SPATIO-TEMPORAL VARIATIONS IN SNOW DEPTH AND ASSOCIATED DRIVING MECHANISMS IN A TEMPERATE MESOSCALE MOUNTAINOUS WATERSHED

KELSEY A. CARTWRIGHT

Bachelor of Science, University of Lethbridge, 2016

A Thesis
Submitted to the School of Graduate Studies
Of the University of Lethbridge
In Partial Fulfilment of the
Requirements for the Degree

MASTER OF SCIENCE

Department of Geography
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

SPATIO-TEMPORAL VARIATIONS IN SNOW DEPTH AND ASSOCIATED DRIVING MECHANISMS IN A TEMPERATE MESOSCALE MOUNTAINOUS WATERSHED

KELSEY A. CARTWRIGHT

Date of Defence: May 25, 2018

Dr. C. Hopkinson Supervisor	Professor	Ph.D.
Dr. S. Rood Thesis Examination Committee Member	Professor	Ph.D.
Dr. S. Kienzle Thesis Examination Committee Member	Professor	Ph.D.
Dr. C. Mahoney Internal Examiner	Post Doctoral Fellow	Ph.D.
Dr. C. Coburn Chair, Thesis Examination Committee	Associate Professor	Ph.D.

ABSTRACT

Seasonal snow is a significant source of runoff in Western Canada. Mountainous snow depth distributions are challenging to quantify over large areas. Enhanced monitoring methods can provide the necessary data for more accurate flood and drought forecasts. Using multiple datasets, this research provides the foundation to optimize LiDAR snow depth data collection. Snow depth distribution consistency during mid-winter and melt onset was assessed and depth driver (elevation, aspect, slope, TPI and canopy cover) importance was determined. Consistent inter-annual relationships between aspect, TPI, elevation, treeline and snow depth distributions could be exploited in future sampling designs. Random forest models were utilized to predict depth over a 103 km2 area, based on high resolution (3m) watershed scale and partial datasets. Statistically significant correlations were found between parent and modelled datasets in all trials. This thesis illustrates that machine learning is a promising means of optimizing airborne LiDAR snow surveys in headwater environments.

ACKNOWLEDGEMENTS

There are many people who've made this Master's Degree fulfilling and enjoyable. Firstly, I'd like to thank my supervisor Dr. Chris Hopkinson for the chance to study two of my favourite things, mountains and snow, as well as for the many opportunities to gain invaluable professional experience and travel. I am also grateful for the guidance he and Dr. Laura Chasmer provided as I learned the ropes of teaching and academia in general. To have the first professor that really piqued my interest in water resources – Dr. Stewart Rood, as well Dr. Stefan Kienzle, who showed me how fun and interesting GIS can be, on my committee was an honour.

Craig Mahoney, Reed Parsons, Dennis Quick, Dave McCaffrey, Josh Montgomery and Maxim Ohkrimenko – thank you all for the precious advice and insights, as well as memorable field excursions out at Castle. This thesis would not have been possible without the additional support of my friends and family.

This thesis research was funded through two grants awarded to Dr. Hopkinson: i) the Water Innovation Program, Alberta Innovates – Energy and Environment Solutions; ii) NSERC Discovery Grant. Field and Laboratory equipment was funded through the Canada Foundation for Innovation. Field and aerial data collection logistics were supported by the Alberta Environment and Parks, Castle Mountain Resort and Airborne Imaging Inc. I also wish to thank the School of Graduate Studies and the Department of Geography for the Teaching Assistantship funds, as well the Government of Alberta, the Canadian Space Agency, the Rockyview 4-H District and Loram 99 for the scholarships and grants I received during my graduate studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BRT	Binary Regression Tree
DEM	Digital Elevation Model
DSSM	Digital Snow Surface Model
FL	Flightlines
GIS	Geographic Information System
LiDAR	Light Detection and Ranging
LSDM	LiDAR Snow Depth Model
LW	Late Winter
m a.s.l	meters above sea level
MLR	Multiple Linear Regression
MO	Melt Onset
MSE	Mean Square Error
MW	Mid-Winter
PCC	Pearson Correlation Coefficient
QC	Quality Control
TLS	Terrestrial Laser Scanning
TPI	Topographic Position Index
RF	Random Forest
RMSE	Root Mean Square Error
SDM	Snow Depth Model
SWE	Snow Water Equivalent
WCW	West Castle Watershed

1 INTRODUCTION

1.1 Background

In temperate zone mountainous watersheds, seasonal snowmelt is often the largest source of runoff. The Oldman Watershed in Southern Alberta is primarily snow fed, with 70-90% of streamflow originating in the headwaters from seasonal snowpack (Byrne et al. 2006). The spatial availability of snow data is quite limited relative to the size of the Oldman watershed and the variable terrain within its mountainous headwaters. Across the ~26,000 km^2 Oldman Basin (Figure 1.1), the Government of Alberta currently measures seasonal snow water equivalent (SWE) and snow depth at four upper elevation and five mid to low elevation sites. Daily depth and SWE data are provided at the upper elevations (snow pillows and sonic depth rangers) while the low elevation sites (snow courses) are only measured approximately once a month with manual methods. The density of snow monitoring stations in the Oldman basin is low at approximately one data point per 2889 km^2. Water supply forecasts are generated using linear, long term regression formulae from these data points. Enhanced monitoring could be achieved by utilizing multiple data streams to fill in temporal and spatial data gaps to improve water supply forecasts in data-poor regions.

Instead of adding to and relying on historic water supply trends, more spatially extensive annual datasets have the potential to save lives as well as millions of dollars in flood damage and drought induced crop losses by improving the ability to predict high and low flow events. Over 90% of water extracted from the predominately snow fed Oldman River in southern Alberta is

used for irrigation of agricultural crops (OWC, 2010). In Alberta, floods cost the province over $6 billion in 2013 (Calgary Herald, 2013) and drought related crop losses in 2015 were valued at $700 - $900 million (CBC News, 2015). The crop loss figures do not include losses in Saskatchewan and Manitoba, provinces that also receive streamflow from mountainous headwaters in Alberta via the Saskatchewan River. As part of the greater South Saskatchewan watershed (Figure 1.1), snowpack monitoring is vital to ensuring that sufficient natural flows remain available to meet interprovincial and international transboundary water sharing agreement quotas (Figliuzzi, 2002) while also meeting domestic, municipal and ecological demands in Alberta.

Integrated snowpack monitoring frameworks, utilizing manual, telemetered continuous and remote sensing data are not a new idea. The United States utilizes a network of automated stations, remotely sensed data and field measurements in snowpack monitoring workflows (Molotch and Bales, 2005). The research presented in this thesis provides a foundation for combining machine learning methods and high-resolution airborne LiDAR data as part of an integrated mountain snowpack monitoring framework. Random forest (RF) is the machine learning method that will be explored. Unlike other common statistical modelling techniques, such as multiple linear regression (MLR) and binary regression trees (BRTs), random forest can overcome challenges surrounding correlated predictor variables through random variable selection. This is particularly useful for modelling in mountainous settings. A predictor such as slope, for example, may be correlated with elevation due to the occurrence of cliffs and steeper slopes at higher elevations. Canopy cover, conversely, is generally abundant at lower elevations and declines towards

mountain summits. RF is similar to BRT approaches, although a BRT is a single decision tree whereas RF is an ensemble, or forest, of decision trees which each equally contributes to the final output of variable order importance. As a machine learning technique, RF is still relatively under explored in the statistical snowpack modelling literature where MLR and BRT based publications are abundant. This research aims to determine RF's suitability for use with airborne LiDAR data as part of an integrated environmental monitoring network.

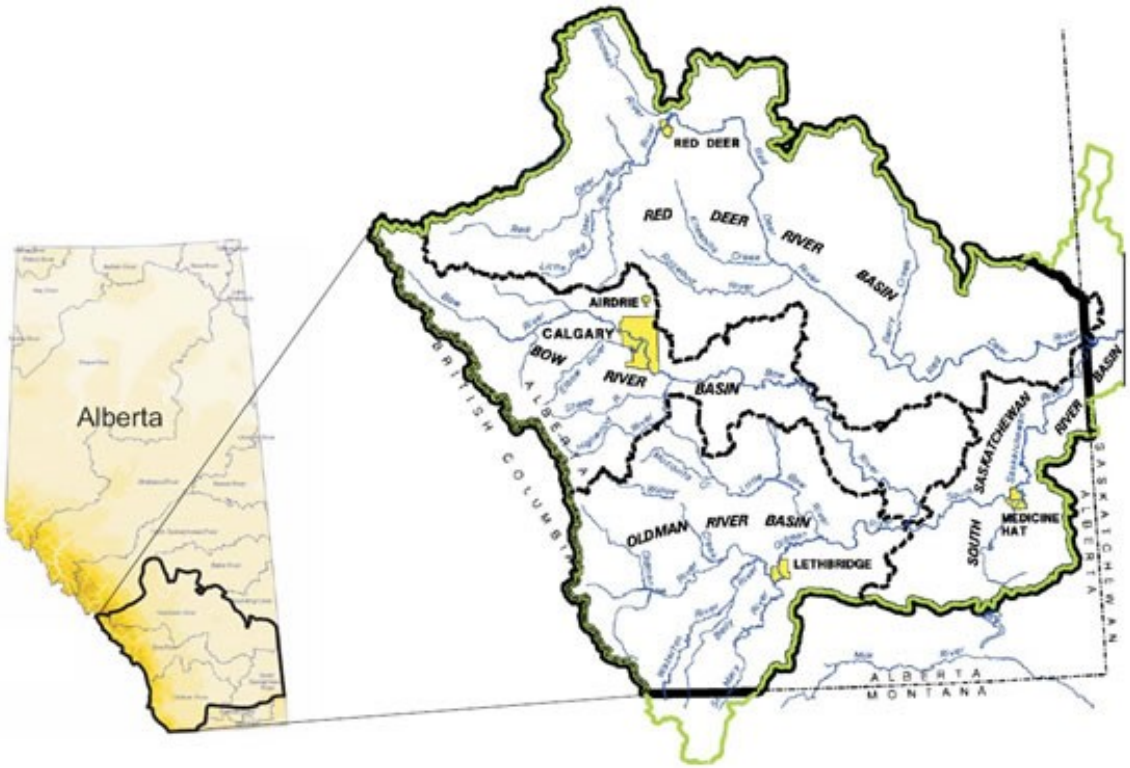

Figure 1.1 The South Saskatchewan watershed is comprised of the Oldman, Bow and Red Deer River basins (Government of Alberta, 2015).

1.2 Snowpack Applications of Airborne LiDAR

Understanding variations in snow depth in mountainous settings is complex due to the many factors that influence snow accumulation such as elevation, slope, aspect and vegetation (Elder et al. 1991) making its measurement and modelling a challenge. Airborne LiDAR is advantageous for measuring snow depth compared to manually acquiring depth measurements as it can provide millions of depth values over large spatial extents in a matter of hours. This research seeks to identify and lay the foundation for optimization of LiDAR snow depth sampling needs by combining high resolution snow depth data with a machine learning based model for spatial imputation. While still in its infancy for modelling purposes, airborne LiDAR derived snow depth data has the potential to revolutionize snowpack monitoring.

LiDAR has been utilized for snow and glacial studies since the early 2000s (Hopkinson et al. 2004; Deems et al. 2013). LiDAR is highly advantageous for mesoscale snow studies as it provides high resolution (1m) spatial grids of snow depth data with up to decimeter precision (Hopkinson et al. 2004; Hopkinson et al. 2012) in far less time than manually acquiring the measurements. With the use of LiDAR for snow depth estimation, researchers can collect and analyze data acquired in alpine areas that were previously inaccessible due to access and safety concerns. For detailed information on the components of a LiDAR system and survey parameters, please see “Hydroscan: Airborne laser mapping of hydrological features and resources” by Hopkinson, Pietroniro and Pomeroy (2006). Data for this thesis was primarily collected with airborne LiDAR, using three different systems. The extent and resolution of

these data (1m) enabled a more thorough analysis of snow depth distributions than would be possible using manual methods.

1.3 Objectives

Declining flows of the Oldman River since the early 1900s (Schindler and Donahue, 2006; Rood et al. 2008), due to damming, anthropogenic withdrawals and increased evapotranspiration as well as decreasing snowpack and glaciers support improvements to integrated snowpack monitoring frameworks in this headwater study area (the West Castle watershed) to ensure future water security for Southern Alberta. Considering the low sampling density of monitoring sites, a logical priority of an improved monitoring regime is to determine if the public sites are good proxies for watershed snow conditions. If so, the potential to integrate those datasets into a new spatially explicit framework should be considered. Understanding the study area's dominant snow depth drivers and their spatiotemporal consistency can help with the optimization of snow depth sampling locations for dense, airborne LiDAR derived datasets. If snow depth at the watershed scale can be inferred from LiDAR sample datasets, combining that approach with knowledge of sampling priority could reduce the spatial extent of airborne snow surveys necessary for an entire study area. With more robust and spatially complete snow depth data, water supply and climate change models would be enhanced to allow for more precise management of freshwater resources.

Government snow data with low spatial resolution (and in some cases, high temporal resolution) were combined with high spatial resolution LiDAR

data to determine the current monitoring infrastructure's representativeness of watershed scale snow depth. Multiple LiDAR datasets enabled examination of terrain and land cover-based snow depth trends as well as evaluation of random forest model performance under various snow pack conditions, mid-winter and melt onset. LiDAR's utility will continue, if not increase, into the future as machine learning and data collection platforms are further evaluated. Objectives are as follows:

i) Given that estimates of snow-based water supply rely on historic trends, we wish to assess how spatially consistent snowpack depth is across different terrain and land cover attributes in the Castle headwaters.

ii) Explore snow depth driver importance and new statistical modelling methods to simulate snow depth across Rocky Mountain Headwater environments using full and partial LiDAR datasets as random forest statistical modelling inputs.

Thesis Organization

First, a review of the terrain and land-cover based snow depth drivers which influence snow depth is given (Chapter 2) along with background on the use of LiDAR in snow studies. A spatiotemporal index of snow depth distributions (Chapter 3) is presented as the first of two research chapters, followed by Random Forest modelling (Chapter 4).

2 LITERATURE REVIEW OF SNOW DEPTH DRIVERS

2.1 Elevation and Hydrometeorology

Much peer-reviewed research exists describing the influence of various topographic and terrain features on the accumulation, persistence, and melt of seasonal snow. Elevation, slope, aspect, canopy attributes, wind exposure, and topographic position all affect snow depth distribution. Elevation has been well established as the depth driver with the greatest influence when examining snow accumulation in both forested and open sites (Daugharty and Dickson, 1982). Understanding the distribution of snow along elevation gradients is important from a water resource management point of view as it is often the case that watersheds with higher variability in elevation carry lower flood risk as spring snowmelt onset and rates of melting are distributed over the full gradient (Hendrick et al. 1971; Alili et al. 2009). Elevation often overrides the importance of other factors as it affects both the magnitude of snowfall events and processes of accumulation and melt due to its effect on orographic precipitation and environmental lapse rates. Other drivers such as canopy cover, slope, and aspect do not influence the magnitude of individual snowfall events but as with elevation, these drivers do affect snow depth accumulation processes and therefore spatial distribution (Anderson and West 1965; Zheng et al. 2016).

Toews and Gluns (1986) found that SWE increased with elevation in both forested and open sites, with 11-15mm SWE increases per 100m elevation gain in forested areas relative to 21-27mm increases in open areas. A 2004 study by D'Eon that examined snow accumulation in open and forested areas along an elevation gradient found that depth was significantly correlated with elevation,

yet canopy and snow accumulation were only significantly correlated at lower elevations (D'Eon, 2004). The author attributed this to greater accumulation at higher sites with less canopy cover as a function of lower temperatures reducing ablation, suggesting that the importance of canopy cover varies with elevation (D'Eon, 2004). Greater snow accumulation at higher elevations, where terrain is more complex and variable, makes it difficult for practical and safety reasons to obtain widespread and comprehensive snow measurements. Means of acquiring more robust depth measurements are crucial to producing accurate SWE estimates.

The first study to utilize airborne LiDAR in Albertan mountains to explain terrain influence was conducted in 2008 (Hopkinson et al. 2012). Their findings provided a baseline for similar results in future studies, particularly concerning elevation and aspect. They found that snow depth increased linearly to the treeline, where depth peaked and subsequently decreased with further elevation gains (Hopkinson et al. 2012). Their results were most accurate in the upper alpine zones as canopy cover decreased and the terrain being sampled by LiDAR was closer to the sensor and thus less prone to error propagation (see Goulden and Hopkinson, 2010). In a LiDAR based study of multiple study areas and repeat scans in the Swiss Alps, Grunewald et al. 2014 found the same elevation distribution as described by Hopkinson et al. 2012 to be the most common pattern (Figure 2.1) LiDAR snow depth measurements are most accurate in open areas (Hopkinson et al. 2004), which has prompted some researchers to only examine LiDAR derived snow depth trends in open areas (Kirchner et al. 2014). Kirchner and others, 2014, studied an area with an elevation range of 1800-3500 m a.s.l., also found a linear increase in snow depth to a point (2050m) after

which depth decreased. They detected orographic effects from maximum elevation down to 2050m, below which differences in aspect, slope, wind redistribution and ablation explained varying magnitudes of depth changes in 100m elevation bins. These factors had a greater influence than orographic precipitation below 2050m, but not above that elevation, which was the point where snow transitioned to rain or descending air masses became depleted of moisture and snow input stopped completely. Type A snow depth distributions (Figure 2.1) have been observed in mountainous watersheds all over the world – in the Central Rocky Mountains of the U.S. (Sospedra-Alfonso et al. 2015), the Spanish Pyrenees (Lopez-Moreno et al. 2005), Switzerland (Moran-Tejeda et al. 2013a) and Iberia (Moran-Tejedea et al. 2013b).

