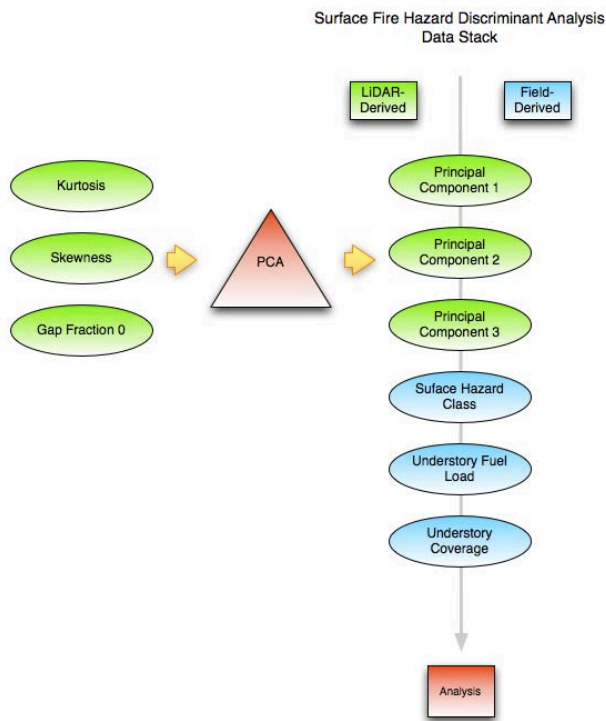


Figure 5 - Example of input variables for statistical analysis

Gap fraction 0%, skewness, and kurtosis were chosen particularly because they pertain more directly with the understory vegetation that needed to be characterized. Gap Fraction 0 is a measure of the porosity of the ground surface. Skewness pertains to the distribution of returns and characterizes the forest vertically. Kurtosis, or the ‘peakedness’, relates to the distribution of height values. Higher kurtosis means more of the variance is the result of infrequent extreme deviations, as opposed to frequent modestly sized deviations. For canopy associated analyses, the gap fraction 50% and 75%, mean, skewness, kurtosis, and rugosity and variance were chosen. The gap fraction bands characterize the canopy closure. Rugosity and variance indicates the surface roughness of the canopy, or variability in the horizontal dimension.

3.5 Summary of Research Method

In this chapter, the methodology for this study was presented in detail. The location for the study in the Sooke Watershed was presented, and the sampling process as well as a description of the data collected is detailed. Stratified random sampled points from the LiDAR dataset have been decorrelated using principal component analysis and discriminant analysis has determined to which fire hazard class the representative pixel belonged in this study. The discriminant analysis was evaluated through discussion of the classification accuracies of hazard classes and its classifying variables.

Chapter 4: Findings and Discussion

4.0 Introduction

The effect of discriminant analysis in determining fuel hazard classes established by the CRD is addressed and results of the analysis are presented. The discriminant analyses of the grouping variables (surface fire hazard, understory vegetation cover, fuel loading, canopy fire hazard, canopy percent coverage and canopy base height) are presented in sections below. The discussion of these findings includes an evaluation of the performance of the discriminant analysis and an examination of the limitations that affect those results. The examination of the results is further broken down into subsections. The implications of the study for forest management practices in the Sooke watershed is also discussed. Additionally, one can relate the results back to the literature and place it into the context of existing strategic fire remote sensing. Finally, recommendations for further research of the discriminant analysis of LiDAR biometrics are discussed.

4.1 Discriminant Analysis Findings

4.1.1 Discriminant Analysis Results

Table 16, below, is a summary table of the results of the discriminant analysis of the 6 grouping variables created by Blackwell (2006) and utilized by the CRD. The overall classification accuracy of pixels successfully classified into the observed classes is presented as a percent in Table 16. The significance level indicates that the discriminant function does better than random chance at separating the groups where $p < 0.05$ statistical significance. In other words, if this value is below 0.05 it suggests that there is a statistically significant difference in the means of the groups and

discriminability is evident. Wilk's lambda is a measure of how well a function separates cases into groups and is equal to the proportion of the total variance in the discriminant scores not explained by differences among the groups; the smaller the value of Wilk's lambda, the higher the discriminatory ability of the function. Kappa coefficient is a measure of the data in the classification table adjusted for the agreement that could be expected by random chance alone.

Table 16 - Summary table for discriminant analysis (DA) results using various discriminating variables

	Overall Classification %	D.A. Statistical Significance	Wilk's Lambda	Kappa Coefficient	Classes	Level of Measurement
Surface Fuel Hazard*	39.2%	0.000	0.950	0.059	Low, mod, high	Derived
Understory Veg. Cover	35.8%	0.026	0.986	0.023	0-30%, 31-60%, 61-100%	Visual
Understory Fuel Load	72.1%	0.003	0.986	0.053	<10kg/m ² , 10-15kg/m ² , 10-15kg/m ²	Derived
Canopy Fuel Hazard*	42.2%	0.000	0.908	0.158	Low, mod, high	Derived
Canopy Base Height	54.8%	0.000	0.643	0.313	<2m, 2-6m, >6m	Visual
Canopy Closure	43.5%	0.000	0.829	0.180	<50%, 50-80%, >80%	Visual

*Dependent on the two discriminating variables listed below

There are many differing interpretations of what constitutes a good level of agreement. In medical sciences statistics, Altman (1991) suggests that a poor agreement is less than 0.20 and a fair agreement is between 0.20 and 0.40. In general terms, the summary shows an overall inability of discriminant analysis to predict the field-derived

classes in which the pixel was associated with. The results do show some promising scores when evaluating the grouping variables individually, especially in the context of how the classes were created in the first place. This is evident when comparing the poor results from surface fire hazard class prediction vs. the strong results of canopy base height class prediction and is discussed further in section 5.1. In the understory fuel loading classification, shown in Table 17, the good classification performance observed is a function of the lack of a 3rd class and a disproportionate loading of the <10kg/m² class.

Table 17 - Summary of discriminant analysis of understory fuel loading

Predicted Group Memberships		
Class	<10kg/m²	10-15kg/m²
<10kg/m²	692 (74.2%)	240 (60.7%)
10-15kg/m²	37 (25.8%)	24 (39.3%)
Total	729 (100%)	244 (100%)

4.1.2 Results of Surface Fire Hazard Class vs. Canopy Base Height

The column labeled level of measurement in Table 16 refers to the precision to which these classes were created in the level of measurement column, either being visually estimated or derived from the combination of the variables themselves. Surface fire hazard classes are derived; the fire hazard model created by Blackwell (2006), Figure 4, determined class thresholds using understory percent coverage and woody fuel loading based on their professional opinion. It additionally separates class matrices for low, moderate and high flammability classes (table 3, 4, 5). In comparison, canopy base height

is one visually estimated grouping variable with less arbitrary classes revolving around 3 easily identifiable thresholds (<2m, 2-6m, >6m). The discriminant analysis shows markedly better performance for grouping variables that were visually estimated. The contrasting results of the discriminant analysis of surface fire hazard and canopy base height are of particular interest and presented in Tables 18 and 19 respectively.

Table 18 - Summary of discriminant analysis of surface fire hazard

Class	Predicted Group Memberships		
	Low	Moderate	High
Low	100 (17.9%)	345 (61.8%)	113 (20.3%)
Moderate	56 (14.4%)	265 (68.3%)	67 (17.3%)
High	9 (19.1%)	14 (29.8%)	24 (51.1%)
Total	558 (100%)	388 (100%)	47 (100%)

The discriminant analysis of surface fire hazard predicted 39.2% of the observed values of hazard correctly overall. Looking at Table 18, one can observe that the discriminant function predicted moderate hazard correctly 68.3% of the time as well as and high hazard correctly 51.1% of the time. The poorest discriminating ability is between the low and moderate class thresholds where ‘low’ is predicted incorrectly to be moderate 61.8% of the time.

The discriminant analysis of canopy base height performed better than any other analysis with an overall accuracy of 54.8%. Table 19 indicates that there is particularly good discriminating performance in the <2m canopy base height class where 76.3% of observed values were predicted correctly. Classification performance is also notable in

the >6m class where 53.9% of values were correctly classified. The major misclassification that occurs in the discriminant analysis is occurring in the boundary between the 2-6m and >6m classes.

Table 19 - Summary of discriminant analysis of canopy base height

Class	Predicted Group Memberships		
	<2m	2-6m	>6m
<2m	183 (76.3%)	47 (19.6%)	10 (4.2%)
2-6m	100 (20.7%)	215 (44.6%)	167 (34.6%)
>6m	21 7.7%	104 (38.4%)	146 (53.9%)
Total	204 (100%)	482 (100%)	271 (100%)

4.2 Discriminant Analysis Discussion

In response to the primary objective of this research, it was possible to discern risk from the large multivariate LiDAR dataset but, in this study, there were significant problems hindering discriminating performance. These limitations are due to the nature of the field-derived dataset. Further problems were evident with regards to the discriminant analysis of surface associated grouping variables, a result of the canopy-centric LiDAR data. The following sections develop these findings further.

4.2.1 Surface Fire Hazard and Surface Forest Attributes

The discriminant analysis of surface fire hazard performed poorly, classifying correctly only 39.2% of the time. In Figure 7, one can see the discriminant model placing group centroids in areas of the principal component feature space and one can identify a problem immediately. Low fire hazard is situated between moderate and high fire hazard; this is counter-intuitive. The principal components are from the LiDAR data and it

reflects the structural nature of the forest stand that the hazard relationships are dependent on, but not the field derived hazard class assignments. If hazard increases from low to moderate to high risk based on increases in the dependent variables, understory coverage and fuel load, then one cannot anticipate that low and high risk will be sharing the boundary in principal component feature space.

The explanation for this is associated with the limitations and error inherent in the field-derived forest attributes and hazard class definition. We are forcing a complete statistical dataset through a field-derived classification with unknown statistical reliability. Our methodology attempts to combine two datasets of very different spatial resolutions together and it is creating error. In particular, it may be because the hazard class designation of a given plot, that becomes associated with a representative pixel in the LiDAR data, was in fact not physically surveyed and lumped into the area because it was identified as being similar by photo-interpretation. Figure 8 reflects this as one can see classes dispersed throughout all areas of the principal component feature space. We are not able to draw any meaningful observations about the behaviour of differing magnitudes of the principal components and the predicted classification of surface fire hazard. This necessitates moving down the hazard pyramid and evaluating the performance of the analysis on the variables constituting hazard in the model.

Understory vegetation coverage performed the worst out of all the grouping variables. We are attempting to assess understory forest characteristics and the majority of LiDAR returns are associated with the canopy. The discriminant analysis of surface fire hazard can be seen as a test of the viability of using airborne LiDAR data that isn't a sufficiently direct measurement of understory vegetation characteristics (<2m).

Understory fuel load performed better than both surface fire hazard and understory vegetation cover in the discriminant analysis. The high overall classification of 72.1% should be taken with restraint; it is a result of the domination by the $<10\text{kg/m}^2$ class in the data and the lack of a 3rd group to misclassify values into as seen in Table 17. The lower discriminating scores for this analysis than surface fire hazard support this. Understory coverage, particularly in the absence of data associated with the canopy, is directly measured by LiDAR and expressed by gap fractions. Fuel load, however, cannot be directly estimated using these datasets. Instead, one can test the relationships between kurtosis, skewness and gap fraction to discern class differences of fuel load indirectly, which did not succeed.

In wildfire research literature, estimating fuel load is the focus of entire research programmes. The dry mass of woody material in the field is measured for an area and regressed against LiDAR data and produces satisfactory accuracies enabling the relationship to be leveraged against the entire LiDAR dataset. With respect to this research, expectations of the discriminant analysis' ability to infer fuel loading was low when one has so many limiting factors due to class definition and visual estimation error in the field polygons. Consequently, the poorest performance with discriminant analysis was attempting to classify this variable. Notably, this grouping variable does highlight the benefit of having a reduced number of classes for discrimination where appropriate.

4.2.2 Canopy Fire Hazard and Canopy Forest Attributes

The discriminant analysis of canopy fire hazard class performed marginally better than any of the surface hazard classifications with an overall accuracy of 42.2% as seen in Table 16. The suitability of the data and additional discriminating variables has

increased the discriminating ability (0.829) over surface analyses considerably. Class definition problems also do not affect the canopy analysis in the same way as they did for the surface analysis because thresholds are considerably more rigid. That being said, the results do indicate a trend of predicting the observed low hazard classes in the moderate or high hazard groups indicating the group centroid for low hazard is not well separated from moderate and high hazard's respective centroids. Unfortunately, like the surface hazard's analysis, no observations about the behaviour of discriminating variables on hazard class can be said with any confidence. While the classification has a statistical significance better than random chance, there isn't the performance necessary to allow watershed wide classification of hazard to be empirically valid. Working down the hazard triangle, however, reveals some interesting results in the discriminant analysis of the dependent variables of canopy fire hazard. The grouping variable of canopy closure can be used as a control for analysis. Canopy closure is directly measured with both extremely high accuracy and consistency over the entire extent of the watershed. One can observe that the LiDAR dataset has been difficult to reconcile against visually estimated values of canopy closure produced in the field-derived dataset.

Relative to the other discriminant analyses performed, however, one can observe increased overall accuracy and discriminating ability for this particular classification; an anticipated result noting the straightforwardness in visually estimating this attribute. The performance of the classification dropped in the boundaries between the moderate class and the adjacent classes of low and high. This is logical as it would be considerably more difficult to decide a class when a forest stand occupies an area around the thresholds of the moderate class. Additionally, the field data acquisition is such that it assigns a class to

a polygon with an average size of 6.2 ha, whereas the biometric data consists of 20m². Discriminant analysis is based on the sampling a single pixel for its LiDAR biometrics and field assigned class. Accuracy in class definition is therefore paramount and there exists too much within-class variability for the field assigned classes. To this end, it may be said that the traditional methodology of large landscape characterizations is ill-suited to combining high resolution of LiDAR with respect to this type of analysis when combining ‘off the shelf’ data inputs. The classification would likely benefit from a field study specific to determine the components of hazard as opposed to the general surveys that have been convention.

A visualization of gap fraction, seen in Figure 8, corroborates this story of a very poorly delineated representation of the observed canopy closure in the field with LiDAR biometrics that are highly representative and accurate. Looking at the rest of the statistical LiDAR biometrics visualized against the polygon statistics suggests the premise is the same. There is too much class variability in the field-derived dataset that cannot be explained by the limitations of the LiDAR data. Either the methodology needs to be modified when scaling this type of high resolution data or the field derived hazard class assignments need to be assigned at high resolution to better suit the specific methodology.

The discriminant analysis of canopy base height outperformed all others analyses with a higher overall classification accuracy (54.8%) and discriminating ability (0.643). The <2m class’ centroid was well delineated from the other groups and this accounts for is excellent classification accuracy (76.3%). We observe with confidence that with values of principal component 1 ≥ 0.75 , principal component 2 of < 0.50 and principal component

3 of ≥ 0.35 , the canopy base height will be less than $< 2\text{m}$ 76% of the time. To further interpret this, one can evaluate the loadings of the principal components and can assess which LiDAR biometrics are the most instrumental to the successful classification of canopy base height.

Principal component 1, which accounts for 52.7% of the variance in the LiDAR dataset, is influenced by gap fraction 75, gap fraction 50, mean, rugosity and skewness, in that order. Of principal component 2, which accounts for 16% of the total variance in the dataset, kurtosis and variance are the most influential. Therefore, gap fraction, which relates to the % of gaps in a given quantile, the mean height of the trees, the surface texture of the trees and the vertical nature of the canopy is fundamental to predicting canopy base height using discriminant analysis. In addition, the nature of variance with regards to the vertical distribution of points and the variance in the dataset also contribute to the performance of that classification. Class determination becomes more difficult beyond the $< 2\text{m}$ class and drops to around 50%. Like canopy closure, the strength in the performance is associated with the increased measurement accuracy of the variable in the field. A canopy base height under 2 metres is very easy to identify in the field; it is a simple question of whether or not vegetation is continuous or less than 2 metre of gap between the understory and canopy. The next two classes, 2-6 metres and greater than 6 metres, becomes more difficult to determine whereas LiDAR does this easily. Relative to the other dependent variables, however, is less susceptible to accuracy error and so one can see a classification accuracy outperforming other analyses. Interestingly, one might logically assume that it might be easier to distinguish an under 2 meter class from a 2-6 metre class than it would be from a 2-6 from an greater than 6m. In terms of the

classification accuracy one can see a possible reflection of this finding, namely that less misclassification in the discriminant analysis occurs around the former than the latter.

Canopy base height is a good example of the ability of LiDAR biometrics to evaluate forest attributes, particularly an attribute that is associated with the understory as opposed to the canopy. Ideally the behaviour of all the grouping variables could be evaluated like canopy base height but unfortunately the field-derived dataset's limitations are prohibitive. Increased spatial resolution and increased estimation accuracy will produce increased classification and contribute to the improved discrimination performance of fire hazard classes.

4.3 Results of Zonal Polygon Statistics of Forest Attributes

The following graphs are a visual indication of the poor class delineations in the dataset and indicate some very basic relationships between the field and remote sensing data. Figure 6, below, is a visualization of the mean values of each polygon for skewness, gap fraction 0 and kurtosis. It is colour coded to show the different surface fire hazard class the respective polygon belongs to. One can observe a relationship where a positive mean skewness and mean kurtosis and an increase in mean gap fraction 0 beyond 25% produce a separation of high surface fire hazard class polygons. There is no discernable separation between low and moderate surface fire hazard class that can be observed however. Figure 7, below, is a visualization of the mean values of each polygon for skewness, gap fraction 0 and kurtosis. It is colour coded to show the different canopy base height classes the respective polygon belongs to. One can observe a relationship where a positive mean skewness and increasing mean gap fraction beyond 15% produce a separation of <2m canopy base height class polygons.

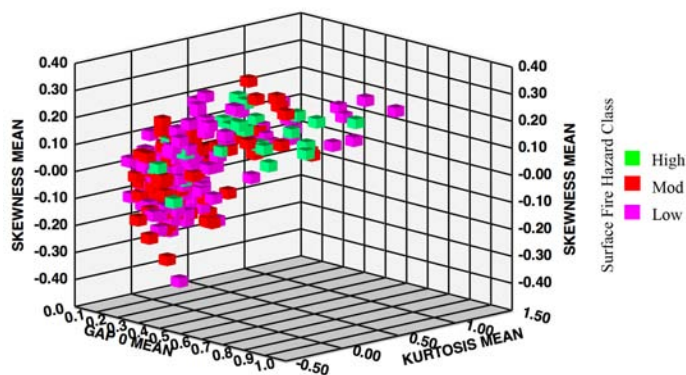


Figure 6 - Mean values of skewness, gap fraction 0, and kurtosis for polygons colour coded to show surface fire hazard class

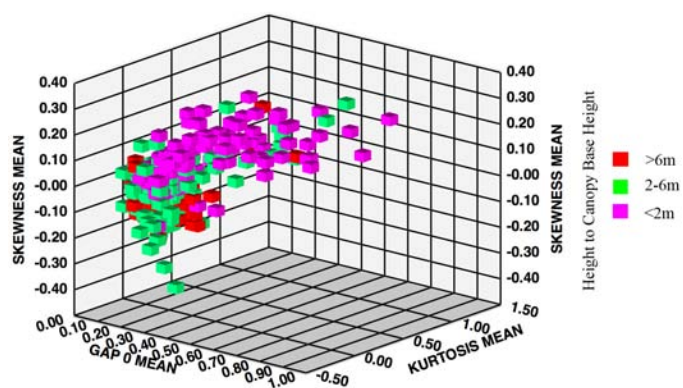


Figure 7 - Mean values of skewness, gap fraction 0, and kurtosis for polygons colour coded to show canopy base height class

Figure 8, below, is a visualization of the mean values of each polygon for skewness, gap fraction 50 and gap fraction 75. It is colour coded to show the different canopy closure classes the respective polygon belongs to. We can observe a relationship where positive mean skewness and increasing mean gap fraction 50 & 75 of the polygon is generally associated with the <50% canopy closure class.

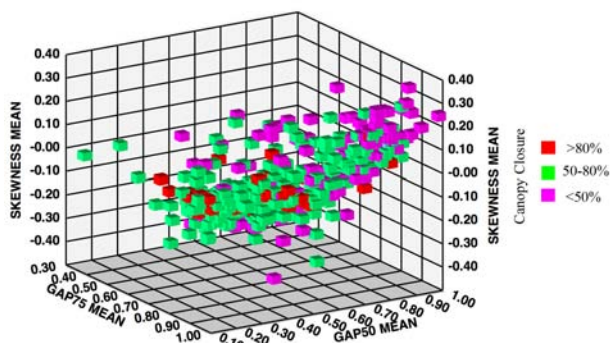


Figure 8- Mean values of skewness, gap fraction 50 & 75 for polygons colour coded to show canopy closure

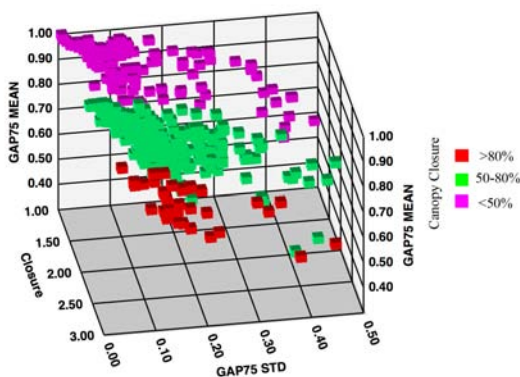


Figure 9 - Mean values of gap fraction 75 and the mean standard deviation of a polygon separated and colour coded by canopy closure class

Figure 9, above, is a visualization of the mean gap fraction in the 75th percentile and the mean standard deviation in that gap fraction separated and colour coded by canopy closure class. In this figure one can observe within canopy closure class variability; similar values of gap fraction 75 and standard deviation can be found across all canopy closure classes.

4.4 Zonal Polygon Statistics Discussion

Zonal statistics comprise the mean, the standard deviation, and the frequency statistics for each LiDAR biometric derivative. By assessing the zonal statistics, in other

words all of the LIDAR biometric data occurring within the bounds on a per polygon basis, one can further address one of the primary research objectives; the examination of the relationship between the different LiDAR biometrics and fire hazard imputed from other observed field attributes. This is important in assessing class delineation and their thresholds, for this and future field plots for this type of research. Examining the mean LiDAR biometric values and standard deviation of field-derived polygons reveals some relationships, information about the within-class variability and nature of the homogeneity.

Relationships that are not apparent when looking at the pixels themselves become more apparent with the perspective provided by this visualization. For instance, for the first time one can see in Figure 7 conditions that are conducive to high fire hazard occurring with positive values of mean skewness and kurtosis and a gap fraction 0 beyond 25%. Despite measurement and class definition issues this fits logically with what would be expected to be found. Understory is indicated by the increased skewness and kurtosis and relates to the shape of the height distribution. Additionally, a low gap fraction 0 relates the porosity near the ground surface that indicates the nature of the vegetation at the ground surface. In Figure 8 there is a well-defined relationship between the positive skewness and gap fraction 0 beyond 15% and <2m canopy base height class. Similarly, this fits with the relationship one would expect to see; namely, a porous ground surface indicating vegetation, increasing understory relating to a lack of definition in the understory and canopy.

Finally, increasing skewness and gaps in the 50th and 75th percentiles are associated with low canopy closure as seen in Figure 8. As the porosity of the canopy and

the presence of understory vegetation increases, the amount of canopy closure decreases; a relationship that is logical expression of the variables involved. We also can observe that with low gap fraction in the 50th and 75th percentiles there is a great deal of mixing in the 50-80% and >80% canopy closure class. This is very indicative of the field-derived dataset and the visual estimation of variables as a whole (polygon) because canopy closure class delineation should be well represented in gap fraction biometrics. Gap fraction is one of the most directly measured variables in the LiDAR dataset and is extremely consistent in its measurement precision and accuracy.

Class mixing with respect to canopy closure is a great example of the limitations of the field campaign for visual estimates of important variables. It is not, however, the only limiting factor negatively affecting class separability. The homogeneity of a polygon can also be highlighted in this type of visualization as seen in Figure 9. Within a polygon, for a give hazard class, variability in LiDAR data associated with a given variable should be clustered as either not very, moderately, or highly variable. In other words, one would anticipate for a high canopy closure class the standard deviation of gap fraction 50th and 75th to be low. As the mean standard deviation of gap fraction increased, in other words the variability of amount of gaps, it would delineate itself from the high canopy closure class. Instead, there exists a great deal of variability in the LiDAR amongst the different polygons for all grouping variables. This suggests that polygons are too heterogeneous with high variability, at least for the particular methodology utilized.

4.5 Discussion of Factors Limiting Findings

A discussion of the factors limiting the performance helps put the results into context as they affect all the iterations of discriminant analysis, albeit some more than

others. The problems encountered generally revolve around the issues in the field-derived dataset, the LiDAR dataset, and how the methodology handles the unification of those two datasets.