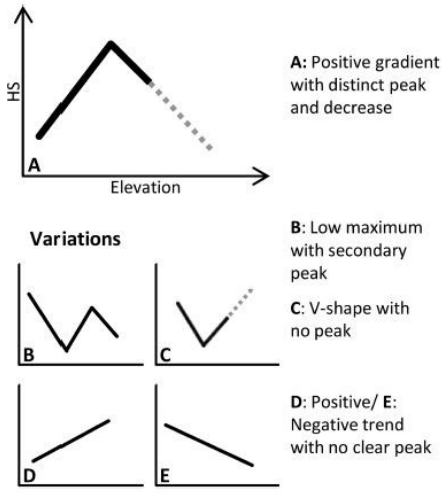

Figure 2.1 Elevation trends observed by Grunewald et al 2014.

A snow accumulation study by Zheng et al. (2016) based in the Sierra Nevada Mountains of California also provided valuable information regarding the interaction between elevation and snow depth. This study examined the influence of other variables such as aspect and vegetation along the elevation

gradient and found that elevation had the highest relative importance in basins with elevation ranges >500m. Both Sierra Nevada, California based LiDAR based studies attributed the location of maximum depth below maximum elevation to a combination of snowfall magnitude below and above the rain on snow transition zone and redistribution from the wind scoured upper elevations (Kirchner et al. 2014; Zheng et al. 2016). Zheng et al. (2016) concluded that elevation was the most important factor influencing snow depth, with the relative importance of aspect, slope and penetration fraction (of light through a forest canopy) varying across the elevation gradient.

The elevation trends presented thus far all suggest similar conclusions: a defined snow depth peak is observed below maximum elevation. Similar hypotheses have explained this result as orographic precipitation and wind redistribution influencing the location peak depth while lapse rates drive the linear increase in depth from the elevation minima to peak snow depth. Environmental lapse rates describe the nearly linear decrease in air temperature with increased elevation, which can vary widely in alpine settings (Pigeon and Jiskoot, 2008). Differential temperatures along an elevation gradient (Figure 2.2) can lead to more persistent snow and/or delayed melt at higher elevations. Orographic precipitation events promote differential precipitation intensities at the watershed scale, in the form of rain or snow, varying with moist adiabatic lapse rate of saturation specific humidity, and the wind speed perpendicular to the mountain and the vertical displacement of saturated air parcels above the windward slope (Shi and Durran, 2013). X-band radar has demonstrated that orographic snowfall is enhanced near mountain summits (Mott et al. 2015), but

these inputs are then subject to wind and gravitational redistribution, so orographic effects cannot fully explain snow depth distribution.

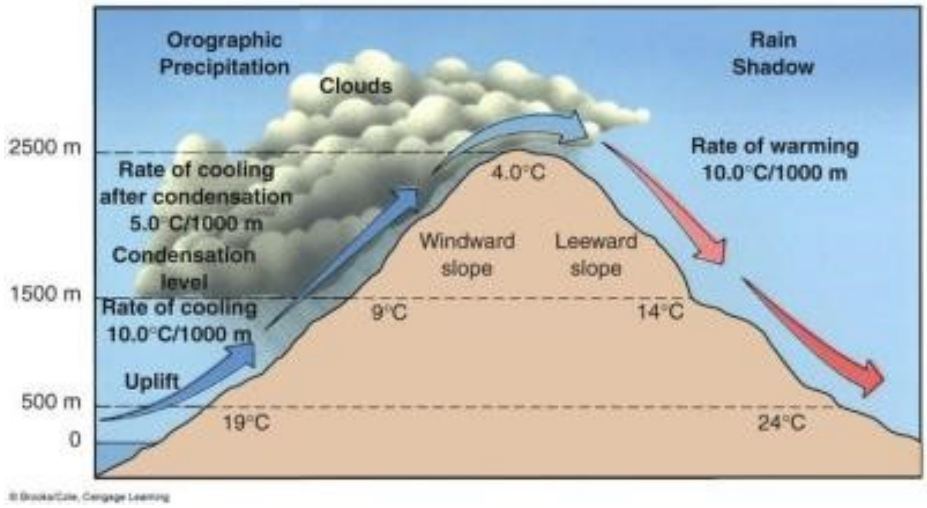

Figure 2.2 Orographic precipitation and environmental lapse rates (Butler, 2018).

With orographic effects mostly limited to upper elevations (Mott et al. 2015), it is believed that a point exists along mountainous elevation gradients above and below which precipitation and temperature are, respectively, primary influences of snow depth distributions. At this point, a break in the linear increase of snow depth with elevation is observed and depths begin to decline (Grunewald et al. 2014). This elevation-based boundary zone has many potential implications for water supply and flood forecasting. An approximate 0.8˚C increase in temperature in the American Pacific Northwest over the last century (Mote et al. 2014) is believed to have caused the runoff pulse of seasonal snow to occur 5-15 days earlier in basins whose maximum elevation is less than ~2000m (Steward et al. 2004; Fritze et al. 2011). Sospedra-Alfonso et al. (2015) determined that this threshold has increased by 140-280m since the late 1960s. Above the threshold, where precipitation is the driving force, snow inputs are

determined by storm events, the track of each storm, and how wind, land cover and topography interact to redistribute snow. If the elevation break for temperature driven snow distributions continues to migrate upwards, there is potential that the perennial snow line, a macroscale phenomenon, will shift downwards. Downwards perennial snowline migration has already been observed (Pelt et al. 2016). Should these changes persist into the future, less snow (and glacial) inputs to streamflow will be available throughout summer as a source of runoff. This input is valuable for its slow release as rivers naturally approach baseflow.

2.2 Canopy Cover

Canopy cover is perhaps the most complex driver to analyze with respect to its effect on snow depth distribution. The interaction between canopy cover and accumulation or ablation can be complicated by:

> "snowfall magnitude (Anderson, 1956), year to year variations (Berndt, 1965), elevation (Daugharty and Dickinson, 1982), aspect and slope (Anderson et al. 1958a; Moore and Wondzell, 2005), size of the clear-cut used as a reference (Golding and Swanson, 1986), wind speed (Woods et al. 2006), local weather conditions (Lundquist et al. 2004; Lundquist and Flint, 2006), spatial distribution of trees (Dunford and Niederhof, 1944; Veatch et al. 2009), and canopy geometry (Essery et al. 2008)" (Varhola, 2010).

The size of open areas, whether they be clear cuts or naturally occurring, can affect the magnitude of accumulation relative to forested stands (Pomeroy et al. 2002). Smaller open areas can be impacted by the shelter of surrounding forests, which may reduce ablation as outlined below leading to differential melt rates relative to forested stands, whereas large open areas can have reduced

accumulation due to increased wind erosion (Pomeroy et al. 2002). The feedback between forest cover and the processes of accumulation and ablation makes separating trends and attributing them to a single process extremely difficult in treed areas. Intermediate sized open areas have maximum accumulation, with the sizes of plots being relative to each other and specific to the study.

Ablation is primarily driven by the net availability of energy from incoming and reflected shortwave radiation, incoming and outgoing longwave radiation, sensible, latent and ground heat fluxes (Brooks et al. 2003; Boon, 2007). Forest cover can affect the energy budget (Figure 2.3) relative to open areas in that vegetation increases long wave radiation inputs but reduces short wave radiation from the sun leading to a net loss of energy available for melt at certain times of year (Essery, 2008). While snow accumulation is typically greater in clear cuts and open areas, radiative budget differences often lead to these areas melting out faster than in surrounding forested stands. Some studies show a 10-day difference between melt onset in open and forested areas (Berdnt, 1965). The magnitude of sensible and latent turbulent heat fluxes, which are also part of the net available energy for snow ablation, are lower in forested areas due to reduced wind speed by canopy cover. Because of canopy cover, melt rates in forested areas can be up to 70% lower than in open areas (Hendrick, 1971; Boon 2007; Teti, 2008). The distribution of snow depths in mountainous watersheds due to the effects of radiation on snow accumulation and melt can be more difficult to identify than elevation because insolation is affected by so many interacting variables.

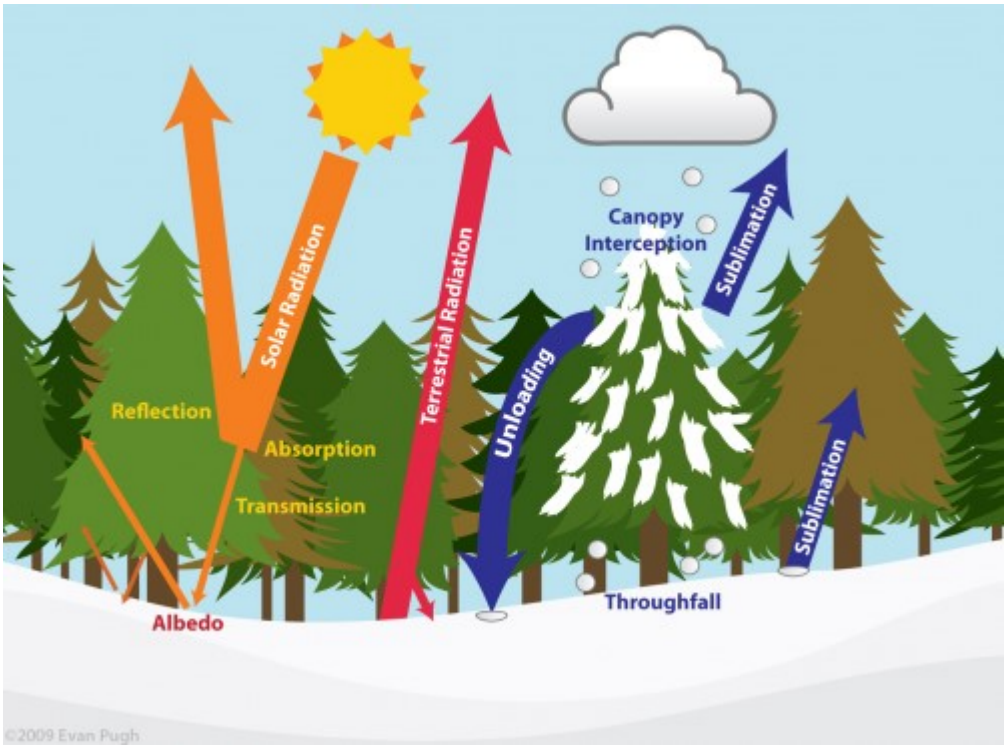

Figure 2.3 Radiation sources in forested zones (Pugh and Gordon, 2012).

As previously discussed, LiDAR snow studies have produced results indicating that snow depth increases to a certain elevation after which depth starts to decrease (Hopkinson et al. 2012; Kirchner et al. 2014). Another explanation for hypsometric depth distributions is that snow from higher elevations is gravitationally redistributed and intercepted by vegetation below resulting in maximum snow depths at treeline (Hopkinson et al. 2012). The transition from sparse vegetation in upper alpine zones to a denser canopy at treeline leads to enhanced interception (Moeser et al. 2015) but also increases the potential for snow retention due to a variety of vegetation heights trapping snow below and within their branches. As the canopy thickens, reduced accumulation can be observed due to the sublimation of snow following canopy interception (Essery et al. 2003). Up to 60% of cumulative snowfall can be

intercepted by the canopy (Hedstrom and Pomeroy, 1998), much of which is then subject to sublimation (Pomeroy, 1998). As a result, forested sites can have up to 40% less snow than open reference areas (D'Eon, 2004; Winkler et al. 2005, Jost et al. 2007). Small snow events allow for a higher proportion of interception and subsequent sublimation (Brooks et al. 2003) as the capacity of the canopy to withstand the weight of the snow loading is greater, providing more time for sublimation to occur. The influence of canopy cover varies with the size of snow events because tree branches do not have the ability to intercept an unlimited amount of snow (Boon, 2009). As observed with topographic depressions, the degree of influence from features such as short vegetation or a small topographic depression diminishes after snow depth exceeds the vegetation height (or depression depth), creating a smooth surface over top with older snow layers below that were influenced to a greater degree by these small features. The effects of various tree species and/or vegetation classes on snow accumulation and ablation is an area of snow hydrology that remains poorly understood.

2.3 Aspect and Slope

In the Northern Hemisphere, north facing slopes experience the lowest exposure to solar radiation thus greater snow accumulation in these areas is observed relative to south facing zones where incoming solar radiation is greater (Golding and Swanson 1986; Anderson et al. 1958a). Solar radiation fluxes peak around mid-day when the south-facing slopes are exposed to sunlight while north-facing slopes remain shaded (Figure 2.4). Short and long wave radiation provide the energy for melt and refreezing (Marshall et al. 1999) which can create preferential depth losses and potential density gains in aspects

with high radiative inputs (south) versus those with little (north). This trend is still observable when aspect zones are stratified by elevation (D'Eon, 2004).

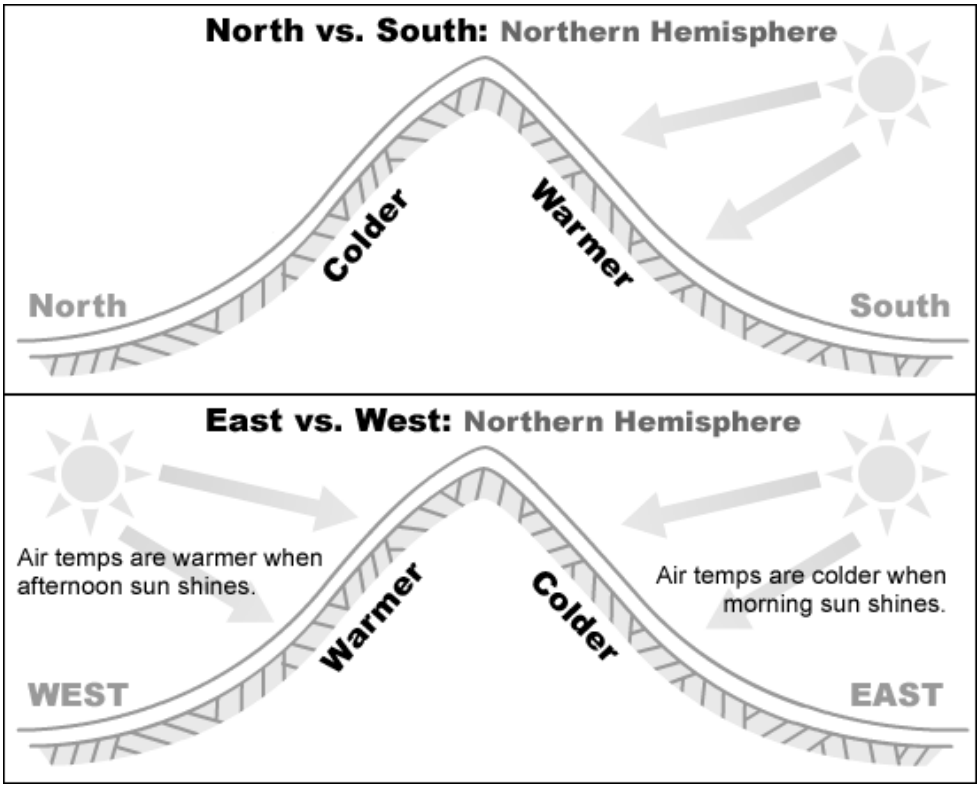

Figure 2.4 Radiation loading by aspect (From Avalanche.org).

Significantly high snowmelt rates in south facing areas relative to those with a north aspect have been confirmed in many studies ((Haupt, 1951; Anderson and West, 1965, Hendrick et al. 1971, Rowland and Moore, 1992; D'Eon, 2004; Jost et al. 2007). Work that examined both snow accumulation and melt rates in open and forested areas have found that melt rates on south facing sites with and without forest cover exceed those of clearcuts with a northern aspect (Murray and Buttle, 2003), likely due to enhanced short and long wave inputs to the snowpack.

Slope is another topographic variable that influences the radiation budget, and thus snow accumulation, of a given site. The zones with the least accumulation and/or highest melt rates (Anderson et al. 1958a) are south facing steep areas as a function of low incidence angles and thus high solar radiation inputs (Varhola, 2010). Snow accumulation is often more variable, and higher, on gentle slopes (Anderson et al. 1958a). Steep slopes display reduced accumulation as snow is transported downhill due to wind and gravity, and in the case of southern aspects, higher temperatures, which enhances melt (Golding and Swanson, 1986) and can trigger avalanches.