4.5.1 Field Derived Data Limitations

The CRD has created a fire hazard model for the forest understory and canopy that is dependent on the classification of its dependent variables. For instance, understory vegetation coverage, flammability, and fuel loading combine to create a class of hazard based on thresholds determined by CRD and Blackwell. These thresholds delineate classes and these classes can be problematic to discriminant analysis techniques. This thesis is attempting to manipulate a LiDAR dataset, which is very structural, consistent and continuous in nature, into a subjective class of hazard that is very conceptual and aggregated. This manipulation of LiDAR data into a pre-existing model is the central focus of this research. The results have shown that this approach has exacerbated the issues of accuracy and precision in the field and photo-interpreted data. Class definitions and labeling have large downstream impact on the classification accuracies. The understory fuel loading class, for instance, is separated into $<10\text{kg/m}^2$, $10\text{-}15\text{kg/m}^2$, and $>15\text{kg/m}^2$ classes, whereas the dataset is virtually without an $>15\text{kg/m}^2$ class.

There are many questions about why fuel loading is separated into these three groups. Was this necessary to be able to visually estimate this variable in the field? Is there an observed relationship about fire risk and fuel load above $>15\text{kg/m}^2$ that warrants its own class to the forest managers or field team? Ultimately one loses a level of precision when assessing this dependent variable and this affects the results of later analysis. In the future one would need to define classes in the field with consideration for

it subsequent use with LiDAR and discriminant analysis. As one moves up the hazard triangle and looks at how surface fire hazard is currently defined in the watershed, there exists an even larger setback to discriminant analysis. Flammability of understory vegetation acts to severely soften the quality of class threshold specified by the surveyor when determining the components of hazard.

The combination of field and LiDAR data into the discriminant analysis would be better served to have data relating to flammability shape the definition of surface fire hazard after the DA initial classification which predicts what class of hazard a pixel belongs to based structural inputs. This is possible by using a decision tree, a cascading series of discriminant classifications that combine to map fire hazard by relating the different multivariate data inputs to each other in the optimal order. The components of hazard are mixed too early, or at too root a hierarchical level for predicting from the highly structural LiDAR data. The field data needed to properly evaluate the performance of discriminant analysis should have well delineated thresholds for each of the components of defined hazard. In figure 10, below, a visual representation of this class definition problem is illustrated. Certain values of the dependent variables, shown on the graph as X & Y, create a resultant surface hazard that can be 3 different hazard classes based on what flammability the polygon was assigned in the field.

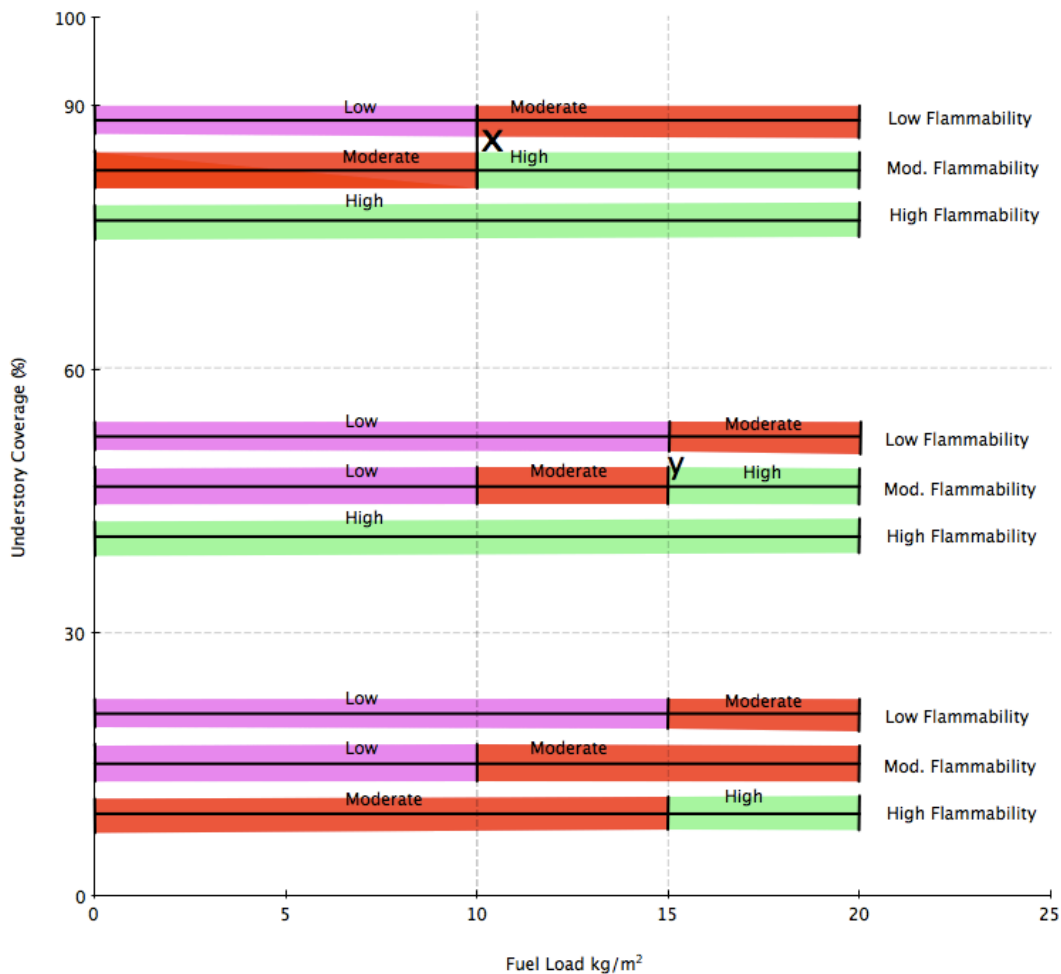


Figure 10 - Visualization of Surface fire hazard class definition problem

Classification performance will be significantly improved by removing flammability from fire hazard triangle when creating the initial discriminant classification. Flammability is an important component to the subjective definition of surface fire hazard, specific to the watershed used as a supplementary data layer for the purposes of analysis being conducted. With that being said, the discriminate analysis should be weighted towards inputs variables that define the physical nature of the stand, particularly when LiDAR is the only remote sensing input. In its current form, with respect to the discriminant analysis, the flammability classes are not adequately defined in the current definition of fire hazard. The field survey assigns flammability classes

through the presences or absence of certain foliage. Flammability is given too much weight relative to understory coverage (%) and fuel load (kg/m^3).

Issues of class definition aside, the much larger concern with regards to this discriminant analysis are the accuracy to which one measures the dependent variables. Unfortunately, the nature of multiplicative error when merging or stacking the data associated with each grouping variable will combine to produce a hazard class in which there can be no confidence in the overall accuracy. Surface hazard, for example, has three variables that have been estimated in the field and merged into a model to produce of class of hazard. The importance in minimizing the error associated with the accuracy of the values of these variables is paramount if fire hazard is to be an effective model. If one over-predicts hazard, the CRD may initiate a expensive treatment program that may not be necessary or fail to mitigate because they believe the situation is safe. Any model that is utilized must have bound of use and be subject to sensitivity analysis to highlight its strengths and weaknesses.

4.5.2 LiDAR Derived Data Limitations

Much like the field derived dataset there are many key factors in the LiDAR data that affect the performance of the discriminant analysis. Among them is the inability of LiDAR, which is best suited to describe the vertical nature or structural characteristics of a forest stand, to estimate the flammability of understory vegetation. Simply put, LiDAR remote sensing systems are not ideal for inventorying vegetation species types. Based on the fire hazard triangle used by the CRD and the field data, flammability of different species types created different class thresholds. This has a significant negative effect in discriminating ability of this type of analysis because LiDAR describes the vertical

characteristics and cannot account for species type in its dataset. The removal of flammability as it currently is integrated in the fire hazard model should improve discriminant analysis performance.

Another factor that could limit the performance of this analysis is the nature of how the LiDAR data is pre-processed for analysis. An analyst classifies raw discrete returns into ground and vegetation points and this can be a somewhat subjective process. Since the biometric statistics are generated from the vegetation-classified points there is potential for more error. In this case, the classification process was conservative with respect to designating ground points so as not to discard potentially valuable vegetation data. In other words, discrete returns that represent the forest floor may have been misclassified and may unduly influence the biometric statistics.

Figure 11 is a profile of the raw LiDAR data and shows the proportion of vegetation to ground points for an area with standing trees and understory vegetation. The magenta points represent discrete returns classified as vegetation whereas the white points are classified as the ground surface. In this image ground points can appear to be above vegetation points but that is a function of displaying depth, or 'width' of the cross-

section in the Z dimension.

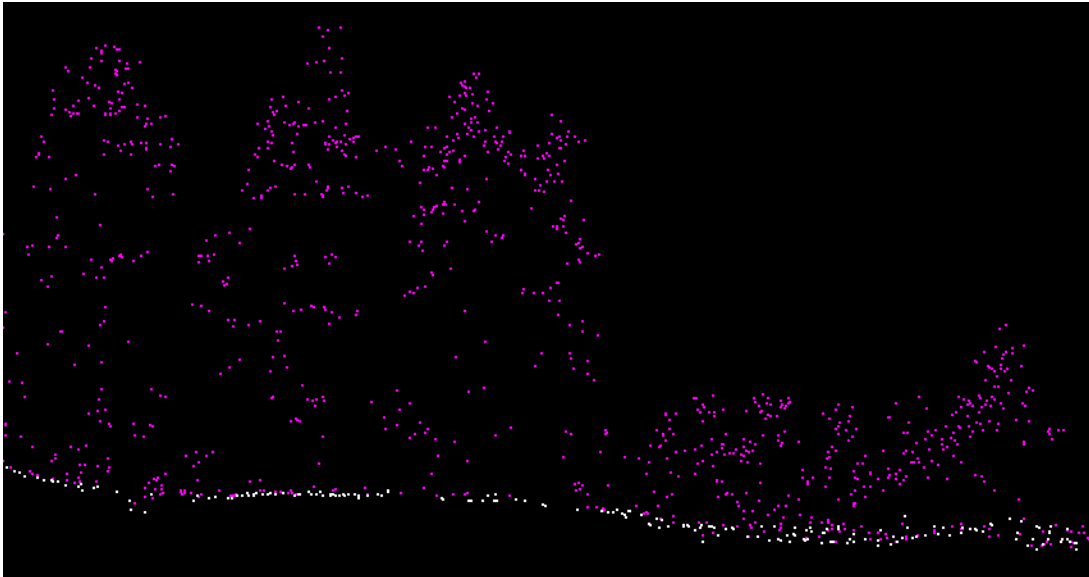


Figure 11 - Profile of raw discrete LiDAR returns undergoing classification procedure (white – ground, pink – vegetation)

In the figure, one can see that a large majority of the discrete returns are associated with the canopy of the standing trees and this is especially true where the canopy is denser as it limits the laser beam's interaction with matter below. Besides the standing trees, one can see shrub vegetation in higher detail than would be obtainable by an obscuring canopy. Recent advances in high point-per-square meter helicopter borne LiDAR systems will produce higher densities of discrete returns per square metre than the system that generated this dataset. More discrete returns will mean more opportunity to penetrate the canopy and offer a richer understory dataset to generate statistics from for the purposes of discriminant analysis. It will also capture more of the forest's inherent variability, straining any methodology that seeks to classify those points into conventional hazard classes.

4.5.3 Spatial Resolution and Methodological Limitations

Through the visualization of within class variability for canopy closure, one can see an example of the lack of homogeneity in the field-derived polygons. For example, canopy closure may be multiple classes within a polygon but the polygon is only assigned one representative class. The average number of 20m² pixels in a polygon is 155 ranging from 3 to 1700 pixels. A polygon represents on average 62,000 m² and is too generalized for use in discriminant analysis as the data representing what are the accurate class designations. Additionally, when the polygon was defined for the field study it is unknown exactly what variable(s) were prioritized when defining its spatial extent or where exactly the plot exists which formed the basis for those variables determination.

Intensively surveying 20m² plots for the input variables used by Blackwell to determine input variables determining hazard would allow for less generalization and error in that dataset. The most significant influence on the performance of this discriminant analysis is that the forest attributes (understory coverage, fuel load, canopy closure and canopy base height) are not measured with the precision or accuracy needed for this type of analysis. Direct measurement is needed, not visual estimation. There is simply too much error associated with visually estimating these variables, especially at a landscape level, and that error is further compounded because these variables must be combined to define a fire hazard class. Both conventional and cutting-edge methods of measuring the dependent variables of the hazard model do exist and should be employed. Hemispherical photography is an example of conventional method of canopy closure estimation in the field that is academically accepted for producing accurate and precise measurements.

The greatest boon to quantifying forest attributes in the field is the use of ground based LiDAR systems. They are highly accurate and precise and can produce estimations of understory coverage, fuel loading, canopy base height with high consistency. With these estimations and canopy closure, which is already available through the airborne LiDAR, the principal variables responsible for fire hazard class definition will be vastly improved. A new field campaign could establish a series of 20m² plots where qualitative derived classes of fire hazard and associated dependent variables occur. The discriminant analysis of this dataset with the airborne LiDAR could have very effective levels of performance warranting the characterization of the watershed as a whole. This could positively change the quality of information being used to prioritize fire hazard reduction efforts and lower the risk of wildfire in this critically important area, especially combined with other known and measurable variables influencing risk such as flammability, aspect, slope, road and stream adjacency.

4.6 Summary of Findings

In this chapter, discriminant analyses of the grouping variables (surface fire hazard, understory vegetation cover, fuel loading, canopy fire hazard, canopy percent coverage and canopy base height) were presented. Results showed some promising scores when evaluating the grouping variables individually. The discriminant analysis of surface fire hazard predicted 39.2% of the observed values of hazard correctly overall. The discriminant analysis of canopy fire hazard class performed marginally better than any of the surface hazard classifications with an overall accuracy of 42.2%. The discriminant analysis of canopy base height performed better than any other analysis with an overall accuracy of 54.8%. There is, however, a great deal of variability in the LiDAR amongst

the different polygons for all grouping variables, which suggests that polygons are too heterogeneous with high variability. With these significant LiDAR-derived data limitations, it can be said that alternative data analysis methods are needed in order to pursue a best-practice approach to fire hazard assessment with currently available data sources and within the existing framework of fire research and management in Canada.

Currently in Canada, the Canadian Fire Danger Rating System (CFFDRS) is the principal source of fire intelligence for all forest fire management agencies (Taylor & Alexander, 2006). It is used to support fire management decision making at strategic and tactical levels and has two primary subsystems, the Canadian Forest Fire Weather Index system (FWI) and the Fire Behaviour Modeling (FBP) system. The outputs from these systems are used in a variety of fire management activities. They are not intended for mapping fire hazard or informing managers on the spatial pattern of risk, rather informing managers of potential active fire behaviour (Taylor & Alexander, 2006). The results from this research will complement the current fire research being conducted by the Canadian Forest Service as it was first identified that the FBP fuel type descriptions do not rigorously or quantitatively follow forest inventory patterns and knowledgeable fire managers will need to develop methods to classify their land base and vegetation data for fire planning (Canadian Forest Service, 1992). These classifications can be used to inform fire managers, hazard like those at the C.R.D.'s Sooke Watershed, to the potential ignition, initial spread hazard, and thinning priorities. They can also be used to assign fuel data to conventional fire behaviour modeling using the FBP models. Taylor and Alexander (2006) state that the national fire danger rating systems must continue to

evolve and improvements will occur as technology advances. The use of LiDAR data as a direct import to fire behaviour modeling in United States national FARSITE fire modeling software by Mutlu, Popescu, & Zhao (2008) is an example of a national fire danger system to evolving through the application of technology and research.

The presence, spatial arrangement and density of fine fuels are the biggest risk apart from fuel moisture content (Mutlu, Popescu, Stripling, & Spencer, 2008), and Varga and Asner (2008) map fire hazard based on a model developed specific to the local geography.

Surface fire hazard classifications in the Sooke Watershed using discriminant analysis build on the conceptual framework outlined by those authors and have the potential to allow mitigation to occur based on information that is more accurate than what is currently available. As these methods continue to progress and utilize new technologies, it is important that they continued to be safe guarded against misuse as identified by Fernandes (2009); Hollingsworth, Kurth, Parresol, et al. (2012); Skowronski et al. (2011). The data inputs into discriminant analysis, similarly, must be approached cautiously and must be verified by a dedicated field campaign to ensure the interpretation of those results are used in a meaningful manner by the fire management at the C.R.D..

Chapter 5: Conclusions

5.0 Introduction

The research problem that this thesis sought to examine was a best-practice method of predicting conventional fire hazards using data drawn from a specific region, namely the Sooke and Goldstream watershed region in coastal British Columbia. This thesis investigated whether LiDAR data can be used to describe conventional forest stand fire hazard classes. Three objectives were posed, namely to discuss the variables associated with fire hazards, specifically the distribution and makeup of fuel; to examine the relationship between derived LiDAR biometrics and forest attributes related to hazard assessment factors defined by the CRD; and, to assess the viability of the LiDAR biometric decision tree in the CRD based on current frameworks for use. Extant data from the Sooke Watershed was used for analysis. The primary methodology used was discriminant analysis, although additional techniques were used as necessary.

This chapter presents a detailed assessment of the findings in relation to the scholarly literature, and presents conclusions to each of the three research questions guiding this study in turn. The chapter also presents the implications of these findings for CRD forest management practices and for future research. Conclusions are drawn on the overall importance of these findings.

5.1 Analysis of Research Question Findings

5.1.1 Research Question 1

Research Question 1 asked: What are the variables associated with fire hazards that should be identified for hazard assessments? Research Question 1 was answered

through an assessment of variables associated with fire hazards which took place in the literature review.

Fuel is defined in terms of the physical characteristics of the live and dead biomass that contribute to the spread, intensity and severity of forest fire (Andrews & Queen, 2001; Arroyo et al., 2008; Burgan, Klaver, & Klaver, 1998). These characteristics are often defined as fuel types (Arroyo et al., 2008; Pyne, Andrews, & Laven, 1996). Merrill and Alexander (1987) define a fuel type as identifiable association of fuel elements of distinctive species, form, size arrangement and continuity that will exhibit characteristic fire behaviour under defined burning conditions. In terms of fuel, understory biomass estimation techniques and the characterization of forest understory are critical components of fuel distribution assessments (Sah et al., 2004). In Canada's FBP system, system fuel maps are interpreted qualitatively, having elements of stand structure and composition, surface and ladder fuels, and forest and floor cover (Nadeau & Englefield, 2006).

Despite some agreement as to how major fuel types should be defined on a broad basis, the literature shows that expressing canopy fuel as a single value for each cell may or may not capture the effects of prescribed fires or other fuel management activities that affect sub-canopy and understory fuel loading (Skowronski et al., 2011). Species types can potentially become arbitrary when describing a local ecosystem, as characteristics of surface fuels and stand age can become more dominant factors in determining the most representative fire behaviour class. To this end, it is necessary to collect data based on local knowledge of fuel types in a region by veteran suppression managers, fuel maps generated from a combination of satellite or airborne imagery (Nadeau & Englefield,

2006), and on-site by field personnel. Given these findings, as suggested by Mutlu, Popescu, Stripling, and Spencer (2008), a DST may more appropriately distinguish between closed canopy fuel models where separability is weak, or where extra conditions are placed on fuel model assignment decisions than other classification techniques.

Given these findings, the best practice defined by the literature is that the variables associated with fire hazards that should be identified for hazard assessments need to integrate both local and global knowledge. These hazards should be encompassed in a DST model, or a similar model, in which changes can be implemented simply and easily based on changes in the physical characteristics of the live and dead biomass in a specific forest area.

5.1.1 Research Question 2

Research Question 2 asked: What is the relationship between derived LiDAR biometrics and forest attributes related to hazard assessment factors defined by the CRD? Question 2 was answered through discriminant analysis, where the independent variables were the six fire hazard variables defined and measured by the CRD; namely, surface fire hazard, understory vegetation cover, fuel loading, canopy fire hazard, canopy percent coverage and canopy base height, and the dependent variable was accurate mapping, as measured by seven LiDAR metrics; namely, mean, mode, rugosity, skewness, kurtosis, 85th percentile, gap fraction, and variance. The discriminant analysis of surface fire hazard predicted 39.2% of the observed values of hazard correctly overall. The discriminant analysis of canopy fire hazard class performed marginally better than the surface hazard classifications with an overall accuracy of 42.2%. The discriminant analysis of canopy base height performed better than any other analysis with an overall

accuracy of 54.8%. There was, however, a great deal of variability in the LiDAR amongst the different polygons for all grouping variables, which suggests that polygons are too heterogeneous with high variability.

The findings from the present study reflect the literature, which suggests that the flawed practical application of research in the field has led to a disconnect between the ways in which fire hazard models have been intended to be used by academic model developers and the ways in which they are used by those tasked with prevention of forest fires, according to the literature (Ager et al., 2011; Alexander & Cruz, 2013; Costanza et al., 2013; Gill & Stephens, 2009). When data is used in a different way than intended by the researchers who create data models and there is a lack of attention given to model application in practice, which is common (Costanza et al., 2013), this can lead to increased risk over time (McKenzie & Kennedy, 2011). Like the findings in the present study, flawed results in data-based models have a limited use in assessing the flammability of natural forests and the effectiveness of fuel treatments in reducing fire potential, even when the data used is altogether accurate (Alexander & Cruz, 2013). In addition, it is important to note that power laws for scaling may or may not have an impact on the efficacy of a model to predict burn, depending on landscape factors, such as fire frequency (McKenzie & Kennedy, 2011) or pine beetle infestation (Alexander & Cruz, 2013). As the landscape of the Sooke basin changes with climate change shifts, proper modelling will become more important.

In general, the findings from this study accurately reflect the findings in the literature that LiDAR data assessment, and assessment of similar data sources, is particularly vulnerable to data invariability (Fernandes, 2009; Hollingsworth, Kurth,

Parresol, et al., 2012; Skowronski et al., 2011). As a whole, interpretation of LiDAR data can therefore be problematic not only because of computational intensity, but also because of the difficulty in displaying these data visually without resorting to some type of classification scheme that can skew the ways in which the data is read and assessed (Skowronski et al., 2011). The lack of viability in the data observed in the current study, with some insignificant exceptions, shows that a significant tradeoff exists between computational requirements for wildfire simulation models and the algorithms commonly used by field teams to apply these models with remote sensing data.

5.1.1 Research Question 3

Research Question 3 asked: Is the LiDAR biometric hazard analysis in the CRD based on current frameworks for use viable or not viable? Research Question 3 was answered through an assessment of the findings from Research Question 2 in combination with the findings from the research literature.

With significant LiDAR-derived data limitations, it can be said that alternative data analysis methods are needed in order to pursue a best-practice approach to fire hazard assessment, and therefore the LiDAR biometric hazard analysis in the CRD based on current frameworks for use is not viable, if used in isolation. It was found that the analyses' classification accuracies primarily suffered from problems inherent from the visual estimation of the forest attributes for the field dataset. Issues associated with class definition and spatial resolution also negatively influenced the overall accuracy of all classifications. A secondary field campaign was utilized to better define class ownership for differing structural stages and a decision tree classification subsequently used the LiDAR biometrics skewness and 85th Percentile height, unbiased by field stratification to

established low and high surface risk for the watershed data. This physical classification delineates structural risk in the watershed, where it can be combined with other variables in a new fire hazard triangle for the CRD. To this end, CRD forest management practices should change, as delineated in the following section of this chapter.

5.2 Implications of Findings for CRD Forest Management Practices

Of particular importance is the assessment of the class definition and thresholds for many of the classes used for the CRD fire hazard model. Many of the class thresholds, for instance understory fuel loading, are ineffectual for further downstream analyses because of issues of precision or statistical significance. Other thresholds are professional judgements and should undergo review to improve within class variability when measured against other datasets, particularly LiDAR data. Ultimately, the most significant implication borne from this research is that the building blocks for empirically valuable survey design to effectively map fire hazards in the watershed must be assessed more carefully. Optimizing how surveyed variables are generated in field plots through the use of statistically reliable methodologies benefit discriminant analysis and other potential classification techniques. This will allow for large-scale, effective and efficient characterization of surface, canopy and total fire hazard for the Sooke Watershed. The establishment of a higher spatial resolution baseline field-derived dataset will greatly benefit the abilities of downstream scientific inquires occurring for this region.

A decision tree classification is a technique, alternative to using LiDAR alone, that is worthwhile with this dataset because it enables the explicit extraction of knowledge and the creation of rules from experts (Jensen, 2005). Additionally, decision tree classifiers simplify classification computations and maintain classification accuracy

(Lillesand et al., 2004). By creating a series of rules that classify unknown observations into distinct and more readily distinguishable groups, it is expected that the classification will perform better than the classification produced through discriminant analysis. Additionally, this type of classification is a more straightforward approach to the concept of surface fire hazard, as it specifically describes the structural component of surface fire hazard with input data that is directly measured. The skewness biometric is a measure of the relationship between the vertical loading of the canopy and the understory vegetation. This biometric effectively indicates the presence or absence of ladder fuels, which is the defining characteristic of surface fire hazard. The skewness value that constitutes the presence of ladder fuels changes for forest stands of differing heights; therefore, thresholds are established over differing height series. 4 major height bins (<6m, 6.01-12m, 12.01-25m, 25-∞ m) were created and are measured by the 85th percentile height statistic. Visiting the watershed and surveying for low and high surface fire hazard plots over the four height series established the value for these thresholds.

Based exclusively on the vertical continuity of fine fuels, a CRD fuels specialist defined the fire hazard as low or high for each plot, and this information is summarized in the Table below, Table 20.