2.4 Scale, Wind and Topography

The accumulation of snow varies greatly at different scales in an alpine watershed within four categories: global scale, macroscale, mesoscale, and microscale (Pomeroy et al. 2002). LiDAR provides micro, meso and macroscale information on snow accumulation all within a single data layer. At the macroscale, from 10 km^2 to 1000 km^2, snow accumulation varies by latitude, elevation, orography, and lacustrine inputs. At the mesoscale, from 100 m^2 to 10 km^2, snow accumulation varies by terrain characteristics and vegetation cover. The microscale occurs from 10 m^2 to 100 m^2, where the snow accumulation and distribution differs greatly with regards to "forested and open environments because of processes of interception, sublimation and wind redistribution" (Pomeroy et al. 2002) and variable boundary layer conditions (Oke, 1988).

Variety in spatial distribution of snow depth is governed by interacting drivers, namely topography, slope, aspect, vegetation, short and long wave radiation, wind, and precipitation inputs. The dominant processes of a given snowpack are heavily dependent on local weather patterns, terrain and land cover. These features vary widely in mountainous environments, particularly in the alpine, thus their effect on snow distribution is difficult to quantify and confidently model. Multiple studies have employed variogram analyses to identify thresholds after which snow depth distribution becomes random and hard to predict. Break values in the tens of meters (15-40m) (Shook and Gray, 1994, 1996; Deems et al. 2006) are commonly observed for variogram analyses of snow depth, topography and vegetation. Arnold and Rees (2003) also observed this effect when comparing microscale plots to mesoscale plots with a 100m threshold.

Schirmer and Lehning (2010) used terrestrial laser scanning (TLS) to examine three 300m long slopes representing cross-loaded, lee side and windward terrain. Variograms of the three slopes showed lag distances of 6.4m, 13.8m and 20.4m for lee, windward and cross-loaded slopes respectively. The researchers attributed increasing scale break distances to the effect of wind speed on bare earth terrain, with high wind speed areas experiencing greater ablation. Examining changes pre and post storm revealed that increases in scale breaks on a cross-loaded slope were due to the filling of small-scale terrain features (Figure 2.5). The occurrence of the smallest break on the leeward side was attributed to lower wind speeds and the production of more, smaller wind eddies which interact with local topographic features to create a variety of microclimates and thus snow profiles.

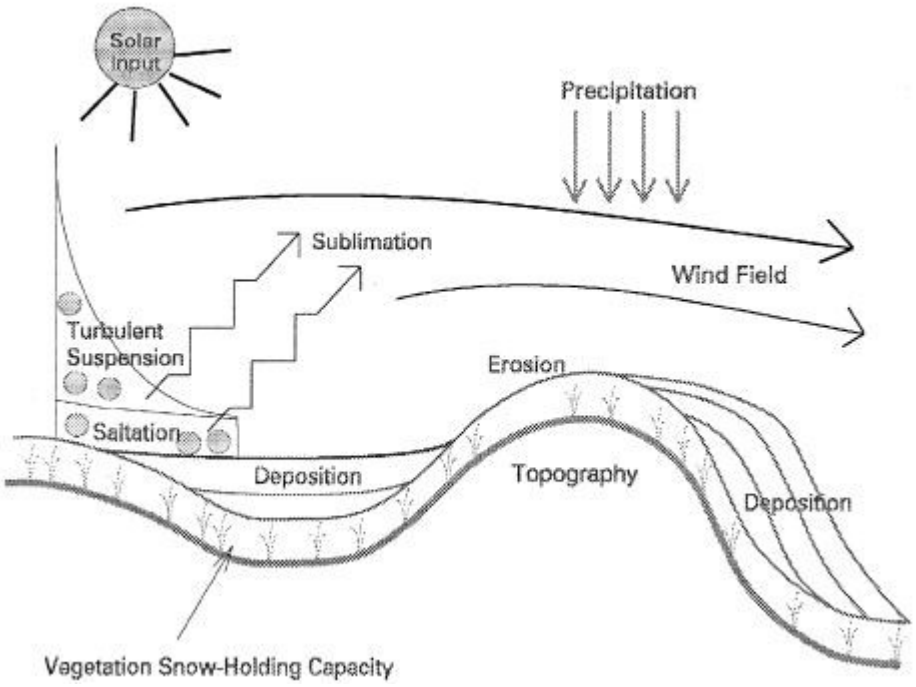

Figure 2.5 Topographic and land cover smoothing due to snow accumulation (Greene et al. 1999).

When trees are considered, a similar study found more persistent snow structure compared to an area with pronounced wind effects (Trujillo et al. 2007). The discrepancy between the two studies (areas with few wind effects as having persistent and antipersistent structure) is believed to arise from the absence and presence of vegetation, but also could be due to potential noise in the TLS data (Schirmer and Lehning, 2010). At the beginning of the accumulation period, these small terrain and land cover features have a more noticeable impact on the overall variation of snow depth within the plot, however as time goes on they are successively filled by wind redistribution and the storm events such that the surface becomes more smooth and scale breaks increase due to more homogeneity of the snowpack at the plot level (300m stretches of terrain). Surface roughness and topographic indices are more variable in montane and

alpine environments compared to flat areas, which can put the scale breaks as high as 80m in prairie environments where topography is more homogeneous (Shook and Gray, 1996). Deems et al. 2006 utilized LiDAR to identify scale breaks of 15m and 40m when analyzing snow depth, and 31 and 56m for vegetation topography (vegetation height + topography). Ultimately, this study highlighted the pronounced influence of snow redistribution by wind, which interacts with large-scale topographic orientation (Deems et al. 2006). Trujilo et al. (2007) demonstrated that while wind is an important influence of alpine snow distribution, snow depth itself is more related to breaks in topography and vegetation. Wind influenced snow patterns are dependent on topographic depressions and/or vegetation to intercept the snow that has been entrained by the wind. This observation highlights the importance of examining interacting variables, especially when the behaviour of a variable such as wind is greatly affected by terrain and land cover features that produce a variety of microclimates within a larger overall area causing seasonal snow depth distributions to exhibit a high degree of spatiotemporal variation.

3 EVALUATION OF TEMPORAL CONSISTENCY OF SNOW DEPTH DRIVERS OF A MOUNTAINOUS WATERSHED IN SOUTHERN ALBERTA

Abstract

Collecting spatially representative data over large areas is a challenge within environmental monitoring frameworks. Identifying consistent trends in a phenomenon such as snow depth enables increased sampling efficiency by minimizing field collection time and costs associated with remote sensing methods. Seasonal snowpack depth estimations during mid-winter and melt onset conditions were derived from Airborne LiDAR over the West Castle Watershed over three years. Each dataset was divided into classes with respect to five snow depth drivers: elevation, aspect, TPI, canopy cover and slope. Although mid-winter class trends for each driver were similar, mid-winter depth distributions within each class were significantly different due to the occurrence or not of recent snowfall events and snowpack redistribution and settling processes. Class trends of melt onset data were also similar to mid-winter trends but are expected to vary with seasonality. This is due to the differing stages of accumulation or ablation and the upward migration in the 0° C isotherm during spring such that depth can be declining at valley levels while still increasing at higher elevations. The observed consistency in depth driver controls and identification of zones of characteristically high and low snow storage can be used to guide future integrated snow sampling frameworks.

3.1 Introduction

Seasonal snow in headwater environments is crucial to water security in Western Canada, and anywhere with mountainous headwaters. Snow also has destructive potential when abundance aligns with certain climatic and terrestrial conditions. In 2013, Alberta experienced devastating floods due to a combination of high soil moisture and accelerated melt of a deep snowpack by a rare, but long-lasting low-pressure system which produced days of consistent rain (Fang and Pomeroy, 2016). Increased monitoring with airborne LiDAR has the potential to provide improved snow pack extent, depth and volume data to improve water supply forecasts. Understanding the spatiotemporal variation of snow depth with respect to terrain and land cover-based variables is vital to developing a snowpack monitoring framework that can be utilized in different locations and into the future. This information can help determine the timing and sampling strategy required for an airborne LiDAR based monitoring framework. For example, if a large portion of snow is found to be stored at treeline, ensuring representative sampling of such terrain would be desirable in an operational monitoring framework. Conversely, landcover and terrain type classes with an overall small percent area may translate to less emphasis on those zones in an optimized sampling strategy unless they demonstrate anomalously high or low snow depth characteristics. A strong case for expanded snowpack monitoring will be developed and explained, and a basis for evaluating subsequent modelling results will be established.

Upper alpine zones tend to contain the most exposed and steep terrain in a catchment which promotes snow redistribution by wind and gravity, therefore flatter terrain is expected to collect more snow. The effects of orographic

precipitation on snow distribution varies from windward to leeward slopes (Pomeroy and Bruin, 2001), and Roe and Baker (2006) suggest that maximum precipitation rates have high spatiotemporal variability which makes it difficult to assign a static location for maximum precipitation inputs. However, snowpack depth distributions with elevation are not solely due to orography and lapse rates but also the redistribution of snow by wind and gravity, followed by interception by vegetation (Greene et al. 1999; Hopkinson et al, 2010; 2012; Grunewald et al. 2014). Wind and gravity also promote the settlement of snow in areas of topographic transition from exposed to concave terrain (Revuelto et al. 2014b; Lopez-Moreno et al. 2014). Therefore, it is expected that peak snow depths will occur within the treeline and greater snow accumulation within topographic depressions.

Aspect is another well researched variable that impacts snow depth distributions. Short and long wave radiation provide the energy for melt which can create preferential depth losses and potential density gains in aspects with high radiative inputs versus those with little (Golding and Swanson 1986; Anderson et al. 1958a). This effect holds true at all elevations (D'Eon, 2004). Significantly high snowmelt rates in northern hemisphere south facing areas relative to those with a northerly aspect have been confirmed in many studies (Haupt, 1951; Anderson and West, 1965; Hendrick et al. 1971; Rowland and Moore, 1992; D'Eon, 2004; Jost et al. 2007). Higher mean depths should (all else being equal) occur on north facing terrain while the lowest depths should be found on south facing terrain where northern hemisphere shortwave radiative inputs are greater.

Studies of vegetation influences that examined both snow accumulation and melt rates in forested areas have found that melt rates in south facing sites with and without forest cover exceed those of clearcuts with a northerly aspect, confirming that aspect can have a greater effect on melt than forest cover (Murray and Buttle, 2003; Haupt, 1951; Anderson and West, 1965; Hendrick et al. 1971; Rowland and Moore, 1992; D'Eon, 2004; Jost et al. 2007). Vegetation, which has a lower albedo than surrounding snow, can enhance melt by absorbing solar energy, which heats up stems and foliage and then re-emits energy as longwave radiation over nearby snowpack, creating a radiation feedback as some of this reflected radiation may be reabsorbed and re-emitted by the trees. Vegetation also promotes ablation by intercepting snow and leaving it exposed on branches to wind and radiation that may drive sublimation, the transition from solid to gaseous phase without a liquid intermediate. Considering these effects, closed canopies should have the lowest depths, especially later in the season as day length and shortwave energy inputs increase. More open canopies should exhibit greater snow accumulation due to wind trapping, reduced longwave radiation and lower amounts of intercepted snow being subject to sublimation prior to falling to the ground.

3.1.1 Objective

To optimize airborne LiDAR snow depth sampling, the priority in snow depth drivers and the variation in their distributions needs to be understood. Stratifying snow depth drivers into classes is one way to detect spatial variation in snow depth due to driver attributes. The objective of this analysis is to determine if well-established snow depth distribution trends, with respect to

various terrain and land cover-based snow depth controls, exist in the headwaters of the Oldman River Basin, southern Alberta. This is assessed by examining seasonal and inter-annual consistency of class-based snow depth drivers at mid-winter and melt onset. Snow volume estimates of storage are also provided to help with sampling prioritization. LiDAR based storage values are compared to estimates using publicly available snow records, to place the Government monitoring data into a broader context at the watershed scale. Evaluating the temporal consistency of high and low snow storage zones can determine if optimized sampling needs to be adaptive with seasonality. By examining these trends and relationships at the watershed and class scales over multiple years, a foundation for future sampling requirements is established under mid-winter and melt onset snowpack conditions.

3.1.2 Study Area

The West Castle watershed (WCW) is a sub-watershed of the Oldman River Basin in Southern Alberta near the Canada – US border. Within the Oldman headwaters ~8.4% of annual runoff yield is derived from the Castle Watershed, which includes the West and South Castle rivers, (Kienzle and Mueller, 2013), making it the 3rd highest yielding sub-basin. The ~103 km^2 watershed is characterized by a large valley that runs NNW-SSE, montane forest and exposed rocky alpine summits (Figure 3.1). A snow course exists within the watershed, which is manually surveyed every month throughout the winter season.

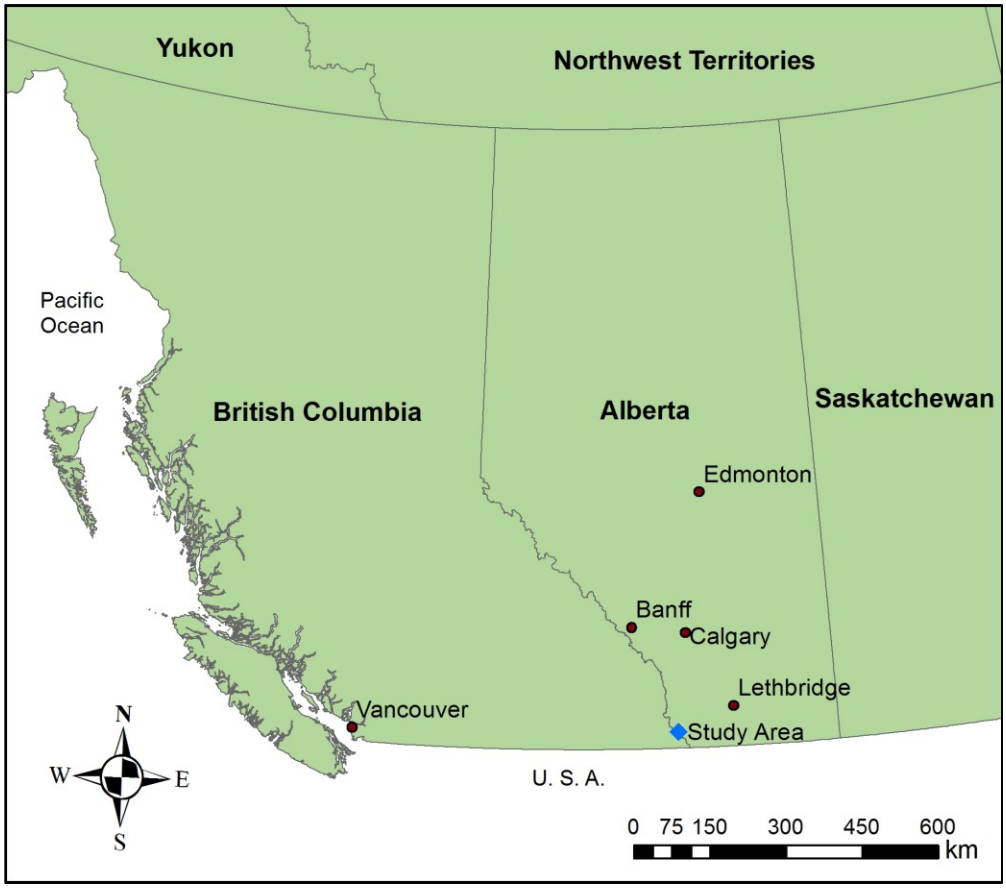

Figure 3.1 The study area is situated in southwest Alberta, along the continental divide between Alberta and BC, just north of the USA/Canada border.

3.2 Data and Methods

3.2.1 LiDAR Data

Winter flights to generate digital snow surface models (DSSMs) during snow- on conditions and another survey in summer to produce a digital elevation model (DEM) are required for LiDAR snow studies. A LSDM (LiDAR snow depth model) is created by subtracting the DEM from the DSSM (Eq 2.1, Hopkinson et al. 2004):

LSDM = DSSM – DEM (Eq 2.1)

The snow-off DEM was generated from a September 2014 survey of the same area when perennial snow was at its minimum. To produce snow-on DSSMs, airborne LiDAR surveys were conducted in February 24, 2014 (mid-winter 2014), April 7, 2016 (melt onset 2016), and February 13, 2017 (mid-winter 2017). A partial survey consisting of two flight lines was flown on March 17, 2017 (late winter 2017) when rapid weather changes interfered with plans for a fourth complete snow-on survey thus enabling a test of interpolation methods. A small validation dataset exists for mid-winter 2014, which was mostly limited to elevations <1450 m a.s.l. The root mean square error (RMSE) of this dataset is 0.27 m, which was measured using survey grade positioning of depth measurements using graduated depth probes. A snow depth validation dataset was collected along an elevation gradient in 2016 where RMSE was 0.25 m. Due to safety and logistical constraints, snow depth validation data were not collected for all surveys but all LiDAR datasets were calibrated to hard runway surfaces 35 km west of the study site at the start and end of each survey flight, so there is no reason to expect variations in depth accuracy from survey to survey. A well-calibrated LiDAR snow survey can produce a decimeter-level root mean square error, especially in zones of open canopy (Hopkinson et al. 2004; Grunewald et al. 2010; Lehning et al. 2011) but many studies report RMSE ranging up to 0.3m (Geist and Stotter, 2008; Moreno Banos et al. 2009; DeBeer and Pomeroy, 2010; Hopkinson et al. 2011; Grunewald et al. 2013). Considering the well-established accuracies of airborne LiDAR for snow depth in mountainous settings as well as the available field validation data in 2014 and 2016, these datasets were determined to be suitable for further analyses. All datasets were gridded at a 1m spatial resolution using triangulation

interpolation after being quality controlled and filtered following standard cleaning and classification methods described in Hopkinson et al (2012).