Table 20 - Summary table for field plots & class thresholds* (surveyed May 31, 2011)

	UTM ZN10 Meters N	UTM ZN10 Easting	Surface Fire Hazard Class	Skewness	85 th Percentile Ht.
Site 1	536836	455258	low	-0.214843	23.84
Site 4	5375728	447825	low	-0.075078	13.92
Site 5	5375756	447899	low	-0.486641	42.93
Site 7	5380141	447066	high	-0.077294	25.26
Site 8	5381710	446506	high	-0.271594	64.849998
Site 9	5381305	446830	high	-0.011683	13.51

* Site 2 & 3 had absence of coincident data

Figure 12 is the decision tree classification employed and defines all the rules and use of thresholds that produce the classification of either low or high hazard. This decision tree is a preliminary classification where the output can be combined with supplemental data layers relating to variable associated with fire hazard. It effectively sorts the LiDAR data into two classes that isolate pixels where surface hazard risk is not a concern based on the vertical profile of vegetation existing on the ground.

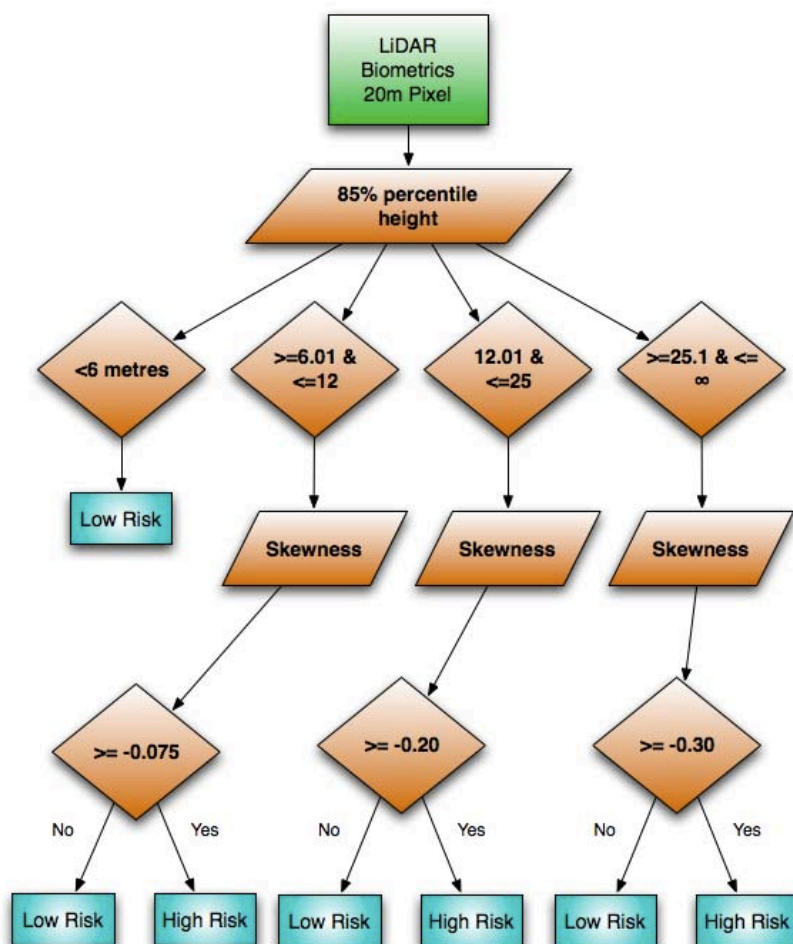


Figure 12 - Decision tree classification

The output of the decision tree classification showing low risk (yellow) and high risk (red) fire hazard based on the vertical profile of vegetation is shown below in Figure 13.

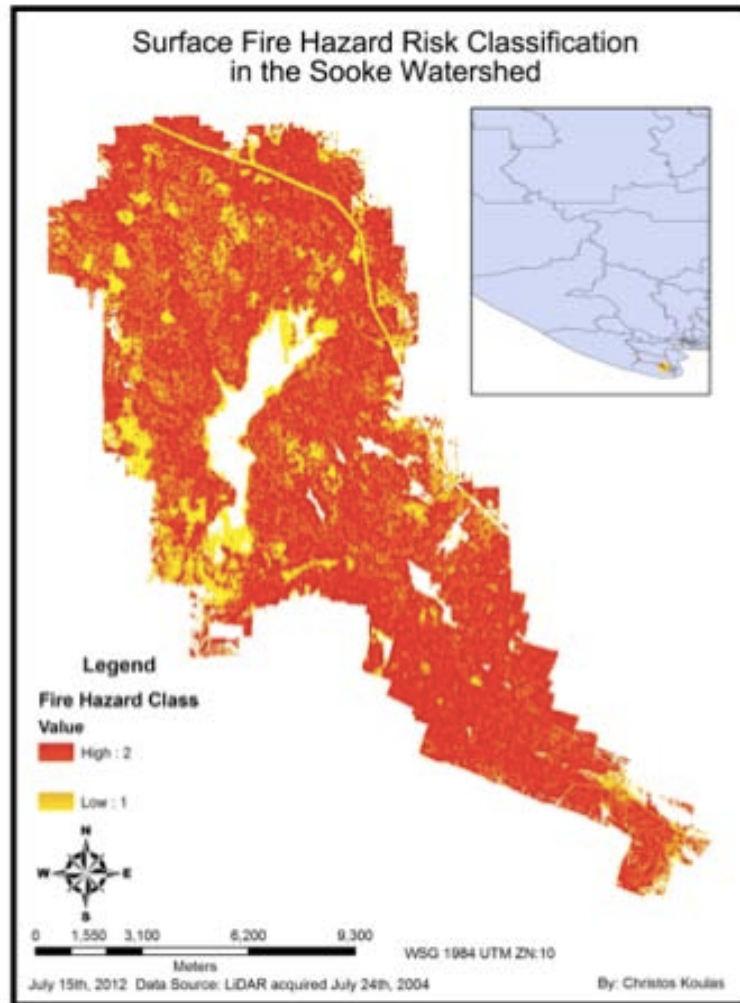


Figure 13 - Surface fire hazard risk classification in Sooke watershed

Based on a modified fire hazard triangle, originally developed by Blackwell and adapted for the recommendations based on the current research, Figure 14 outlines a new hazard triangle and how the decision tree classification approach can be combined with other variables for a modular risk assessment based on the needs of the CRD.

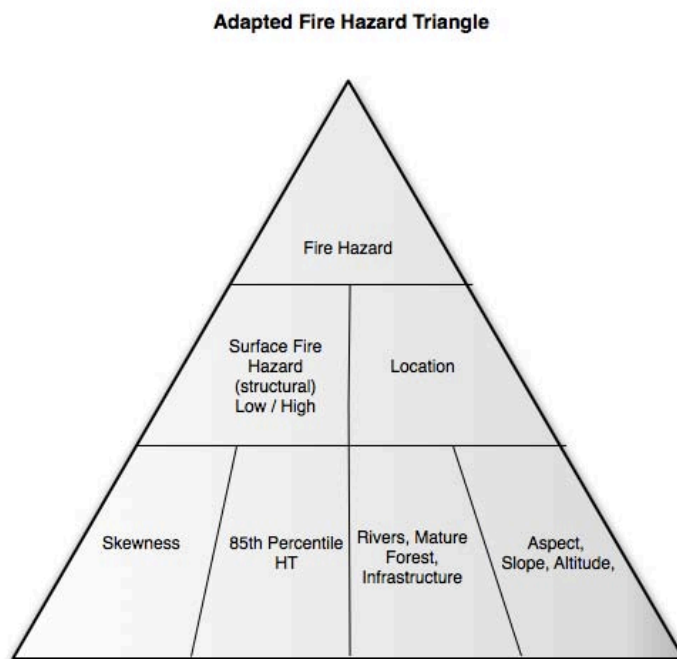


Figure 14 - Adapted Fire Triangle from Blackwell (2006)

The classification output is the structural definition of surface fire hazard and indicates high-risk ladder fuels. In combination with ancillary data such as aspect, slope, road and river proximity, surface fire hazard can be combined to produce landscape level risk assessments. For example, a risk assessment during a drought year may be to highlight areas of high surface fire hazard, with a south-facing aspect, accessible slope, and a proximity to catchment rivers. This could highlight regions where access roads may need to be planned for fire suppression or could help prioritize the limited resources available for thinning efforts. Additionally, the CRD can observe the areas of low surface fire hazard where slope, location and lack of vegetation create natural fire breaks. These areas can be re-enforced through thinning and fire breaks can be planned around sensitive areas like infrastructure.

Finally, biological classifications, like those derived from hyperspectral datasets, can be combined where the spatial characteristics of surface fire hazard and the location of Scotch Broom or valuable old-growth forest might be of interest. The decision tree classification enables systematic generation of surface fire hazard classifications at a reduced cost and an increased efficacy. The greatest benefit, however, is that it provides a more objective acquisition than traditional field teams and photo-interpreted survey.

5.3 Implications for Strategic Fire Remote Sensing Research

To date the bulk of scientific research has used remote sensing data to characterize fuel through the creation of metrics that have been largely of limited use to forest management practitioners. These studies have been very effective in establishing the relationships between differing metrics, such as canopy bulk density and crown base height, but are more suited for use in fire behaviour modeling or assessing biomass for carbon storage inventories. This research establishes that discriminant analysis could be used to effectively monitor and describe surface hazard conditions for forest managers in coastal conifer forests. The central issue in this type of research is retrofitting this dataset into a traditional existing framework. The heuristic ‘boots on the ground’ method of forest hazard assessment is critical, however, the variables tackled are not always representative of what LiDAR is measuring. One must combine the pieces of the risk that can be directly measured by LiDAR and use that information realistically into the definition of fire hazard in use. Further research should focus on growing the framework and exploring the optimal sampling methodologies when using small plots to define entire landscapes.

5.4 Recommendations for Further Research

5.4.1 Recommendations for Segmentation Decision Tree Classifier Research

The limited scope of this field data has meant a segmentation approach could not be properly evaluated when using the decision tree classification. The DST used skewness, which describes the vertical loading over all percentile height bins, rather than using binned components of the vertical fuel loading through Gap75 & Gap0 as depicted in Figure 16. There might be a distinct improvement in the performance of this segmented approach vs. the one used in this research that could be further evaluated.

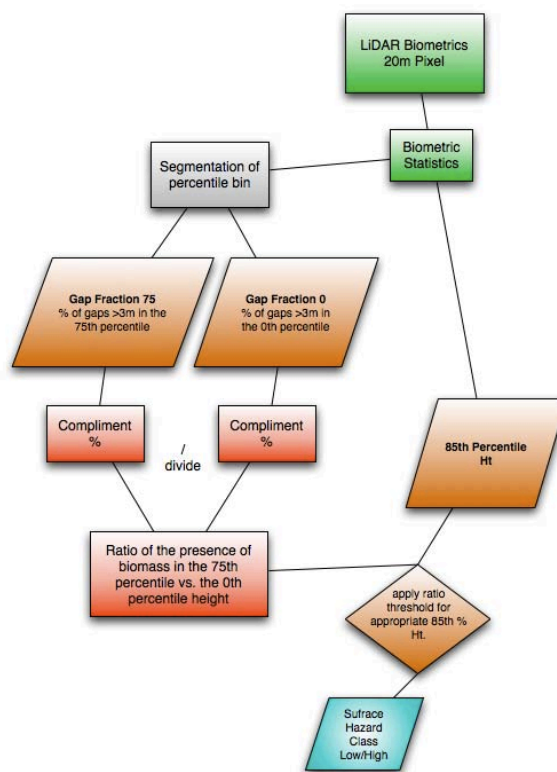


Figure 15 - Segmented approach in DST classification

5.4.2 Recommendations for discriminant analysis research

Discriminant analysis should be further assessed for its suitability in discerning between field derived fire hazard classes through a new field campaign. The campaign

should establish between low, medium and high surface hazard plots through the direct measurement methodologies for the dependent variables of hazard in the adapted fire triangle. A statistically significant number of plots will be necessary to encapsulate for true heterogeneous variability that occurs in the watershed. The capability of discriminant analysis can then be fully evaluated against other classification techniques to determine what is the optimal generalization of LiDAR data when classifying a heuristic hazard class.

5.4.3 Recommendations for Spatial Resolution and Sampling Methodologies

Advances in higher return per square metre systems would better describe the vertical continuity of fuels in higher resolution and would allow for the better understanding between the relationship between the different statistical biometrics, fire hazard, and structural stages. This advance, as well as ground based LiDAR systems which can directly measure many of the dependent variables in the fire triangle, will present new opportunities and new challenges. The coincident use of ground and airborne LiDAR systems will establish high-resolution structural fuel loading profiles that could significantly improve classifications. The new challenges will be dealing with a new robust dataset that might involuntarily introduce noise into classification techniques.

5.5. Conclusions

Many researchers have investigated the quantification of fuel (Andersen et al., 2005; Riaño et al., 2003; Skowronski et al., 2007), classifying fuel types (Arroyo et al., 2008; Lasaponara et al., 2006; Mutlu et al., 2008), and mapping fuel hazard (Varga et al., 2008). However, the importance of this research is that it is the first of its kind to predict fire hazard classes that were established by those responsible for managing a forested

watershed. This research thesis established that discriminant analysis would successfully classify fuel hazard classes better than random chance to varying performance levels, and that CRD forest management practices should change to incorporate a decision tree model in order to decrease risk.

References

- Ager, A. A., Vaillant, N. M., & Finney, M. A. (2011). Integrating fire behavior models and geospatial analysis for wildland fire risk assessment and fuel management planning. *Journal of Combustion, 2011*, 1-19.
- Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster Analysis*. Newbury Park, CA: Sage Publications.
- Alexander, M. E., & Cruz, M. G. (2013). Are the applications of wildland fire behaviour models getting ahead of their evaluation again?. *Environmental Modelling & Software, 41*, 65-71.
- Altman, D. G. (1991). *Practical Statistics For Medical Research*. New York: Chapman and Hall.
- Andersen, H., McGaughey, R. J., & Reutebuch, S. E. (2005). Estimating forest canopy fuel parameters using LIDAR data. *Remote Sensing of Environment, 94*(4), 441-449.
- Andrews, P., & Queen, L. (2001). *Fire Modeling And Information System Technology*. Collingwood, Victoria, Australia: CSIRO Pub.
- Arroyo, L. A., Pascual, C., & Manzanera, J. A. (2008). Fire models and methods to map fuel types: The role of remote sensing. *Forest Ecology and Management, 256*(6), 1239-1252.
- Asner, G. P., & Heidebrecht, K. B. (2002). Spectral unmixing of vegetation, soil and dry carbon cover in arid regions: Comparing multispectral and hyperspectral observations. *International Journal of Remote Sensing, 23*(19), 3939-3958.
- Asner, G. P., Knapp, D. E., Kennedy-Bowdoin, T., Jones, M. O., Martin, R. E., Boardman, J., et al. (2007). Carnegie airborne observatory: In-flight fusion of hyperspectral imaging and waveform light detection and ranging (wLiDAR) for three-dimensional studies of ecosystems. *Journal of Applied Remote Sensing, 1*(1), 013536.

- Asner, G. P., Wessman, C. A., Bateson, C. A., & Privette, J. L. (2000). Impact of tissue, canopy, and landscape factors on the hyperspectral reflectance variability of arid ecosystems. *Remote Sensing of Environment*, 74(1), 69-84.
- BA Blackwell and Associates. (2006). Fuel Hazard Assessment, Greater Victoria Water Supply Area. Unpublished report.
- Beguería, S. (2006). Validation and evaluation of predictive models in hazard assessment and risk management. *Natural Hazards*, 37(3), 315-329.
- British Columbia Forestry Service (BCFS). (2012). *Reports And Reviews*. Retrieved from <http://bcwildfire.ca/History/ReportsAndReviews/>
- British Columbia Wildfire. (2009). *2009 Fire Season Summary*. Retrieved from <http://bcwildfire.ca/History/Summary.htm>
- Burgan, R. E., Klaver, R. W., & Klaver, J. W. (1998). Fuel models and fire potential from satellite and surface observations. *International Journal of Wildland Fire*, 8, 159-170.
- Canadian Forest Service. (1992). Development and structure of the Canadian forest fire behavior prediction system. Retrieved from <http://cfs.nrcan.gc.ca/pubwarehouse/pdfs/10068.pdf>
- Clark, R. N. (1999). Spectroscopy of rocks and minerals, and principles of spectroscopy. *Manual Of Remote Sensing, Volume 3: Remote Sensing For The Earth Sciences* (3rd ed.,). New York, NY: Wiley & Sons.
- Costanza, J. K., Weiss, J., & Moody, A. (2013). Examining the knowing–doing gap in the conservation of a fire-dependent ecosystem. *Biological Conservation*, 158, 107-115.
- Dunteman, G. H. (1989). *Principal Components Analysis*. Newbury Park: Sage Publications.

- Falkowski, M. J., Gessler, P. E., Morgan, P., Hudak, A. T., & Smith, A. M. S. (2005). Characterizing and mapping forest fire fuels using ASTER imagery and gradient modeling. *Forest Ecology and Management*, 217(2-3), 129-146.
- Fernandes, P. M. (2009). Combining forest structure data and fuel modelling to classify fire hazard in Portugal. *Annals of Forest Science*, 66(4), 1-9.
- Filmon, G., Leitch, D., & Sproul, D. (2004). *Firestorm 2003*. Victoria, BC: BC Provincial Government.
- Gill, A. M., & Stephens, S. L. (2009). Scientific and social challenges for the management of fire-prone wildland–urban interfaces. *Environmental Research Letters*, 4(3), 1-10.
- Goetz, A. F. H., Vane, G., Solomon, J. E., & Rock, B. N. (1985). Imaging spectrometry for earth remote sensing. *Science*, 228(4704), 1147-1153.
- Green, A. A., Berman, M., Switzer, P., & Craig, M. D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26(1), 65-74.
- Harding, D. J., Lefsky, M. A., Parker, G. G., & Blair, J. B. (2001). Laser altimeter canopy height profiles methods and validation for closed-canopy, broadleaf forests. *Remote Sensing of Environment*, 76(3), 283-297.
- Hollingsworth, L. T., Kurth, L. L., Parresol, B. R., Ottmar, R. D., & Prichard, S. J. (2012). A comparison of geospatially modeled fire behavior and fire management utility of three data sources in the southeastern United States. *Forest Ecology and Management*, 273, 43-49.

- Hosking, J. R. M. (1990). *L*-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Series B*, 52, 105-124.
- Hosking, J. R. M. (1992). Moments or L moments? an example comparing two measures of distributional shape. *The American Statistician*, (46), 186-189.
- Jensen, J. R. (2007). *Remote Sensing Of The Environment: An Earth Resource Perspective*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Keane, R. E., Drury, S. A., Karau, E. C., Hessburg, P. F., & Reynolds, K. M. (2010). A method for mapping fire hazard and risk across multiple scales and its application in fire management. *Ecological Modelling*, 221(1), 2-18.
- Kim, Y., Yang, Z., Cohen, W. B., Pflugmacher, D., Lauver, C. L., & Vankat, J. L. (2009). Distinguishing between live and dead standing tree biomass on the north rim of grand canyon national park, USA using small-footprint lidar data. *Remote Sensing of Environment*, 113(11), 2499-2510.
- Klecka, W. R. (1980). *Discriminant Analysis*. Beverly Hills, CA: Sage Publications.
- Koetz, B., Morsdorf, F., van der Linden, S., Curt, T., & Allgöwer, B. (2008). Multi-source land cover classification for forest fire management based on imaging spectrometry and LiDAR data. *Forest Ecology and Management*, 256(3), 263-271.
- Koulas, C. E. (2009). Extracting wildfire characteristics using hyperspectral, LiDAR, and thermal IR remote sensing systems. *Infrared Technology and Applications XXXV*, 7298, 72983Q.

- Krivtsov, V., Vigy, O., Legg, C., et al. (2009). Fuel modelling in terrestrial ecosystems: an overview in the context of the development of an object-orientated database for wild fire analysis. *Ecological Modelling*, 220(21), 2915-2926.
- Langford, J. O. & Niemann, K. O. (2006). A systematic approach for using lidar intensity to detect forest structure (Doctoral dissertation). University of Victoria.
- Lasaponara, R., & Lanorte, A. (2006). Remote characterization of fuel types using multi and hyper-spectral data. *Remote Sensing for Agriculture, Ecosystems, and Hydrology VIII*, 6359, 635914.
- Lefsky, M. A., Cohen, W. B., Acker, S. A., Parker, G. G., Spies, T. A., & Harding, D. (1999). Lidar remote sensing of the canopy structure and biophysical properties of douglas-fir western hemlock forests. *Remote Sensing of Environment*, 70(3), 339-361.
- Lefsky, M. A., Cohen, W. B., Acker, S. A., Spies, T. A., Parker, G. G., & Harding, D. (1998). *Lidar remote sensing of forest canopy structure and related biophysical parameters at H.J. Andrews experimental forest, Oregon, USA*. Geoscience and Remote Sensing Symposium Proceedings, IGARSS'98. 1998 IEEE International (Vol. 3, pp. 1252-1254). IEEE.
- McCune, B., & Grace, J. B. (2002). *Analysis of Ecological Communities*. Glendon Beach, OR: MjM Software Design.
- McKenzie, D., & Kennedy, M. C. (2011). Scaling laws and complexity in fire regimes. In *The Landscape Ecology of Fire* (pp. 27-49). Amsterdam: Springer.
- McLachlan, G. J. (1992). *Discriminant analysis and statistical pattern recognition*. New York: Wiley.

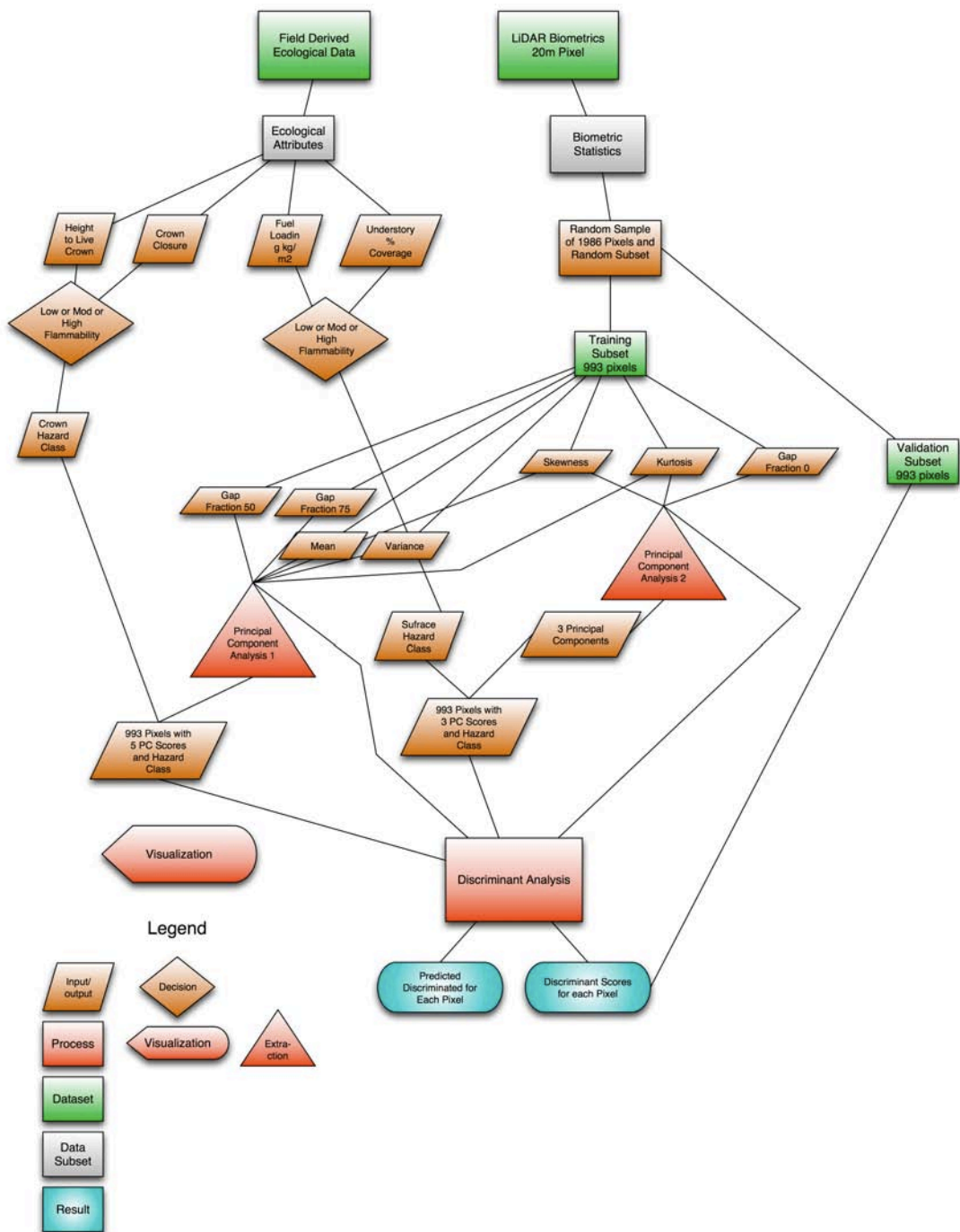
- Merrill, D. F., & Alexander, M. E. (1987). *Glossary Of Forest Fire Management Terms*. Ottawa: National Research Council of Canada.
- Mutlu, M., Popescu, S. C., Stripling, C., & Spencer, T. (2008). Mapping surface fuel models using lidar and multispectral data fusion for fire behavior. *Remote Sensing of Environment*, 112(1), 274-285.
- Mutlu, M., Popescu, S. C., & Zhao, K. (2008). Sensitivity analysis of fire behavior modeling with LIDAR-derived surface fuel maps. *Forest Ecology and Management*, 256(3), 289-294.
- Nadeau, L. B., & Englefield, P. (2006). Fine-resolution mapping of wildfire fuel types for canada: Fuzzy logic modeling for an alberta pilot area. *Environmental Monitoring and Assessment*, 120(1-3), 127-152.
- Naesset, E., & Gobakken, T. (2008). Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research*, 38(5), 1095-1109.
- Niemann, O., & Frazer, G. (2009). *Evaluation of emerging remote sensing technologies for the assessment of forest resources*. Unpublished MSS. University of Victoria.
- Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment*, 56(1), 1-7.
- Popescu, S. C., Wynne, R. H., & Scrivani, J. A. (2004). Fusion of small-footprint lidar and multispectral data to estimate plot-level volume and biomass in deciduous and pine forests in virginia, USA. *Forest Science*, 50(4), 551-565.
- Pyne, S. J., Andrews, P. L., & Laven, R. D. (1996). *Introduction to Wildland Fire* (2nd ed ed.). New York: Wiley.