3.2.2 Snow Depth Driver Classes

Slope, aspect and TPI (Topographic Position Index) are controls that require neighbourhood functions to derive the surfaces from a DEM. To ensure that cells along the outer edge of the area of interest had a sufficient neighbourhood, they were computed on a DEM that extended past the area of interest and then clipped to the watershed extent. Neighbourhood size for slope and aspect calculations is limited to the neighbours immediately adjacent to a given cell, whereas neighbourhood settings for TPI are user defined. Using ArcGIS 10.4, slope (Eq 2.2) and aspect were calculated with the aspect and slope tools.

$$\text{Slope } (^{\circ}) = \arctan(\text{rise} / \text{run}) \qquad \text{(Eq. 2.2)}$$

TPI is calculated by comparing a cell's elevation to that of its neighbours within a user defined window (Eq 2.3; Weiss, 2001).

$$TPI = E - M \qquad \text{(Eq 2.3)}$$

Where E is the elevation of an individual grid cell and M is the mean elevation of grid cells within the window. TPI values are unique to the DEM used as an input so they are generally stratified into three classes: depressions, uplands and transitional terrain which can be flat or sloped.

Random Forest (Breiman, 2001) was utilized to rank the importance of multiple TPI surfaces generated from different window sizes and shapes. With this method, importance corresponds to a higher “Increase MSE” value (Breiman, 2001). This refers to the change in model MSE when a predictor is removed from the analysis. The more important the predictor, the greater the change in MSE in its absence. MSE values for each dataset were summed by window type (which vary in size and shape). Little was published on LiDAR snow depth models and appropriate TPI parameters prior to Lopez-Moreno et al (2017). With minimal research available regarding appropriate window sizes, study area specific sensitivity analyses can be employed to inform this decision. Too small of a window may create unnecessary detail within the three TPI classes whereas too large of a window will include complex terrain features that don’t represent the simple ridge / depression type of relief – and snow depth driver processes – that we are aiming to represent. Maximum window sizes were therefore limited to 100m. The 101m rectangular window was determined to be the most appropriate choice (Figure 3.2). The WCW’s elevation range (1390-2630 m a.s.l.), so a 100m window avoids capturing micro-scale elevation change.

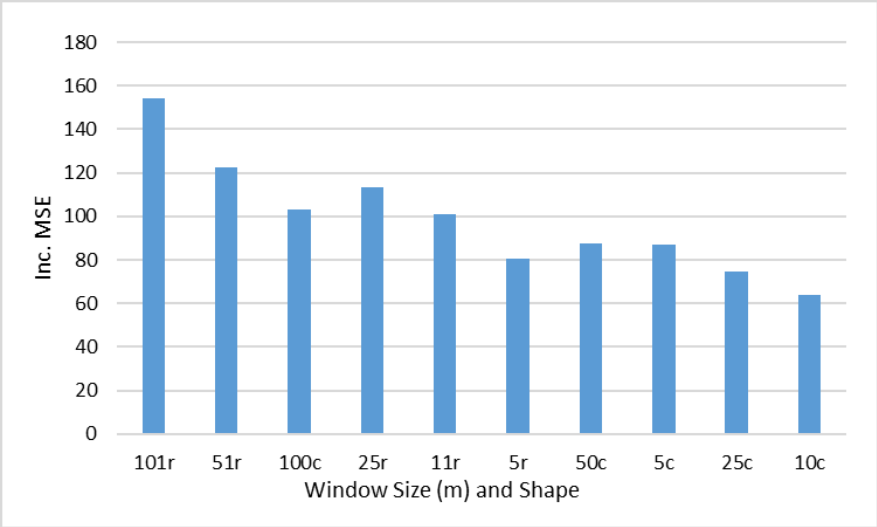

Figure 3.2 Importance of TPI surfaces.

Data collected in September 2014 was used to create a DEM, which was the input for creating snow depth control variable classes. Canopy Cover was calculated as a ratio of LiDAR canopy to total returns (Barilotti et al., 2006) (Eq 2.4).

Fractional Cover = Canopy Returns / All Returns (Eq 2.4)

Given the almost 1400 m of elevation range in the WCW and high local variation in snow depth, elevation was divided into 100m increments. The choice of a 100m elevation class was also found to be congruent with the optimal window size for the TPI window sensitivity analysis (see below). Aspect was divided into eight cardinal directions. “Flat” is a class produced by the aspect tool in ArcGIS, but it was not considered in this analysis as it contains only 7,191 data points whereas the next smallest aspect class (north) has 5,218,741 points. Exclusion of small or outlier classes (<0.5% of total area) has been

implemented in other snow studies (Grunewald et al. 2014). TPI and canopy cover were stratified using a three-class quantile approach, to ensure comparable class population sizes, while maximizing the chance of class separability. Slope also had three classes based on evenly distributed physical slope breaks at 30 degree increments instead of statistical sample breaks. As with aspect and elevation terrain classes, it was assumed that slope influences will be identified through linear, rather than statistical, stratification of physical slope attributes. The reclassified rasters were then used to create polygons for each class to use as a mask for extracting data from the LSDMs. Appendix A outlines the script locations and order for the analyses in this chapter.

3.2.3 LSDM Quality Control

LiDAR data must be carefully analyzed when making conclusions on snow depth distribution. The main sources of error with LiDAR are vertical and horizontal errors as well as incorrect classification of returns from the forest canopy, buildings or other obstructions, which may be labeled as ground returns (Hodgson, 2004). Even after cleaning raw LiDAR data, further quality control on snow depth models is required prior to analysis. It is intuitive that snow depth cannot be a negative number. Multipath (deflection of laser pulse energy from a single travel path due to multiple reflections from high reflectance surfaces) leads to low points or pits in LiDAR snow surface data that lie beneath the true surface (Hopkinson et al. 2004). This occurs because the ground to aircraft portion of the two-way travel time exceeds the time it took for the pulse to reach the ground after emission from the sensor (Hopkinson et al. 2006). Multipath is more likely to occur over a frozen snow surface than bare ground due to the

difference in reflectivity, and this is likely to produce occasional negative snow depth values in the LSDM. Such erroneous values are typically set to 0 (to mitigate their overall impact) or removed entirely (set to 'nodata'). For further information on LiDAR as a snow depth measurement tool, and for specific examples of its use in cryospheric applications, see the review papers by Deems et al. (2013) and Bhardwaj et al. (2013).

Steep slopes and horizontal DEM uncertainty can produce erroneous depth values that exceed plausible snow depths (Hodgson et al. 2005). Horizontal coregistration errors may arise where the location of a steep cliff face has shifted a few meters in the x or y direction in the DSSM compared to DEM. If horizontal coregistration is an issue, when the DEM is subtracted from the DSSM the result may be a large negative value because the elevation associated with the top of the cliff face is in a different cell between datasets. One approach to quality controlling slope induced errors is to eliminate steep areas from the datasets under analyses (e.g. Zheng et al. 2016). At the watershed scale, the total area of steep terrain and cliff faces is not a large proportion (1.8% of grid cells). To test if quality control does indeed produce more accurate snow depth measurements, pre and post quality control mean depths were presented for slope and elevation classes, because of known slope issues and the general occurrence of steeper slopes at higher elevations. It is expected depth values in steep and upper elevation classes to be most affected by quality control.

The abundance of data in a LSDM makes removal of the upper 1% of cells a statistically valid approach to limit the upper range of snow depth. Eliminating these cells entirely instead of simply thresholding implausible snow depths

excludes questionable zones from the analysis (i.e. eliminating a snow free slope with a depth of 60m snow depth instead of assigning a thresholded value to a cell where snow doesn't exist). Revuelto and others (2014b) employed a routine using the field-measured depth of a topographic depression as their maximum depth value while others (Kirchener et al. 2014; Grunewald et al. 2010) have employed knowledge of the area or canopy-based metrics to determine a maximum depth. A 99th percentile-based quality control method implicitly addresses steep slope errors and provides a statistically consistent, reasonable, and automated means of setting a limit on maximum depth. The 99th percentile was calculated after eliminating negative values, which removed ~1% of total 1m grid cells.

3.2.4 Temporal Depth Distribution Analysis

Determining how snow depth distributions and drivers vary temporally at the watershed scale supports future sampling design. For example, knowledge of snow depth driver variable spatial extent and which drivers are most important in controlling snow depth can inform driver class sampling prioritization and location. The consistency of depth distributions was first assessed at the watershed scale using Kendall's Tau correlation, given its non-parametric status and the non-normal distributions of the datasets. Snow depth distributions and storage estimates were derived at the watershed scale for two different stages in winter: mid-winter (MW) and melt onset (MO). Creating classes for each snow depth driver allowed a more refined test of mid-winter

depth distribution consistency for which an ANOVA (analysis of variance) was employed at a 95% confidence level.

Two mid-winter datasets were collected as well as one captured at late winter (LW), March 2017, which covers two flightlines (FL) as opposed to the whole watershed (Figure 3.3). From the snow pillow records at Gardiner Creek (AEP, 2017), MW 2014 and LW 2017 demonstrated similar conditions on the survey dates: a recent accumulation of fresh snow (Figure 3.4). Mean air temperature prior to these surveys was similar but it was warmer on the day of MW 2017 data collection (Table 3.1). SWE at Gardiner Creek was highest at the time of the LW 2017 survey but air temperatures were similar to MW 2017 (Table 3.1). Melt is evident prior to the 2016 survey and was observed during validation data collection, which corresponded to declining SWE at Gardiner Creek (Figure 3.4) and air temperatures were near or above zero (Table 3.1). This dataset is therefore considered to represent melt onset (MO) conditions.

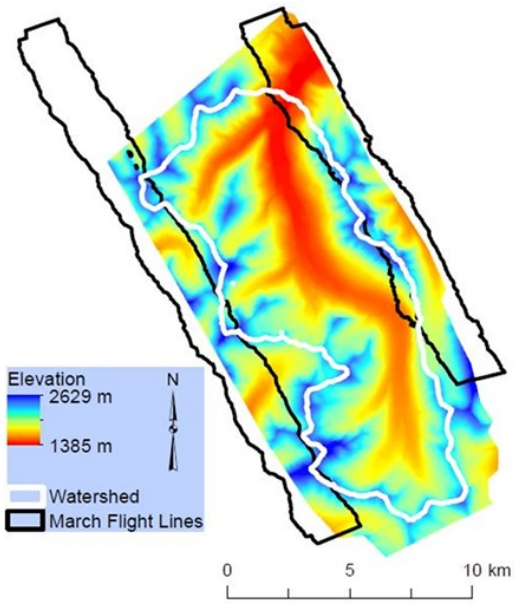

Figure 3.3 The data extents utilized in this study. Except for the March 2017 flight lines, the DEM extent represents the area covered for all winter LiDAR surveys.

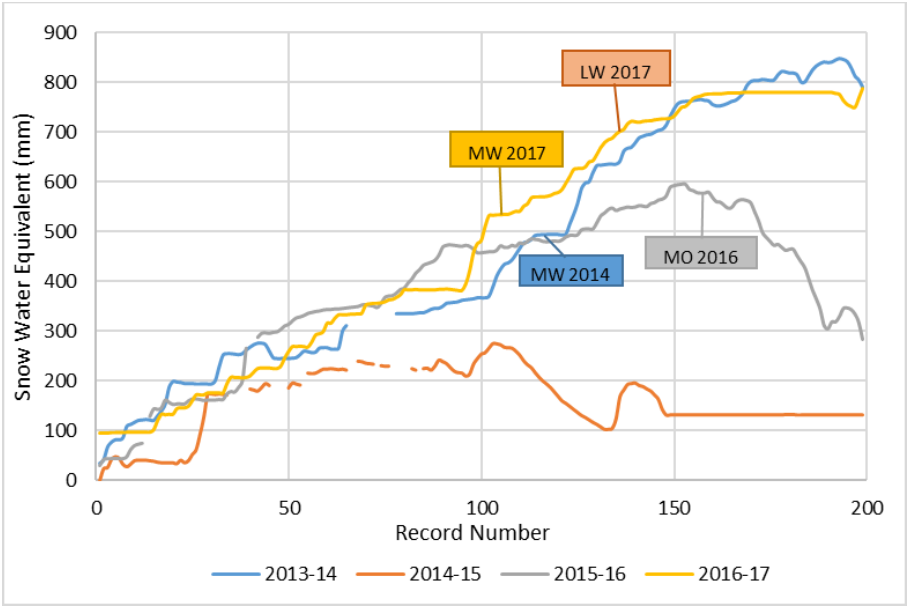

Figure 3.4 Daily SWE from the Gardiner Creek site (1924 m a.s.l.), for 200 days starting on November 1st through to mid-May of the following year (AEP, 2017).

Table 3.1 Air temperature at Gardiner Creek prior to and on the day of the LiDAR surveys (AEP, 2017).

Dataset	Air Temperature (˚C)	
	30 days prior	Day of survey
February 24 2014	-13.9	-12.9
April 7 2016	-1.9	6.16
February 13 2017	-10.6	-3.23
March 17 2017	-7.4	-3.76

The snow depth driver class polygons generated in section 3.2.2 were inputs to a python script (Cartwright, 2018), which provides a text file containing raster summary statistics. Storage is derived by multiplying mean depth of each control class by the number of grid cells (excluding 'nodata' values) in that class. To detect statistically significant differences between depth distributions at mid-winter, ANOVAs were implemented on a sample of each class' populations at a 95% confidence level. Sample n values of each snow depth driver class were as follows: elevation 10,000, aspect 25,000, canopy cover 30,000, TPI 30,000 and slope 30,000. A challenge of working with high resolution data is that there are many data points and more research is needed regarding how to representatively sample classes of varying sizes (i.e. achieve the same level of class representation for predictors with 9 classes versus a different predictor with 3 classes). These effects should be considered in interpretation of the results along with summary statistics such as class means and standard deviations.

3.2.5 Government Snow Monitoring Comparison

A government hydrometeorological station (Gardiner Creek) which measures snow depth, water equivalent and other meteorological variables is located just beyond the northern edge of the watershed boundary, making it the closest Government snowpack monitoring site to the WCW with automated, daily data collection. Depth is measured with a SR-50 and SWE data is collected with a snow pillow. A snow course exists within the watershed as well, which is manually surveyed every 4-5 weeks, depending on weather conditions, throughout the winter season. To give context for how the publicly available snow data relates to watershed snow conditions as determined with LiDAR, average watershed scale snow depth from the LiDAR data was compared to the SR-50 average daily snow depth value at Gardiner Creek (GC). Mean depth derived from LiDAR data as well as the SR-50 depth on the survey date were both multiplied by the number of grid cells (excluding no data points) in each LSDM to calculate volume. Average depth from the LSDM was calculated by finding the mean of all cells (excluding no data points). The depth at the West Castle snow course (WCSC) was also used to estimate volume, although the only date on which snow course measurements coincided with a LiDAR survey was for the first mid-winter LSDM. For the other survey dates, the WCSC was surveyed 2+ weeks after the LiDAR surveys. This comparison is intended to explore how snow volume estimates from the existing monitoring network data compare to LiDAR based estimates for the WCW. Integrating LiDAR and continuous, automated data streams has the potential to enhance monitoring networks by increasing temporal and spatial availability of data in highly complex landscapes. It is likely that multiple weather stations in strategically

chosen locations a priori to commencement of a study would be necessary to observe consistent spatiotemporal trends. Maintenance of the weather station is costly, especially with many locations requiring helicopters for access, and there may be an argument to put funds towards LiDAR monitoring if a large difference in snow volume between the publicly available and remote sensed data exists.

3.3 Results

3.3.1 Snow Depth Driver Classes

Watershed land surface classifications derived from the lidar data produced 14 elevation classes, 8 aspect classes, and 3 classes for slope, TPI and canopy cover (Figure 3.5).

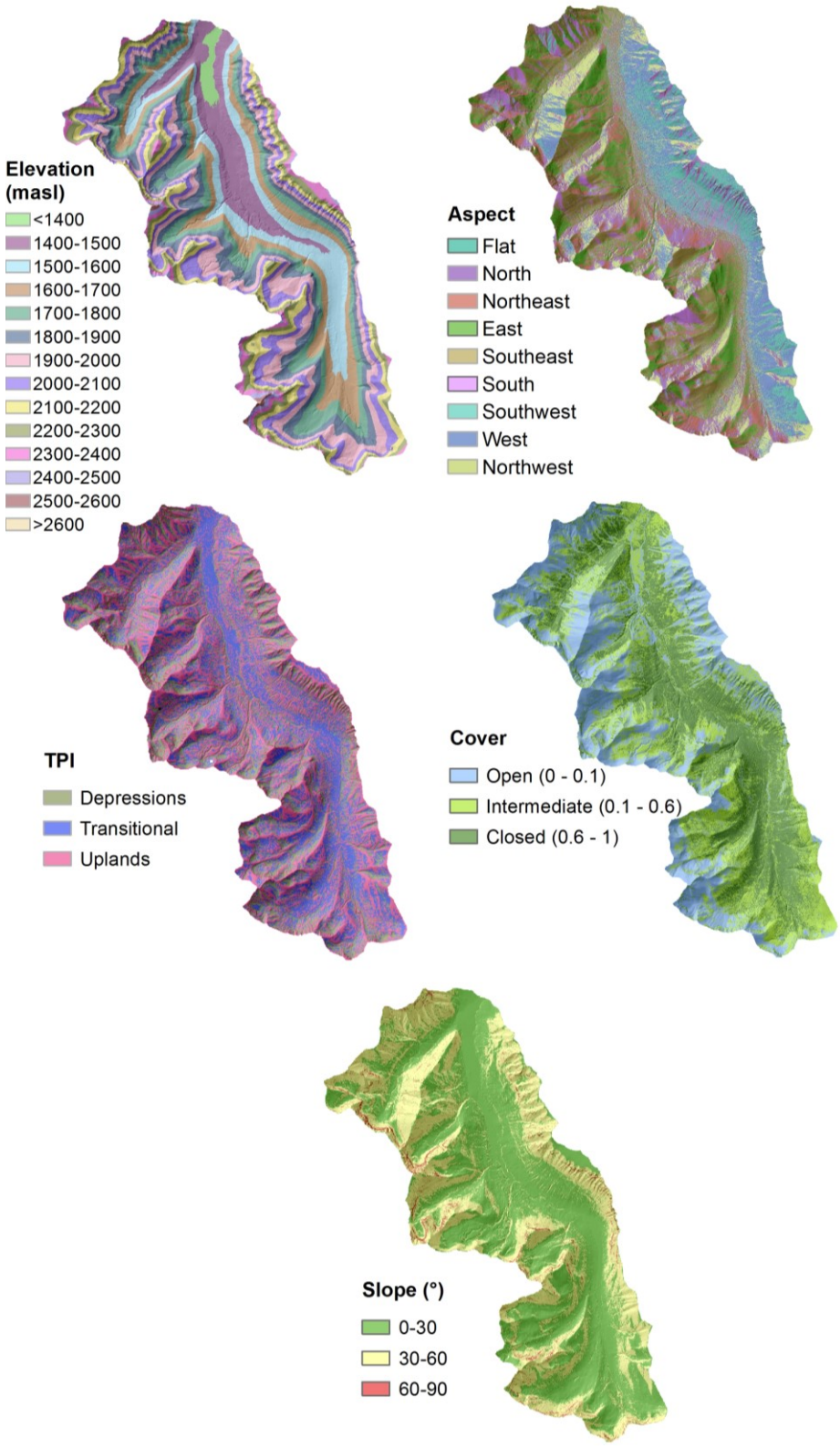

Figure 3.5 Snow depth driver classes.

3.3.2 LSDM Quality Control

Quality control (QC) reduced watershed scale depth by 0.1m for MW 2014 and MO 2016 due to the removal of extreme depth outliers while MW 2017's mean depth was unchanged (Table 3.2). Mid-winter 2017 elevation means were the least affected by QC, whereas half of the melt-onset elevation class means changed by >0.1m (Table 3.3). Larger differences generally occur at upper elevations and means below 1800m were relatively unaffected (<0.05m change in mean) for all years. As slope increased, so did the magnitude of difference between pre- and post- quality control datasets (Table 3.3). Slopes >60˚ make up only 1.8% of the watershed (Table 3.4). Although depth changes were larger in steeper and higher terrain (Table 3.3), QC in this analysis only has a small effect on mean depth at the watershed scale, as shown by Table 3.2. After QC, elevation and aspect class area distributions remained within 0.1% of the 2014 DEM areas (Table 3.4). However, slope, TPI and canopy cover class area distributions differed by up to 0.6% relative to the DEM following QC.

Table 3.2 Mean depth at the watershed scale prior to and following QC.

Dataset	Snow Depth (m)	
	No QC	99th percentile clip
MW 2014	1.6	1.5
MW 2017	1.5	1.5
MO 2016	1.3	1.2

Table 3.3 Mean snow depth change as a result of quality control.

Snow Depth Driver Class		Δ Depth (m)		
		Mid-winter 2014	Mid-winter 2017	Melt Onset 2016
Elevation (m a.s.l.)	<1400	0.00	0.00	0.00
	1400-1500	0.00	0.00	0.00
	1500-1600	0.00	0.00	0.00
	1600-1700	0.00	0.00	0.00
	1700-1800	0.01	0.00	0.01
	1800-1900	0.06	0.01	0.06
	1900-2000	0.11	0.03	0.12
	2000-2100	0.09	0.02	0.12
	2100-2200	0.13	0.05	0.20
	2200-2300	0.07	0.03	0.11
	2300-2400	0.08	0.03	0.12
	2400-2500	0.09	0.03	0.17
	2500-2600	0.13	0.04	0.23
	>2600	0.01	0.00	0.03
Slope (˚)	0-15	0.01	0.00	0.01
	15-30	0.03	0.01	0.04
	30-45	0.08	0.02	0.10
	45-60	0.13	0.05	0.18
	60-75	0.20	0.17	0.31
	75-90	0.85	0.95	1.35

Table 3.4 Percent area of snow depth control classes after quality control.

Snow Depth Driver Class		Fall 2014 (DEM)	Mid-winter 2014	Mid-winter 2017	Melt Onset 2016
Elevation (m a.s.l.)	<1400	1.3	1.3	1.3	1.3
	1400-1500	9.5	9.6	9.5	9.6
	1500-1600	11.5	11.6	11.5	11.6
	1600-1700	11.3	11.4	11.3	11.4
	1700-1800	11.7	11.8	11.7	11.8
	1800-1900	11.6	11.6	11.6	11.6
	1900-2000	12.6	12.5	12.6	12.5
	2000-2100	11.3	11.2	11.3	11.2
	2100-2200	8.7	8.6	8.7	8.6
	2200-2300	6.0	6.0	6.0	6.0
	2300-2400	3.4	3.4	3.4	3.3
	2400-2500	0.9	0.9	0.9	0.9
	2500-2600	0.2	0.2	0.2	0.2
	>2600	0.0	0.0	0.0	0.0
Aspect	Flat	0.0	0.0	0.0	0.0
	N	10.0	9.8	9.9	9.8
	NE	12.7	12.5	12.7	12.5
	E	17.7	17.6	17.7	17.6
	SE	15.4	15.4	15.4	15.4
	S	7.9	7.9	7.9	7.9
	SW	11.5	11.6	11.6	11.6
	W	14.1	14.2	14.1	14.2
	NW	10.8	10.8	10.8	10.8
Slope	0-30	57.1	57.4	57.1	57.4
	30-60	41.2	40.9	41.1	40.9
	60-90	1.8	1.7	1.7	1.7
TPI	Depressions	33.1	32.5	33.0	32.6
	Transitional	33.8	34.1	33.9	34.1
	Uplands	33.1	33.4	33.1	33.4
Cover	Open (0 - 0.13)	33.2	32.7	33.1	32.7
	Intermediate (0.13 - 0.57)	33.1	33.3	33.1	33.3
	Closed (0.57 - 1)	33.7	34.1	33.8	34.0

3.3.3 Snow Depth Distributions

3.3.3.1 Quality Controlled Outputs and Watershed Scale Correlations

Following QC, depth distributions for MW 2014 and 2017 have more in common than MO 2016, which appears to display a more distinct transition from low to high depths around valley sides (Figure 3.6). This is confirmed with statistically significant ($p < 0.05$) Kendall's tau results indicating a higher correlation between mid-winter datasets relative to the mid-winter and melt onset LSDM comparisons (Table 3.5). Of the two mid-winter datasets, 2014 has a slightly higher correlation with melt onset 2016. No correlation was possible for LW 2017 due to the limited spatial overlap between the flightlines and watershed extent (Figures 3.3 and 3.6).

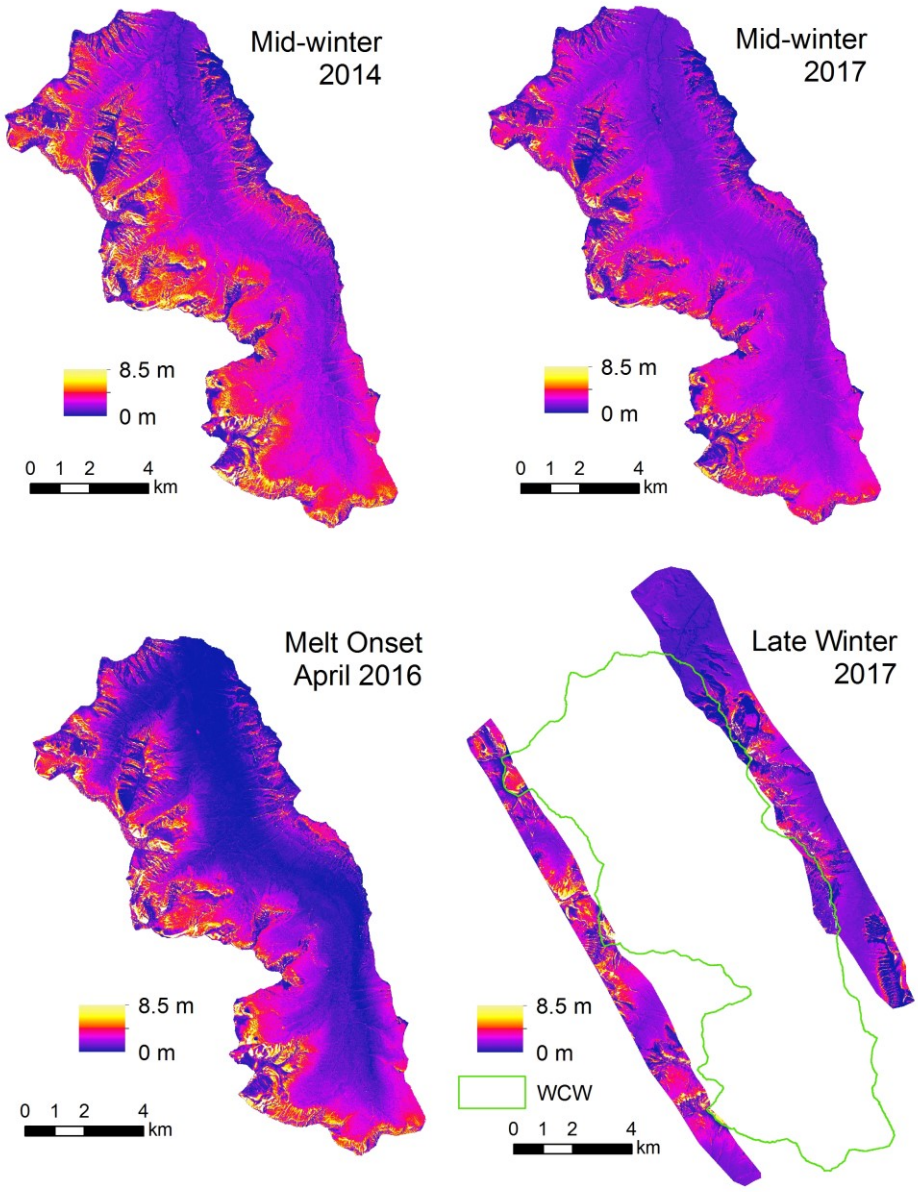

Figure 3.6 LSDMs within the WCW extent.

Table 3.5 Correlation results between WCW snow depth distributions (n = 50,000).

Datasets	Kendall's tau	p-value
MW 2014 vs LW 2017	0.76	*
MW 2014 vs MO 2016	0.68	*
MO 2016 vs LW 2017	0.65	*

3.3.3.2 Elevation

For each year of data, class-based depth distributions increased with elevation gains for the first half of the elevation range (Figure 3.7). Peak depth mean snow depths occur in the 1800-2200 m a.s.l. range (Table 3.6). Although the exact elevation of peak depths varied, they were within the treeline ecotone for all years. Above 2100 m a.s.l., mid-winter 2014 and mid-winter 2017 depths declined more sharply than late winter 2017 and melt onset 2016. Mid-winter 2014 is the only dataset in which a consistent decline in depth after the peak is observed (Figure 3.7). Both of the mid-winter depth distributions are significantly different except for the uppermost elevation class (Table 3.6). Most of these results align with the magnitudes of standard deviations observed in each class for both datasets, except for elevations >2600 which, due to comparatively small class size over extended distances, has a large standard deviation value resulting in no significant difference between the two datasets.

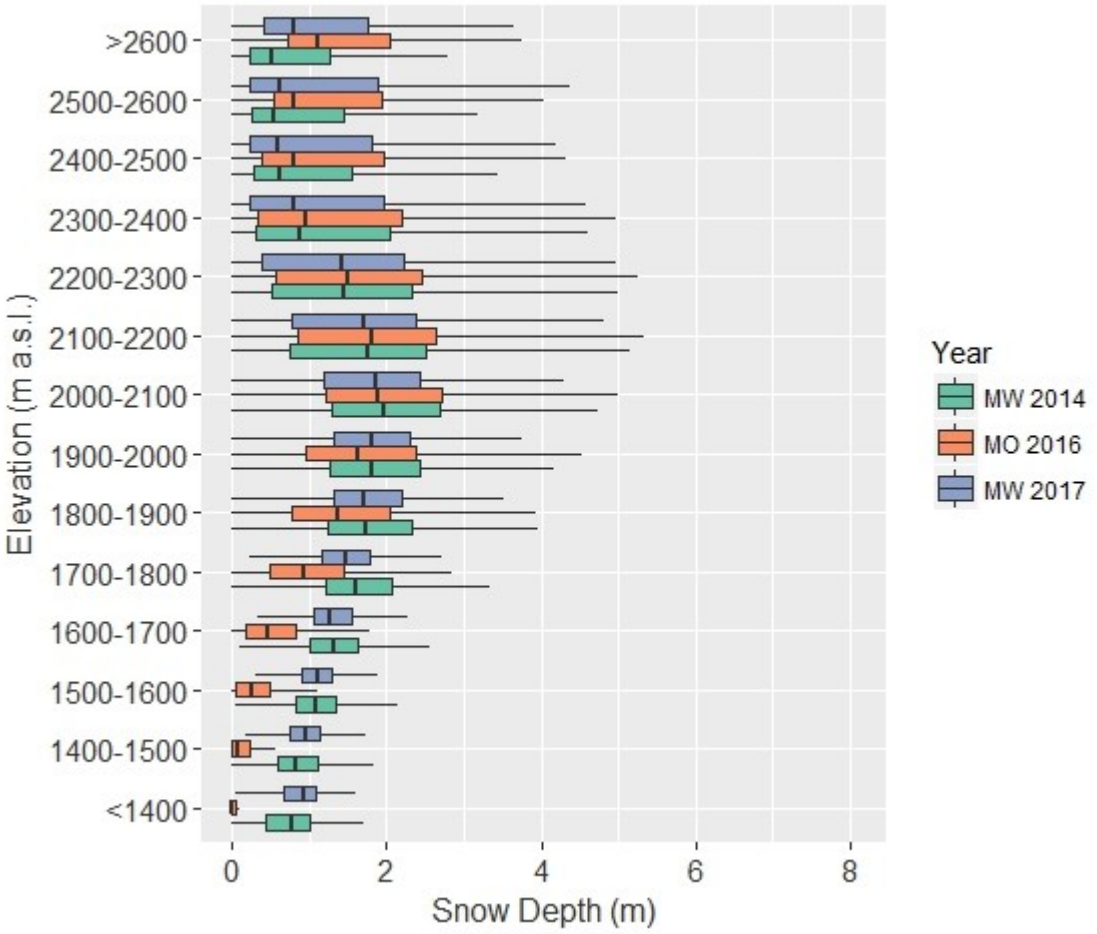

Figure 3.7 Mean snow depth of each elevation class for watershed scale datasets.