- Quinn, G. (2011). *Towards an operational root disease mapping methodology through lidar integrated imaging spectroscopy*. Unpublished MSS. University of Victoria.
- Riaño, D., Meier, E., Allgöwer, B., Chuvieco, E., & Ustin, S. L. (2003). Modeling airborne laser scanning data for the spatial generation of critical forest parameters in fire behavior modeling. *Remote Sensing of Environment*, *86*(2), 177-186.
- Rollins, M. G., Keane, R. E., & Parsons, R. A. (2004). Mapping fuels and fire regimes using remote sensing, ecosystem simulation, and gradient modeling. *Ecological Applications*, *14*(1), 75-95.
- Sah, J. P., Ross, M. S., Koptur, S., & Snyder, J. R. (2004). Estimating aboveground biomass of broadleaved woody plants in the understory of florida keys pine forests. *Forest Ecology and Management*, *203*(1-3), 319-329.
- Skowronski, N. S., Clark, K. L., Duveneck, M., & Hom, J. (2011). Three-dimensional canopy fuel loading predicted using upward and downward sensing LiDAR systems. *Remote Sensing of Environment*, *115*(2), 703-714.
- Skowronski, N., Clark, K., Nelson, R., Hom, J., & Patterson, M. (2007). Remotely sensed measurements of forest structure and fuel loads in the pinelands of New Jersey. *Remote Sensing of Environment*, *108*(2), 123-129.
- Taylor, S. W., & Alexander, M. E. (2006). Science, Technology, and human factors in fire danger rating: the Canadian experience. *International Journal of Wildland Fire*, *15*, 121-135
- Thompson, M. P., & Calkin, D. E. (2011). Uncertainty and risk in wildland fire management: a review. *Journal of Environmental Management*, *92*(8), 1895-1909.

- Ung, C., Bernier, P., & Guo, X. (2008). Canadian national biomass equations: New parameter estimates that include british columbia data. *Canadian Journal of Forest Research*, 38(5), 1123-1132.
- Varga, T. A., & Asner, G. P. (2008). Hyperspectral and LiDAR remote sensing of fire fuels in hawaii volcanoes national park. *Ecological Applications*, 18(3), 613-623.
- Visintini, F. (2011). Assessment of two spectral reflectance techniques for the estimation of fuel moisture content, equivalent water thickness, and specific leaf weight in douglas-fir (*pseudotsuga menziesii* (mirb) franco) needles. Unpublished MSS. University of Victoria.
- Zhao, K., Popescu, S., & Nelson, R. (2009). Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Remote Sensing of Environment*, 113(1), 182-196.

Appendix A: Methodology Flowchart

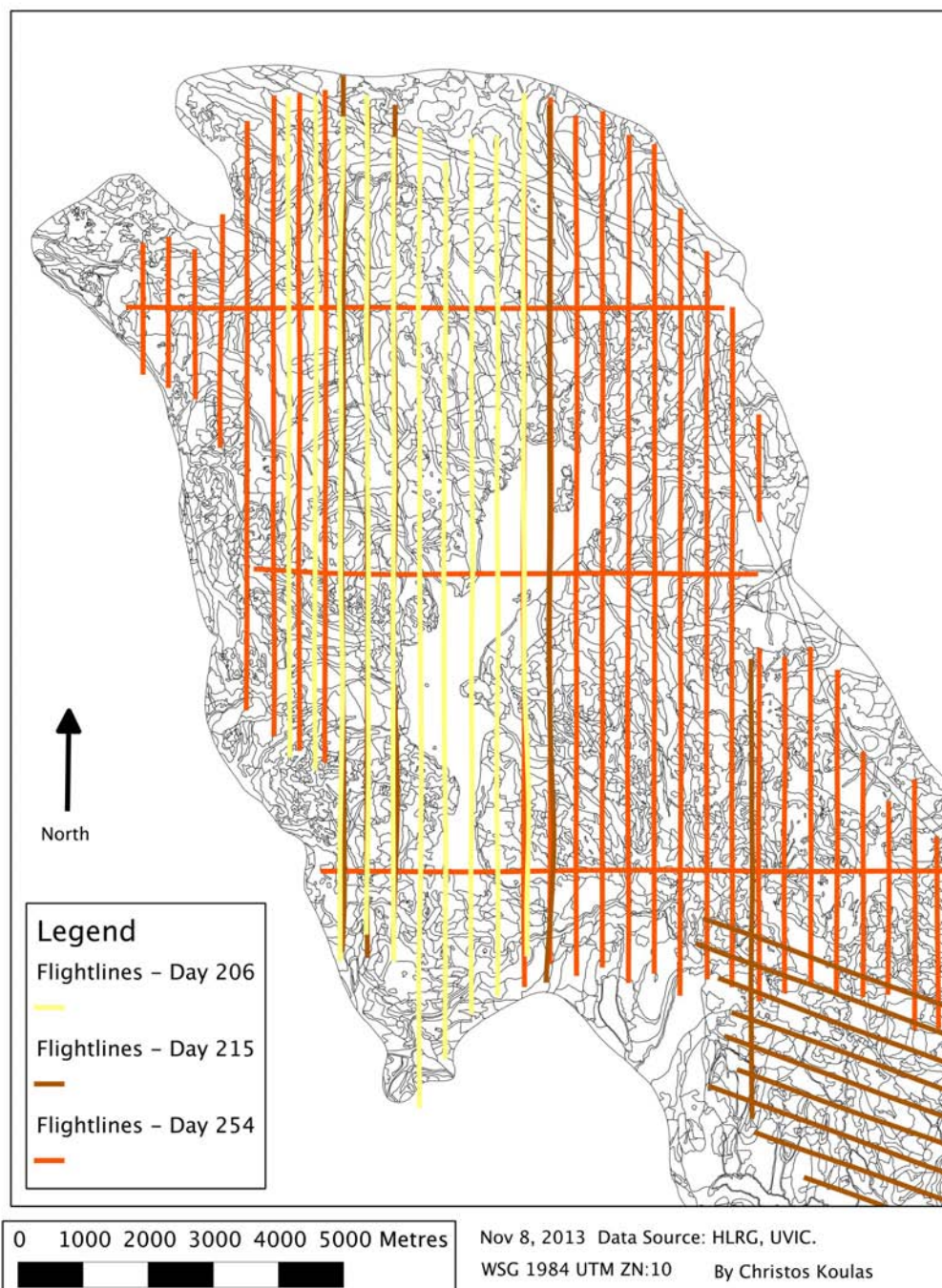
Below is a flowchart depicting the processing of the inputs and outputs of the methodology. The chart is a reference of the different iterations of inputs used to perform the analysis. Looking at the field-derived dataset, need ecological attributes were extracted. These include field-measured classes of height to live crown, crown closure, fuel loading and understory coverage for each surveyed polygon. Additionally, the associated classes of vegetation flammability and hazard for the surface and canopy were extracted. These extracted classes were used in the discriminant analysis with the LiDAR biometrics dataset. Looking at the LiDAR biometrics dataset one can see that the statistics were extracted which include the gap fraction measure, mean, variance, skewness, and kurtosis. They were separated into training and testing datasets as a way of validating the discriminant analysis. The training dataset was subsequently used, with the extracted field-derived classes, to perform the discriminant analysis. This was executed with, and without a principal component analysis to evaluate the principal component analysis' affect, which was ultimately a positive one. The results of the discriminant analysis output a predicted class and discriminant score for each pixel in the dataset.



Appendix B: LiDAR Flightlines

The following map depicts the flightlines flown for the LiDAR acquisition over three days in 2006.

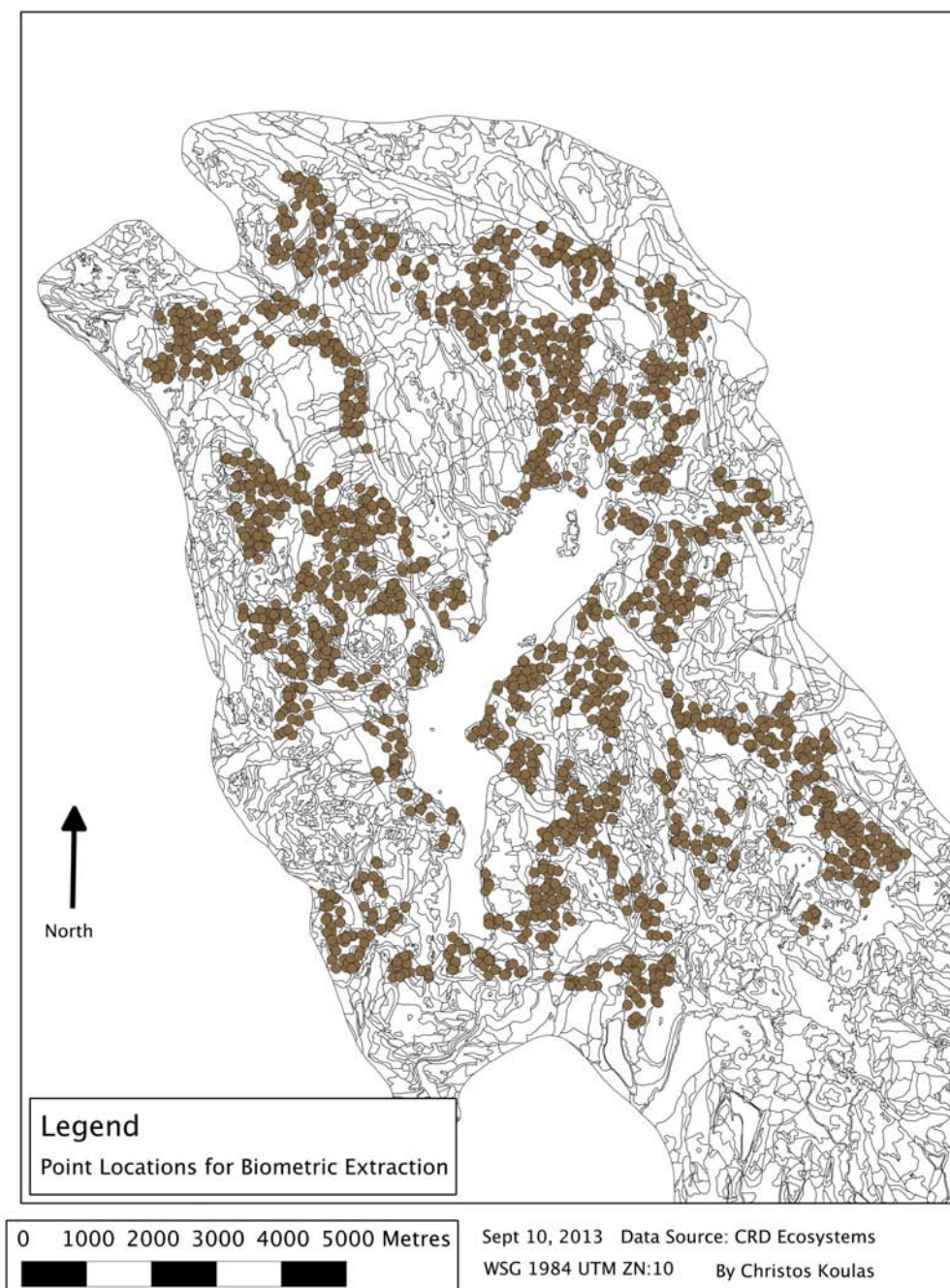
LiDAR acquisition flightlines in Sooke
Watershed 2006



Appendix C: Randomly Sampled Points

The following map shows the 2000 randomly sampled points used to extract the LiDAR biometric statistics and the hazard variables from the Blackwell field data.