*Table 3.6 Elevation class depth summary. * = significant difference ($p < 0.05$). N = 10,000 for all classes except >2600 m a.s.l. where N = 1,200.*

Elevation Class (m a.s.l.)	Mean Depth (m)				Standard Deviation (m)				ANOVA Significant Difference
	MW 2014	MW 2017	LW 2017	MO 2016	MW 2014	MW 2017	LW 2017	MO 2016	
<1400	0.7	0.8	0.6	0.0	0.4	0.3	0.3	0.1	*
1400-1500	0.9	1.0	0.7	0.2	0.4	0.3	0.2	0.2	*
1500-1600	1.1	1.1	0.9	0.3	0.4	0.4	0.6	0.3	*
1600-1700	1.4	1.3	1.1	0.6	0.6	0.4	0.5	0.5	*
1700-1800	1.7	1.6	1.3	1.0	0.7	0.6	0.6	0.8	*
1800-1900	1.9	1.8	1.6	1.5	0.9	0.9	1.0	1.1	*
1900-2000	1.9	1.9	1.9	1.8	1.0	1.1	1.3	1.2	*
2000-2100	1.9	1.9	2.1	2.0	1.1	1.1	1.5	1.2	*
2100-2200	1.8	1.8	1.8	1.9	1.1	1.3	1.4	1.3	*
2200-2300	1.5	1.5	1.7	1.7	1.1	1.2	1.4	1.3	*
2300-2400	1.2	1.2	1.6	1.4	1.2	1.3	1.5	1.4	*
2400-2500	1.1	1.2	1.4	1.3	1.2	1.4	1.5	1.4	*
2500-2600	1.0	1.3	1.7	1.4	1.1	1.5	1.7	1.4	*
>2600	1.0	1.1	1.3	1.6	1.1	1.0	1.1	1.2	

Maximum storage was in the 1900-2000 m a.s.l. elevation class at mid-winter in both years (Table 3.7). At melt onset, maximum storage shifted upwards to the 2000-2100 m a.s.l. class. The uppermost elevation classes (>2500 m a.s.l.) stored the least amount of snow for all years of data. For all datasets, the majority of snow is stored in elevation bands approximating the treeline ecotone (1800-2100 m a.s.l.).

Table 3.7 Storage of elevation classes at the WCW scale.

Elevation Class (m a.s.l.)	Storage (x 10^6 m^3)		
	MW 2014	MW 2017	MO 2016
<1400	1.0	1.1	0.0
1400-1500	8.5	9.5	1.5
1500-1600	13.3	13.4	3.8
1600-1700	15.8	15.1	6.7
1700-1800	20.0	18.7	12.4
1800-1900	22.5	21.8	18.5
1900-2000	24.6	24.2	23.1
2000-2100	22.6	22.2	23.2
2100-2200	15.8	15.9	17.0
2200-2300	9.4	9.3	10.4
2300-2400	4.3	4.2	5.0
2400-2500	1.0	1.2	1.3
2500-2600	0.3	0.3	0.4
>2600	0.0	0.0	0.0

3.3.3.3 Aspect

West facing terrain had the lowest mean depth for all years (Table 3.8), although the median depth value of the southwest class was slightly lower in 2016 (Figure 3.8). The lowest zones of storage were south and south west zones (Table 3.9). The deepest mean snow depth occurred on north and northeast facing terrain (Table 3.8), although the east class had the highest storage value (Table 3.9). All classes had significantly different mid-winter depth distributions.

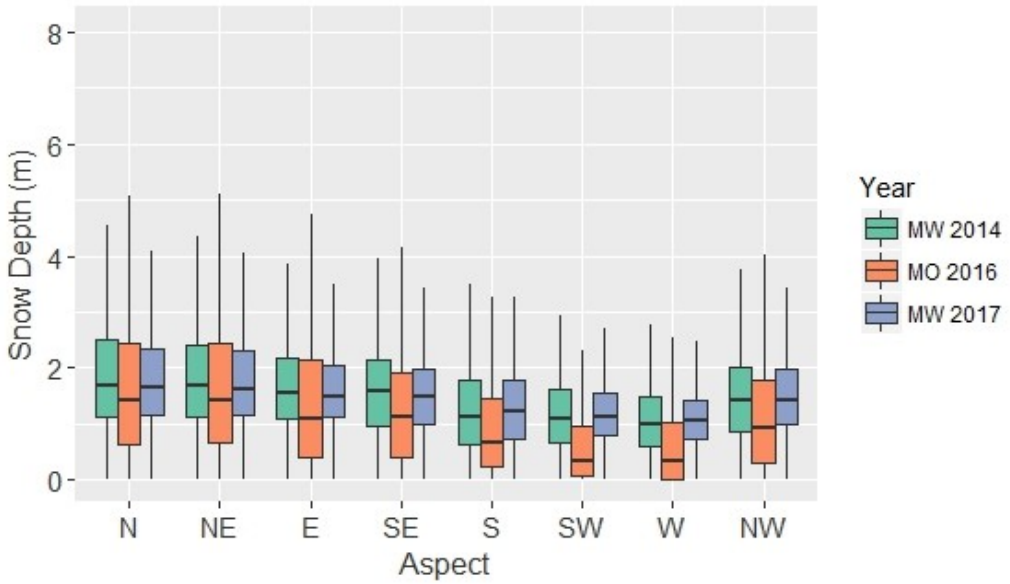

Figure 3.8 Boxplots of aspect class depth distributions.

*Table 3.8 Aspect class depth summary. * = significant difference (p < 0.05). N = 20,000.*

Aspect Class	Mean Depth (m)				Standard Deviation (m)				ANOVA
	MW 2014	MW 2017	LW 2017	MO 2016	MW 2014	MW 2017	LW 2017	MO 2016	
N	1.9	1.9	1.9	1.7	1.0	1.1	1.4	1.3	*
NE	1.9	1.9	1.9	1.7	1.0	1.1	1.3	1.4	*
E	1.7	1.7	1.6	1.4	0.9	1.0	1.2	1.2	*
SE	1.6	1.6	1.5	1.3	0.9	0.9	1.1	1.1	*
S	1.3	1.3	1.5	0.9	0.8	0.8	1.0	1.0	*
SW	1.2	1.2	1.2	0.7	0.7	0.7	1.0	0.8	*
W	1.1	1.1	1.1	0.6	0.8	0.7	1.1	0.8	*
NW	1.5	1.5	1.5	1.2	1.0	0.9	1.3	1.1	*

Table 3.9 Snow depth storage by aspect class.

Aspect Class	Storage (x10^6 m^3)		
	MW 2014	MW 2017	MO 2016
N	19.4	19.2	17.3
NE	24.9	24.4	21.6
E	31.6	30.8	25.3
SE	25.6	24.6	20.7
S	10.6	10.6	7.7
SW	13.9	14.1	7.9
W	15.9	16.3	9.3
NW	17.1	16.8	13.3

3.3.3.4 Topographic Position Index

Across all datasets, topographic depressions demonstrate the highest mean snow depths, and as terrain transitions from depressions to uplands, depths tend to decrease (Table 3.10). Transitional terrain classes have the most variability across all datasets (Figure 3.9). Storage values follow the same pattern (Table 3.11), with greater snowpack storage in depressions. All mid-winter TPI class depth distributions were significantly different.

*Table 3.10 TPI class depth summary. * = significant difference ($p < 0.05$). N = 30,000.*

TPI Class	Mean Depth (m)				Standard Deviation (m)				ANOVA
	MW 14	MW 17	LW 17	MO 16	MW 14	MW 17	LW 17	MO 16	
Depressions	1.9	1.9	2.0	1.6	1.2	1.3	1.5	1.4	*
Transitional	1.5	1.4	1.5	1.0	0.8	0.7	1.1	1.0	*
Uplands	1.4	1.3	1.2	1.1	0.5	0.5	1.0	0.6	*

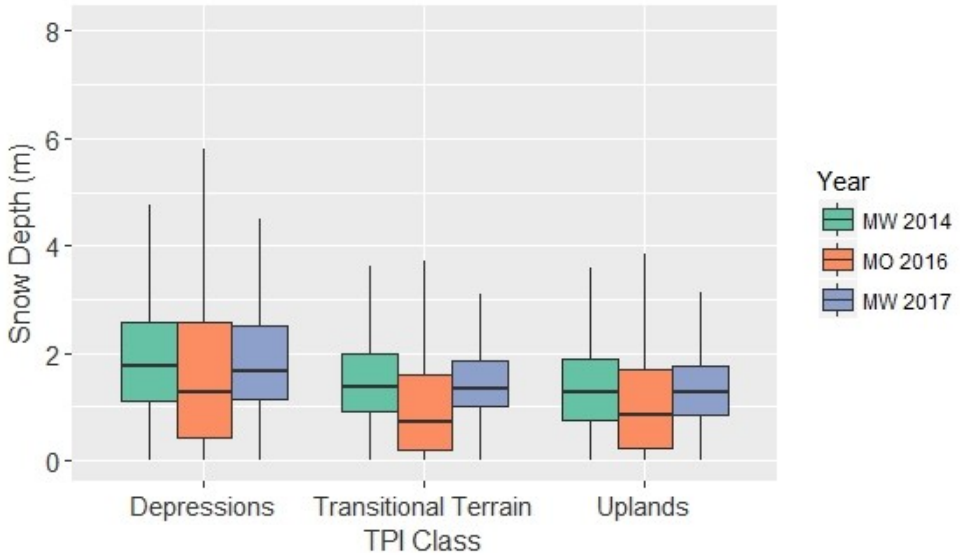

Figure 3.9 TPI class depth distribution boxplots.

Table 3.11 Snowpack storage volume within TPI classes.

TPI Class	Storage (x10^6 m^3)		
	MW 2014	MW 2017	MO 2016
Depressions	64.8	64.8	54.6
Transitional	52.2	48.8	34.8
Uplands	47.7	44.3	37.5

3.3.3.5 Canopy Cover

In mid-winter, peak mean snow depth occurs in the intermediate (0.13 - 0.57 fractional canopy cover) canopy cover class with the lowest mean depth under closed canopy (Table 3.12). With the melt onset 2016 data, mean depth decreases as canopy cover increases. Open canopy class values were the same for all years. Each canopy cover class yielded a significant difference via ANOVA. Storage was highest in areas of intermediate cover (Table 3.13), and open areas

stored much less snow than intermediate and closed canopies in mid-winter. At melt onset, closed canopies displayed lower storage volumes than open areas.

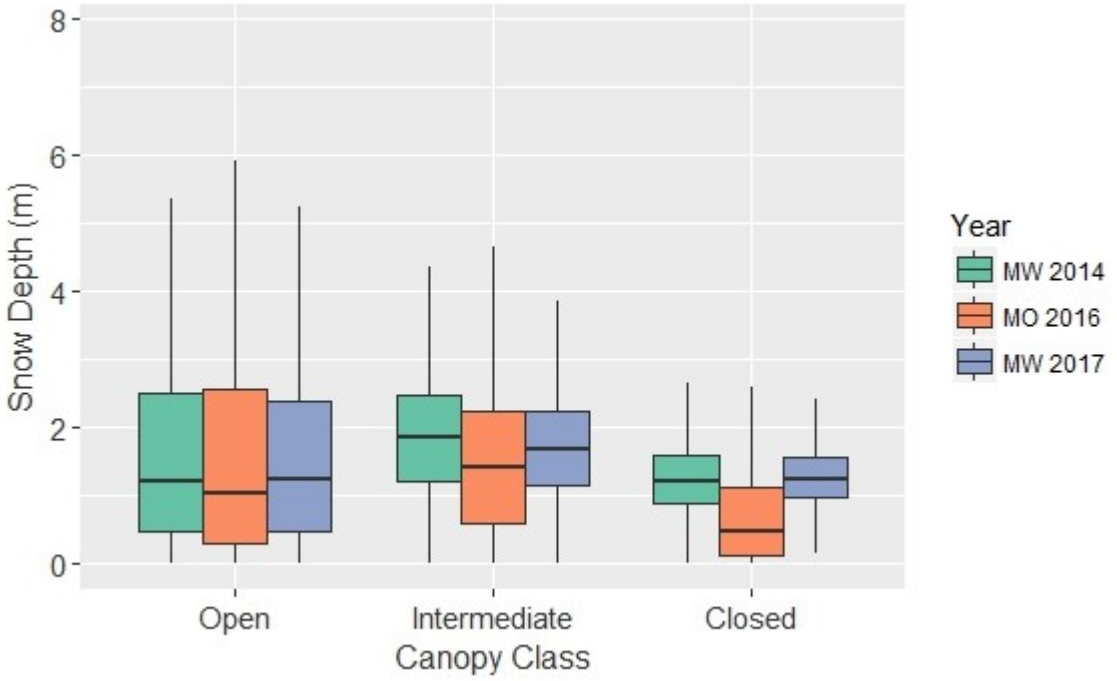

Figure 3.10 Canopy cover depth distributions.

*Table 3.12 Canopy cover depth summary. * = significant difference ($p < 0.05$). N = 30,000.*

Cover Class	Mean Depth (m)				Standard Deviation (m)				AN-OVA
	MW 2014	MW 2017	LW 2017	MO 2016	MW 2014	MW 2017	LW 2017	MO 2016	
Open	1.6	1.6	1.7	1.6	1.2	1.3	1.6	1.4	*
Interm-ediate	1.8	1.7	1.7	1.4	0.8	0.7	1.2	1.0	*
Closed	1.2	1.2	1.1	0.6	0.5	0.5	0.6	0.6	*

Table 3.13 Cover class storage.

Cover Class	Storage (x10^6 m^3)		
	MW 2014	MW 2017	MO 2016
Open	52.0	54.1	53.1
Intermediate	62.4	58.7	52.0
Closed	44.2	44.9	24.3

3.3.3.6 Slope

Patterns of mean snow depth with slope through time are not as consistent as other driver variables. Mean snow depth was lowest above 60˚ in 2014, with intermediate sloped terrain (30-60˚) displaying the greatest depth (Table 3.14). Intermediate terrain also contains the greatest mean depth in late winter 2017. The 60-90˚ depth distributions appear to contain the most variability regardless of season or year. The mid-winter 2014 and 2017 snow depth distributions are statistically different. At melt onset, a positive relationship between slope class and mean depth exists (Table 3.14). Despite this, steep slopes stored the least amount of snow at melt onset and mid-winter due to these classes representing a small (~1.8%) portion of the watershed (Table 3.4). Flatter terrain stored the most snow for all time periods sampled (Table 3.15) but the difference from the 0-30˚ and 30-60˚ classes is less pronounced at melt onset compared to mid-winter.

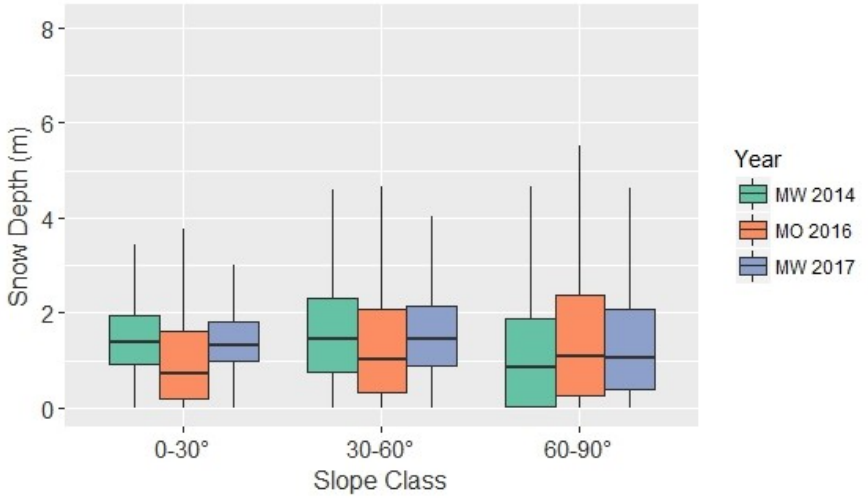

Figure 3.11 Slope class depth distributions.

*Table 3.14 Slope class depth summary. * = significant difference (p < 0.05). N = 30,000.*

Slope Class	Mean Depth (m)				Standard Deviation (m)				ANOVA
	MW 2014	MW 2017	LW 2017	MO 2016	MW 2014	MW 2017	LW 2017	MO 2016	
0-30˚	1.5	1.5	1.4	1.0	1.5	0.8	1.1	1.0	*
30-60˚	1.6	1.6	1.6	1.4	1.6	1.1	1.3	1.4	*
60-90˚	1.2	1.5	1.7	1.5	1.2	1.5	1.6	1.5	*

Table 3.15 Slope class storage.

Slope Class	Storage ($x10^6$ m^3)		
	MW 2014	MW 2017	MO 2016
0-30	87.5	86.3	60.9
30-60	69.3	67.9	59.4
60-90	2.1	2.8	2.6

3.3.4 Public and LiDAR Snow Data Comparison

The lowest depths from all three data sources (GC, WCSC, WCW LiDAR) was during MO 2016 (Table 3.17). It should be noted that there is a temporal disparity between the WCSC and LiDAR as the snow course was visited 2-3 weeks after from the 2017 and 2016 LiDAR surveys. GC depths were consistently higher than the LiDAR derived watershed mean while the WCSC were consistently less than LiDAR means. As a result, GC consistently overestimated storage, whereas the WCSC consistently underestimated storage (Table 3.16).

Table 3.16 Mean depths and volumes of LiDAR and provincial snow data.

LiDAR Survey Date	Mean Depth (m)			Volume ($x10^8$ m^3)		
	GC	WCSC	LiDAR	GC	WCSC	LiDAR
Mid-winter 2014	1.91	1.15	1.53	1.95	1.17	1.56
Mid-winter 2017	1.90	1.18	1.51	1.94	1.20	1.54
Melt Onset 2016	1.70	0.54	1.18	1.73	0.55	1.20

3.4 Discussion

3.4.1 Snow Depth Drivers and Quality Control

The only predictor variable with user-defined calculation parameters was TPI, and it is unsurprising that the largest window was the best predictor. Our TPI results corroborate findings elsewhere (Lopez-Moreno et al. 2014). Spanish snow research has found that study areas with larger elevation ranges have snow depths that correlate best with larger TPI windows (Lopez-Moreno et al. 2014). Revuelto et al. 2014b used a 25m window in a small Spanish Pyrenees catchment with a 400m elevation range whereas a more recent Spanish study

(Lopez-Moreno et al. 2017) over a ~1361m elevation range found that a 200m TPI window produced the best correlations with snow depth, suggesting that as study area and terrain complexity increase, so should window size. It was appropriate based on their observations and exploratory analyses to use a 101m rectangular window for the TPI surface in our simulations for a watershed with an elevation range of 1238m.

For all datasets, the zones excluded by quality control are associated with steeply sloped terrain (Figure 3.12). LSDM extreme outlier values in cliff edge areas ranged from 62-110m prior to quality control. The bigger post-QC change in depths at both upper elevations and in the 60-90˚ slope class is due to a concentration of high snow depth values where steep slopes exist along with small class sizes associated with single grid cells or small patches.

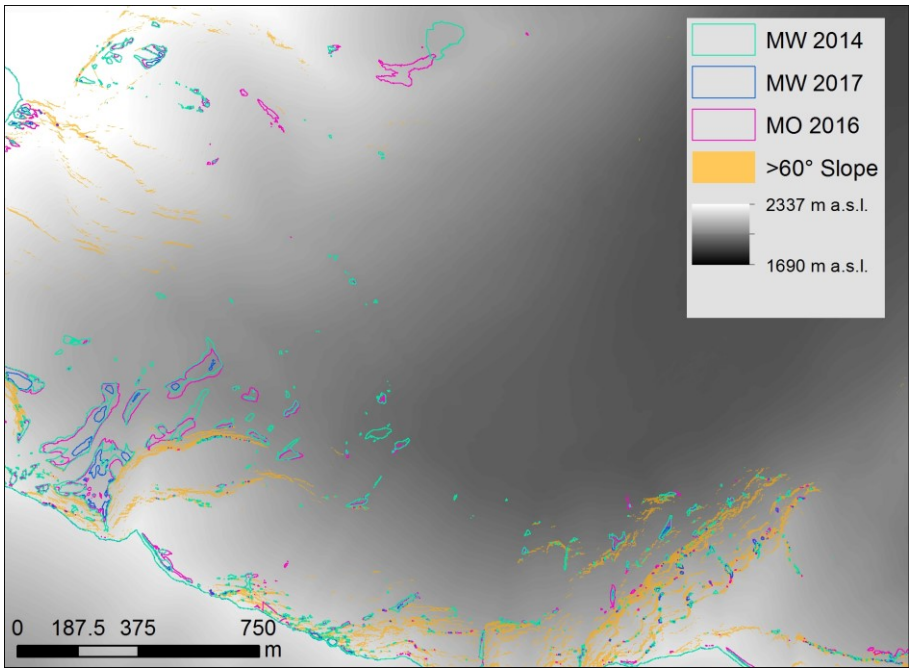

Figure 3.12 Zones removed by 99th percentile quality control (teal, blue or pink outlines) were often consistent between datasets and occurred near steep slopes.

At the watershed scale, mean snow depth was either unchanged or adjusted downwards by ~0.1m (Table 3.2) using the 99th percentile approach. Effects of co-registration errors and steep slopes on mean snow depths are pronounced when the data are broken down into classes (Table 3.3). This suggests that co-registration errors do not affect the mean snow depth at the watershed scale but erroneous depth values may complicate class-based analyses or modelling workflows. Thresholding values using either the 99th percentile or 10m (an arbitrary but reasonable upper depth value) produces values within 0.1m of the pre- or post-QC data. Snow depth values to use as an upper threshold have been determined with field-based measurements of known depressions and zones of wind loading during the season of LiDAR data

collection in other studies (Revuelto et al. 2014). Overall, any of these methods could be suitable if depth means are to be used in further analyses. While this QC workflow does not substantively alter watershed scale depths, further localised analyses in high elevation or steep sloped areas may require dedicated QC. For example, a detailed slope-based assessment of upper elevation snowpack on north-facing aspects might require additional QC, as these areas are generally the last to melt and are important in understanding late season runoff. Ultimately there are several valid QC methods for LiDAR data, and exclusion of the upper 1% of data was desirable in lieu of field-based measurements to determine reasonable thresholds for each year, and to remove extreme values for subsequent modelling analyses.

Table 3.17 Mean snow depth sensitivity to QC method.

Dataset	Mean WCW Snow Depth (m)				
	No QC	99th percentile elimination	60˚ Slope elimination	Threshold 99th percentile	Threshold 10m
MW 2014	1.6	1.5	1.6	1.6	1.6
MW 2017	1.5	1.5	1.5	1.5	1.5
MO 2016	1.3	1.2	1.3	1.2	1.3

3.4.2 Inter-annual Depth Distributions

3.4.2.1 Elevation

Elevation is a well-researched explanatory variable of snow depth distributions. Peak depth for all years occurred within the general location of the treeline ecotone. Treeline is a transitional zone from predominately forested

hill slopes to more rocky, complex open terrain areas devoid of vegetation. As snow is redistributed from exposed upper elevations, the increase in mean fractional canopy cover with descending elevation (Figure 3.13) results in greater potential for canopy interception. Moeser et al. (2015) used LiDAR data to show that the variety of canopy traits (height and cover) at the interfaces between open and forested areas promote complex and variable snow accumulation and ablation patterns relative to zones with more homogenous canopy traits. Grunewald et al. (2014), Hopkinson et al. (2012), Kirchner et al. 2014 and Zheng et al. (2016) have all published papers where peak snow depth along an elevation gradient occurred below maximum elevation.

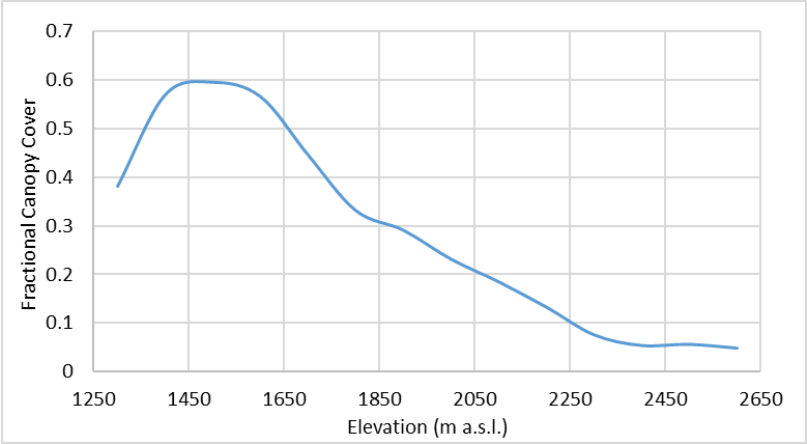

Figure 3.13 Mean of fractional canopy cover within elevation classes.

Declining snow depth below treeline can be explained by hydrometeorological conditions. Environmental lapse rates are the thermal gradients of decreasing air temperatures with elevation gains, which can vary widely in alpine settings (Pigeon and Jiskoot, 2008). Lapse rates frequently result in temperatures >0˚C at the base of mountains while below freezing conditions exist at upper elevations. Low elevation melt can also be enhanced by rain,

which adds thermal energy to the snowpack (Garvelmann et al. 2015). The topographic relief created by mountain ranges promotes orographic precipitation events, where precipitation is enhanced on the leeward sides of mountains. As air masses descend the leeward side of a mountain, they can become exhausted of precipitation prior to reaching the valley bottoms, thus depositing more snow in alpine and sub-alpine zones than mid and lower elevations (Mott et al. 2015). Elevations above the zone of greatest snow depth did not produce a consistent pattern across the years (Figure 3.8). While linear decreases in snow depth are sometimes reported above the peak value (Grunewald et al. 2014; Hopkinson et al. 2010), seasonal and inter-annual distributions influenced by a dynamic phenomenon such as wind are unlikely to be identical at all times of snow depth sampling. LiDAR scans of multiple mountainous study areas have confirmed a variety of upper elevation site-specific micro-scale effects (Grunewald et al. 2014). Peak erosion and redistribution of snowpack typically occurs during and immediately after snowfall while the surface layers are less dense than deeper settled layers (Pomeroy et al. 1997). How far unconsolidated snow is transported depends on shelter, wind speed and potential obstructions (changes in topography or vegetation) along the wind vector after snow crystals are entrained.

Upper elevations tend to have sporadic and heterogenous vegetation cover among exposed, variable rocky terrain. Varhola et al.'s (2010) review highlights that the effects of forest attributes on snow depth are highly variable, in large part due to the influence of canopy geometry on wind patterns. This is supported by significantly different mid-winter elevation class distributions in all classes except one. It is possible that the upper elevation class mid-winter

distributions are not significantly different due to wind effects and the potential for smooth surfaces to develop over topographic and vegetation features (Schirmer and Lehning, 2011). Another possible explanation for this result is that the uppermost elevation class had a smaller N value than other classes (1,200 vs. 10,000) because the >2600 m a.s.l. class is only comprised of 11,288 grid cells whereas the next smallest elevation class contains 247, 468 cells. Too large of an N value, especially with these dense datasets, could produce a Type 2 statistical error and false results (Kaplan et al., 2014). Larger N values were sampled for other depth drivers with bigger class sizes. Statistical power and appropriate sample size is a major hurdle with large datasets and the associated field of research is rapidly expanding as methods to address this issue are developed.

3.4.2.2 Aspect

Aspect delineated mean snow depths have a consistent trend across all datasets. As anticipated, south facing terrain had the lowest mean depth values each year. North and Northeast mean values rounded to the same number for all years, and the means of these classes were the largest. This result was also expected. Significantly high snowmelt rates in south facing areas relative to those with a north aspect have been confirmed in many studies (Haupt, 1951; Anderson and West, 1965; Hendrick et al. 1971; Rowland and Moore, 1992; D'Eon, 2004; Jost et al. 2007). Shortwave radiation provides the energy for melt and refreezing which can create preferential depth losses and potential density gains in aspects with high radiative inputs (south) versus those with little (north)

(Golding and Swanson 1986; Anderson et al. 1958a). This effect holds true at all elevations (D'Eon, 2004). A variety of elevations, topographic features and vegetation attributes within each aspect class is likely to explain why all aspect class distributions are significantly different at mid-winter (Table 3.9). Considering the effects of late-day sensible and solar heat inputs with knowledge of frequent south and southwest winds in the WCW, which enhance ablation and redistribution of snow from wind exposed zones, it is possible that complimentary wind and radiative effects produced lower mean depths (S, SW) directly across from the zones of deeper snow (N, NE).

3.4.2.3 Topographic Position Index

A 101m rectangular window was the best predictor of snow depth for all datasets (Figure 3.2). TPI is still a relatively under researched snow depth driver, and it is complex due to the variation in window sizes across various studies as well as study area size and elevation range. A 25m window has been determined as the optimal size for a small Spanish Pyrenees catchment with a 400m elevation range, whereas a more recent Spanish study found that a 200m TPI window produced the best correlations with snow depth for a different study area with a larger elevation range, suggesting that as study area size and terrain complexity increases, so should window size.

Enhanced snow accumulation was expected to occur in topographic depressions and decline as terrain became more exposed. This occurred in each dataset, however the difference between Upland and Transitional snow depth means were small. These results suggest that the transport of snow to

depressions is a reliable and predictable process, which was also established by Lopez-Moreno et al. (2014) in a study of intra- and inter-annual snow persistence where topographic position index received more attention as an explanatory variable of snow depth than in previous studies.

A challenge of using TPI in snow depth and modelling studies is that depressions vary in depth throughout a given winter season. Once a depression is full of snow, a relatively smooth surface is created. After this point the terrain feature will continue to have a high depth, but new snow loading on top of the smooth surface will not be subject to the same localized shelter that earlier season snow was. With a variety of topographic depressions throughout a large study area, they will likely fill at different rates as a function of depression size, availability of snow for redistribution, and other factors that may influence or intercept the snow while being transported across the landscape. Schirmer and Lehning (2011) confirmed with variorum analyses that scale breaks increased after a snow storm as small scale terrain features were filled. Similar effects have been observed during field and recreation in our study area, as deceivingly smooth snow surfaces occur in areas which are known to have highly complex and variable terrain beneath. Given this, to use TPI as a predictor of event-based snow depth accumulation, a dynamic TPI surface would need to be generated representing different levels of overall snow accumulation throughout the winter season.

3.4.2.4 Canopy Cover

As with topographic depressions, some vegetative features eventually become buried after which topographic smoothing occurs (Schirmer and Lehning, 2011). If all or some of the short vegetation gets buried during the winter and the surface is much smoother than it is in snow-off conditions, subsequent snow accumulation may no longer be influenced by short canopy in that area falling within the Open class. Tree geometry and opening patch size influences snow accumulation patterns (Varhola et al. 2010) by creating localized microclimates with unique wind flows and vectors. Open vegetation is generally more prominent in the sub- alpine and alpine areas as a lack of continuous zones of flat topography starts to limit the formation of stable, fertile soils and thus the size of tree patches, whereas closed canopy is abundant at lower elevations for the opposite reason. Areas of open canopy have greater potential for enhanced redistribution and ablation by wind and radiation. Mean fractional canopy cover is lowest at upper elevations (Figure 3.13), but the mean snow depth for open cover classes is usually higher than the mean snow depth for elevation classes above treeline (Table 3.7). Much of the terrain with intermediate canopy cover has an elevation within 1900-2100 m a.s.l. (Figure 3.13), an approximation of what is a highly variable alpine treeline ecotone within the WCW (McCaffrey, 2017). Treeline elevation classes (Table 3.6), like the intermediate cover class, demonstrated the greatest mean snow depths among the elevation classes. The highly localized nature of turbulence and wind eddies combined with snowfall patterns likely explains the significant difference between canopy cover classes. Previous studies of treeline and snow accumulation using LiDAR (Moeser et al., 2015) found depths to be most

variable at treeline, and this is confirmed visually over a sample area at mid-winter in 2014 (Figure 3.15). With aspect also influencing treeline elevation (McCaffrey, 2017), the distribution of watershed-scale class sizes should be reflected in the sampling while attempting to also capture a representative sampling of elevations and canopy cover strata within each aspect class.

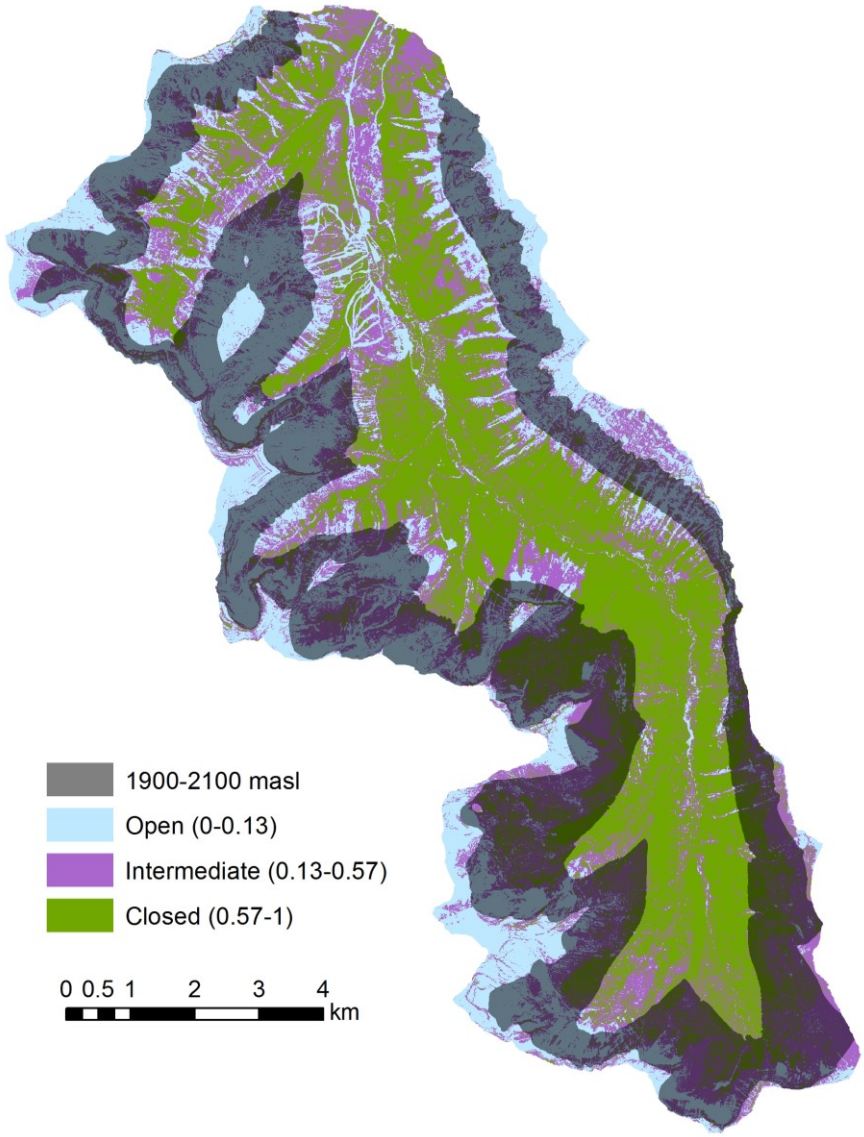

Figure 3.14 Approximate treeline ecotone elevation zone (1900-2100 m a.s.l.) and canopy cover classes.