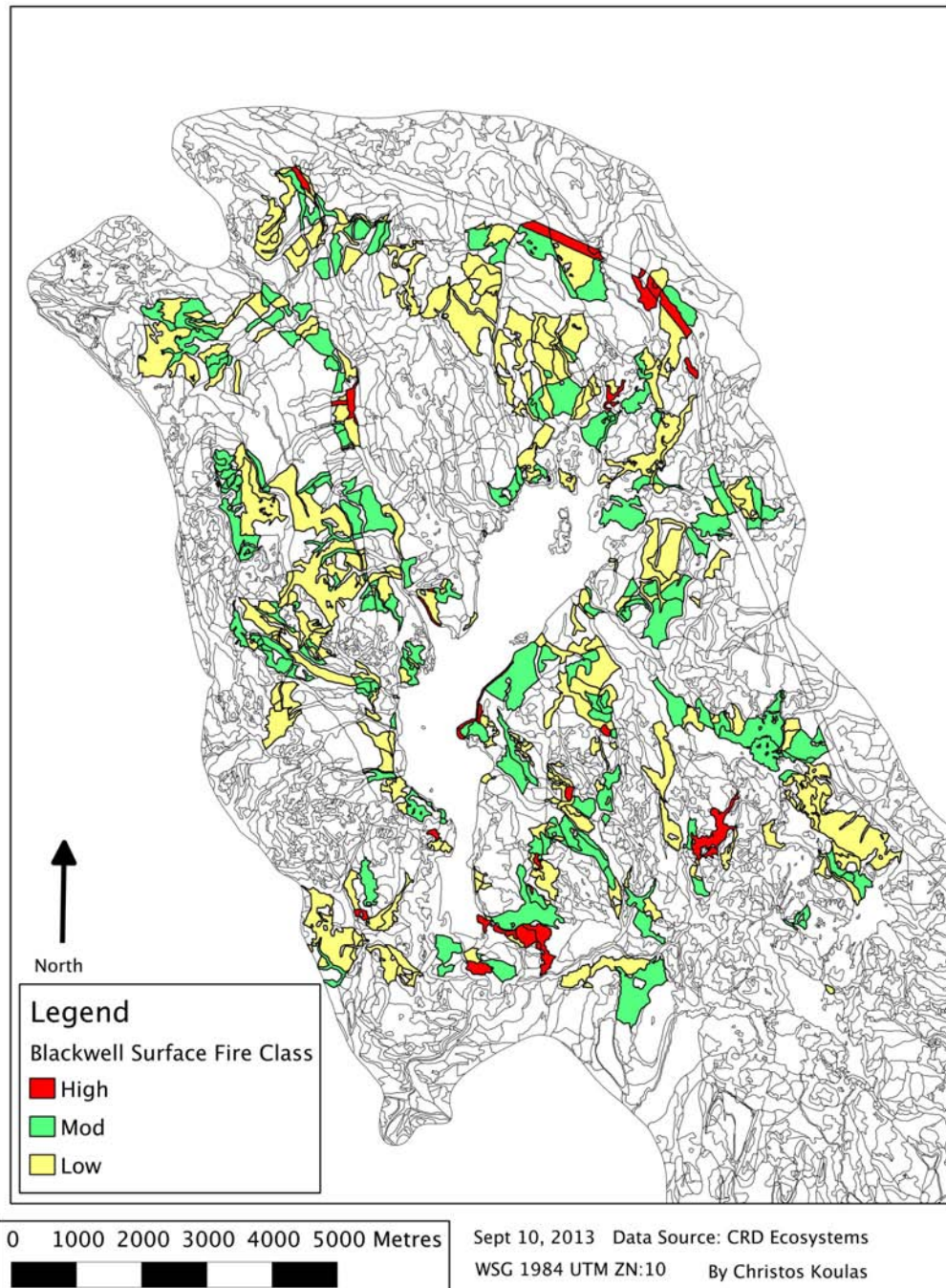
Point Locations for Biometric Extraction



Appendix D: Blackwell Surface Fire Hazard Classes

The following map depicts the surface hazard classes assigned to a field plots by the Blackwell field campaign.

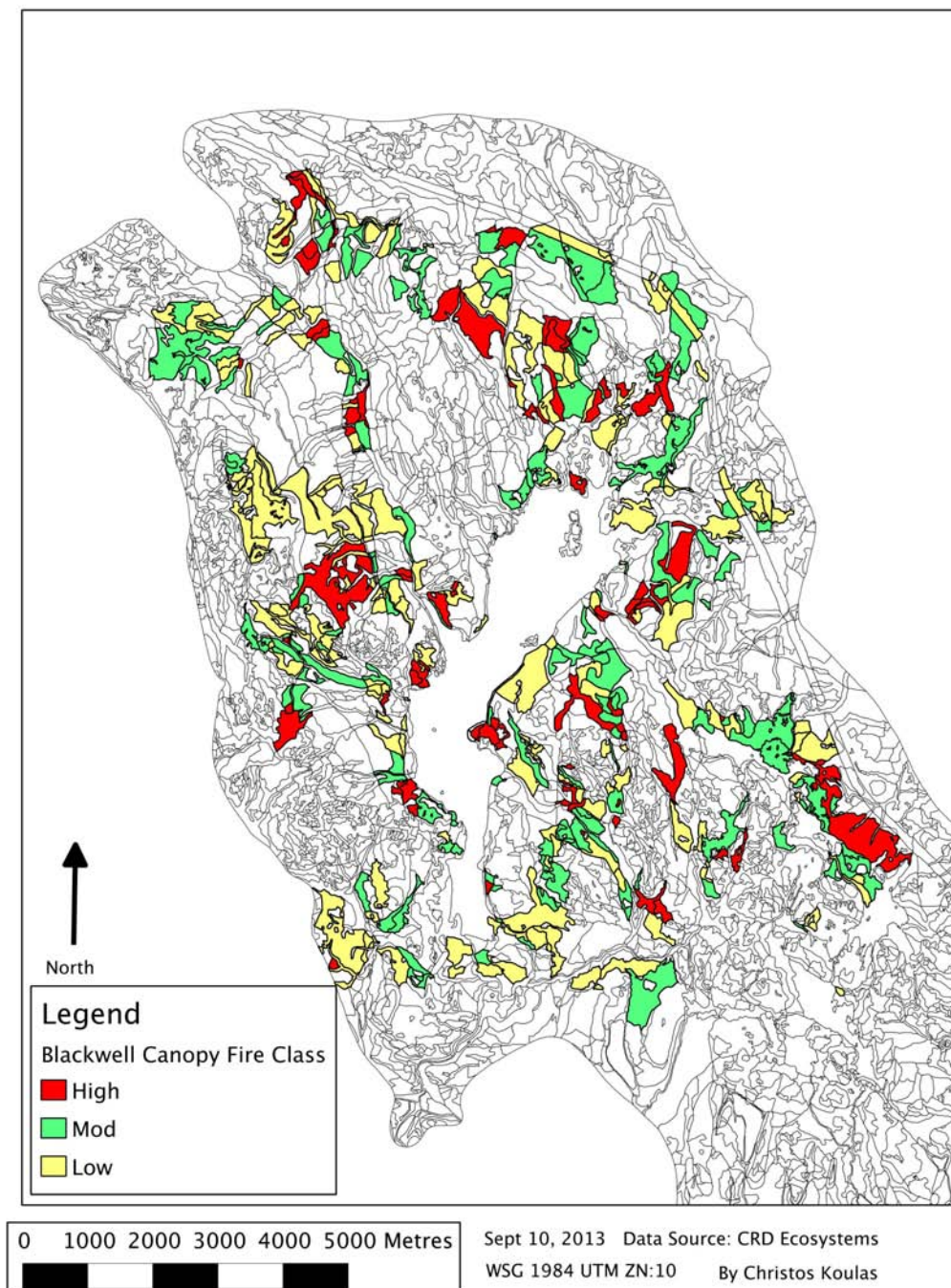
Blackwell Derived Surface Fire Hazard Classes in Sooke Watershed



Appendix F: Blackwell Canopy Fire Hazard Classes

The following map depicts the canopy fire hazard classes assigned to the field plots by the Blackwell field campaign.

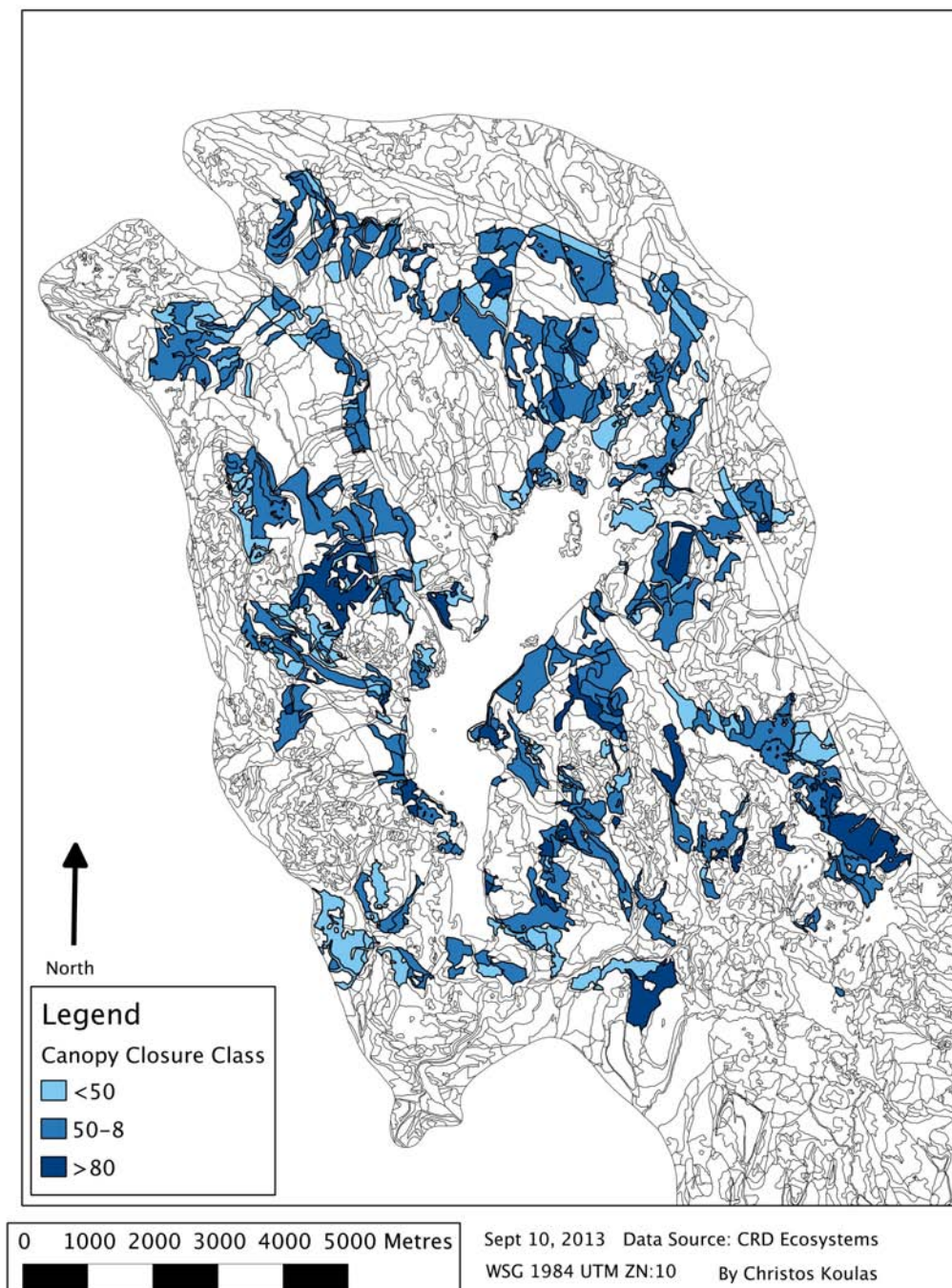
Blackwell Derived Canopy Fire Hazard Classes in Sooke Watershed



Appendix G: Blackwell Canopy Closure Classes

The following map depicts the canopy closure classes assigned to the field plots by the Blackwell field campaign.

Blackwell Derived Canopy Closure Classes in Sooke Watershed



Appendix H: LiDAR Gap Fraction 75

The following map depicts the 75 percentile gap fraction for the LiDAR data as an indication of canopy closure.

LiDAR Biometrics Gap Fraction 75
Percentile