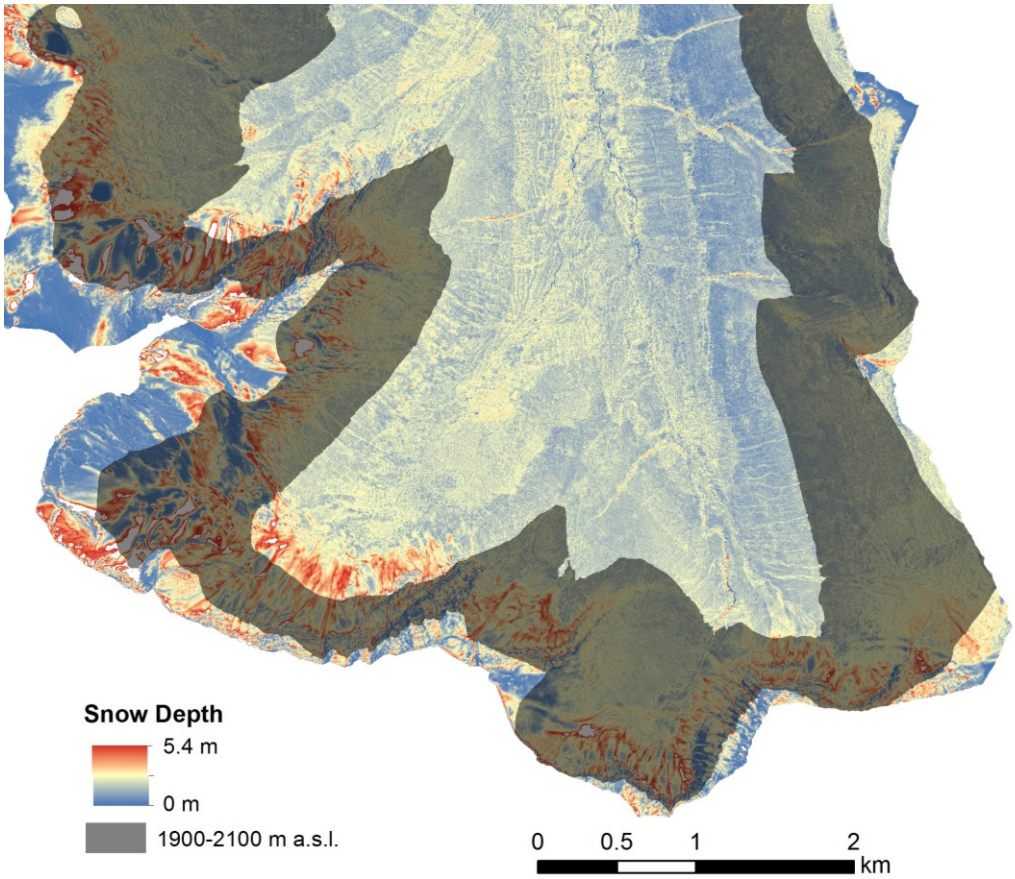

Figure 3.15 2014 depth variability near treeline.

Mean snow depth was lowest in the closed canopy class for all datasets. The closed canopy supports greater interception of snow, increasing the time available for sublimation losses. There is potential to lose 60-80% of intercepted snow in conifer canopies to sublimation during favourable hydrometeorological conditions (low humidity and wind) (Hedstrom and Pomeroy, 1998; Pomeroy et al. 1998). Lodgepole pine (*Pinus contorta*) conifers are the dominant tree type in the study area. If conditions are not favourable for sublimation, snow accumulation can increase below closed canopies as the branch loading capacity is exceeded and intercepted snow reaches the ground. Increased longwave

radiation emitted from tree stems and branches can enhance melt, which is likely why the closed class mean is low during melt onset when temperatures were higher than either of the mid-winter surveys. Canopy cover's relation to snow accumulation and ablation suggest that this variable may less temporally consistent than other snow depth drivers and may require further research or an adaptive sampling approach in an integrated framework.

3.4.2.5 Slope

Considering potential issues arising from slope-induced errors, the higher depth means in the 60-90^0 class at melt onset 2016 and mid-winter 2017 are likely skewed by erroneous outliers in the form of cornices and/or deep drifts adjacent to near-vertical cliff edges. The steep slope class represents <2% of the total area while low and intermediate slopes contain 57% and 41%, respectively. Many of the steepest class cells were eliminated in the quality control step. However, remaining cells in the steep class may still possess slope-induced elevation errors in addition to those associated with cornices and drifts, so the potential for LiDAR snow depth measurements bias in high slope class areas is high. A variety of elevations, terrain and vegetation features as well as aspects is captured in each slope class therefore the significant differences between mid-winter classes is unsurprising. Snow storage increased with decreasing slope angle which is likely more influenced by class size than the depth values themselves. Storage in the steep class is quite small relative to terrain <60˚. Mean snow depths of slope classes didn't produce consistent patterns across the datasets. Considering the potential for terrain induced errors associated

with steep slopes and class size influence storage results, further refinement of this variable and its classes may be warranted if slope is to be used in future analyses. With that said, inclusion of slope in future studies to explain spatial snow depth distributions may not be a valuable use of time based on these generic findings.

3.4.3 Public and LiDAR Snow Data Comparison

Placement of the Gardiner Creek site near treeline is a good start for weather station integration into a monitoring framework, as it captures snow depth in the zone where depths are expected to be higher, although a single measurement in this zone of higher snow accumulation can bias estimates such as snow volume, as seen in Figure 3.16. Conversely, placement in the valley bottom at low elevations, as is the case with the WCSC, can bias estimates in an opposite way, causing underestimations. Grunewald & Lehning 2015 as well as Hopkinson et al 2012 also found that public monitoring sites do not correspond well to LiDAR based estimated of volume and are the only studies to use LiDAR to assess potential relationships. Grunewald and Lehning (2015) emphasize that "It appears that representative cells are rather randomly distributed and cannot be identified a priori" in reference to identifying areas that are representative of static monitoring sites where hydrometeorological data is collected. The elevation of each monitoring station likely explains some of the disparity between public data and LiDAR snow volume estimates, although there is variation in other terrain attributes and temporal alignment (ie sampling WCSC on a different day than the LiDAR survey) to consider as well. A network of

monitoring stations would ideally capture more variation in a variety of terrain attributes, as LiDAR data for the WCW illustrates that snow depth distributions are highly variable even when broken down into classes, and that the relationships between class-based depth distributions from different seasons are not significant.

3.5 Conclusion

Airborne LiDAR is a useful means of acquiring spatially explicit snow depth data due to the large areas covered quickly compared to manual methods. A variety of quality control methods have been implemented on LiDAR snow datasets, each having negligible impacts on watershed scale means. Ultimately the method chosen should be what is most appropriate for the objectives of a given study, although the 99th percentile approach is a consistent approach to creating a comparative metric for multiple datasets. If watershed scale depth is used for modelling or water supply forecasting, the quality control implemented here is not crucial for storage estimates. However, if stratified classes are utilized for local scale estimates, volumes could be skewed at upper elevation and steeper slope sites. An accurate quantification of upper elevation snowpack is useful to understand late season runoff, as these elevations are generally the last to melt. The discrepancy between volume predictions using LiDAR versus public data suggests that the current monitoring regime does not provide a high level of detail of seasonally variable mountain region water resources at the watershed scale. New approaches to evaluating linkages between static monitoring sites and high resolution data may help integrate these sites into

remotely sensed data analyses, such as the examination of snow class signatures.

Many of the well-established spatial snow depth distribution trends that have been observed with manually collected and remotely sensed data are confirmed through the analyses presented here. The occurrence of peak snow depths within treeline was observed across all datasets in the WCW, with this zone also storing most of the seasonal snowpack. The elevation distribution of snow depths above and below treeline, primarily driven by orographic processes, lapse rates and wind redistribution, compliments the work of other LiDAR mountain watershed snowpack distribution studies (Hopkinson et al. 2010; 2012; and Grunewald et al. 2014). Reduced snow depth on terrain subject to the highest amounts of solar radiation as well as frequent winds, south and southwest classes, was another expected result that occurred consistently across all datasets. As expected, TPI results demonstrated depressions contain more snow than other terrain types but of note, this result most consistent for a TPI window of approximately 100m. Closed canopies consistently displayed the shallowest snowpack in mid-winter, likely due to interception sublimation losses and higher longwave radiation fluxes. Slope results were inconsistent, suggesting that for this study area and using these LSDM data, slope is an unreliable explanatory variable of snow depth. The exclusion or further exploration of slope effects are particularly important to consider in modelling exercises too, as it may be confounded by elevation.

In terms of seasonality, a significant Kendall correlation of mid-winter watershed scale depths was observed but the individual class-based strata

illustrated significant differences in snow depth. For this to occur at mid-winter, when distributions are more variable than later in the season (Lopez-Moreno et al. 2017), as well as at melt-onset suggests that terrain depressions and uplands are more reliable drivers of snow depth than other variables, regardless of timing, and that this is a useful variable for locating and stratifying all-season depth sampling strategies or for snow depth modelling routines. As methods of assessing sample size as it relates to statistical power become more established for data mining, N values for both the Kendall correlation and ANOVA could be revisited to ensure the results accurately represent the consistency (or lack thereof) of inter-annual snow depth distributions.

Class sizes, as well as the storage of snow within each, can also help guide future sampling. Considering the occurrence of peak depth and the large storage of snow observed within treeline elevation classes, accurately characterizing this area would be important in an optimal spatially explicit snowpack monitoring framework. The consistent TPI results suggest that this may be a flexible variable to sample, as the clear trend of more snow in depressions and less in upland zones could be compensated for in a model if a TPI class was slightly under-sampled. Slope can and probably should be ignored in sampling design, at least in so far as not all slope classes need to be equally represented given slope is an unreliable driver of snow depth, high slope classes will amplify errors and tend to cover small proportions of watershed terrain, anyway. LiDAR provides a more thorough spatial overview of snow depth distributions and their drivers than is possible by manual measurements, especially in mountainous environments where critical water resources are often difficult to access and quantify in efficient and safe manner.

The spatial, seasonal and inter-annual consistency and variability in snow depth driver classes presented here provide the basis for: a) designing a more time- and cost-effective LiDAR snowpack sampling strategy; and b) extrapolating watershed-wide snow depths from LiDAR sample datasets using spatial imputation techniques. These two objectives, combined with spatially explicit watershed-scale snow density modeling, will be the next steps in building this research into an operational snowpack monitoring framework. As climate change advances, a better understanding of mountainous water supply and therefore a clearer depiction of possible changes in snow depth and volume through more robust monitoring regimes is paramount.

4 SPATIAL AND TEMPORAL CONSISTENCY OF SNOWPACK DEPTH DRIVERS TO SUPPORT OPERATIONAL LIDAR SAMPLING AND MACHINE LEARNING-BASED EXTRAPOLATION

Abstract

Airborne LiDAR can provide high resolution snow depth datasets in a short time period. These datasets provide the spatial coverage necessary to support water supply forecasts, as snow is an integral source of streamflow draining mountainous headwaters. This research utilized LiDAR and random forest modelling to assess the feasibility of using partial datasets (two flightlines) to extrapolate snow depth at the watershed scale under mid and late winter conditions as well as at melt onset. Rasters in integer format proved to be the fastest type of input data to implement the random forest workflow with. Models developed from flightlines had lower R^2 values than watershed based models, with all trials ranging from 0.41—0.61. Spatially imputed datasets were significantly correlated with the original LiDAR values (Pearson Correlation Coefficients of 0.5-0.8, $p < 0.05$). The importance of various snow depth drivers was also assessed, with aspect and TPI shown to be valuable predictors regardless of seasonality.

4.1 Introduction

4.1.1 Snowpack Monitoring

Mountainous terrain is highly complex and the seasonal snow that accumulates in these regions exhibits large spatial variability as a result. This variability can limit the accuracy of spatial snow depth models (Erxleben et al., 2002; Pomeroy et al., 1997). Models can be valuable prediction tools based on a small amount of input data to fill in spatial data gaps over a broad area. Understanding water storage by seasonal snow is especially important in regions where major rivers start in the mountains and are therefore influenced by snow melt (Byrne et al. 2006). Data collection is challenging in mountains due to logistical access and safety constraints. In Alberta, continuous seasonal snow depth is measured by sonic rangers and snow water equivalent by snow pillows. These technologies are automated and typically record values at time increments of 15, 30 or 60 minutes. Both depth and SWE (snow water equivalent) values are also collected manually once every 4-5 weeks, at snow course sites, which are often co-located or at lower elevations than the automated stations. If climate and headwater snowpack trends continue to change, the small number of snowpack monitoring sites in Alberta may not be sufficient for future water supply predictions.

Snow water equivalent is the variable desired for modelling water resource availability and it is the product of snow depth and density. Density can be sampled in the field and has the potential to be modelled from hydrometeorological data (Jonas et al. 2009). The localized range in snow depth can exceed four times that of density (Dickinson and Whiteley, 1972; Steppuhn

1976), thus SWE values are more sensitive to the range in depth variations than density. It is therefore important to prioritize the time and cost-efficient means of acquiring depth data over density, which is where the use of LiDAR and modelling could improve the current availability of headwaters snow data. LiDAR is often used to collect baseline terrain data and if budgets permit, can be an effective environmental monitoring tool that has demonstrated potential for spatially explicit seasonal snowpack monitoring (see Chapter 3 of this thesis), Grunewald et al. 2010, Revuelto et al. 2014b). Considering the high variability of snow depth distributions and the infeasibility of collecting spatially explicit and landcover-representative data manually, coupling LiDAR snow depth models with density data will provide more spatially complete and accurate headwater SWE estimates than is currently possible using traditional methods.

4.1.2 LiDAR and Snowpack Measurement

Data availability and quality has limited snow depth model performance in the past when manual acquisition was the primary means of acquiring model inputs. With manual measurements, values are often irregularly spaced over large study areas. Interpolation techniques are then applied to the field data to generate spatially complete datasets (Erxleben et al. 2002). When depth data collection is limited to manual measurements, and in some cases georadar (Marchand and Killingtveit, 2005), values are often irregularly spaced. Field measurements are, however, invaluable for calibrating and validating remotely observed products. Due to LiDAR's ability to provide high resolution data over large areas in a relatively short data collection period, airborne LiDAR is gaining popularity as a snow depth measurement tool (Grunewald et al. 2010, Banos et

al. 2011, Lehning et al. 2011, Hopkinson et al. 2012, Grunewald et al. 2013, Bhardwaj 2016, Zheng et al. 2016, Lopez-Moreno et al. 2017). LiDAR datasets are generally collected over entire study areas which is less cost and time efficient than using a spatial sampling and statistical modelling approach to infer snow depth at the watershed scale (e.g. Hopkinson et al. 2012).

LiDAR provides dense point clouds from which snow depth can be computed with decimeter accuracy under ideal survey conditions (Hopkinson et al. 2004) at a 1m grid cell resolution. Tens or hundreds of millions of depth measurements are provided by a LSDM (LiDAR Snow Depth Model), whereas intensive field sampling campaigns will only yield datasets with 100s of depth values (Elder et al. 1998), and possibly low 1000s if the measurements are collected over multiple days. Only three of the studies presented in Table 4.1 utilized LiDAR (terrestrial or airborne) for model inputs. LiDAR's ability to make observations quickly and safely are advantages when spatial coverage and measurement resolution are considered, as manual measurements can introduce spatial bias as safety and ease of access are high priorities during field campaigns. It can be expensive, however, to conduct airborne LiDAR surveys over large areas and requires favourable weather conditions for flying. Quality control is an essential step in preparing remotely sensed data for analysis. Some values in LiDAR datasets are the result of vertical and horizontal error, misclassification of canopy returns as ground returns (Hodgson, 2004) and multipath. If the ground – aircraft portion of the two-way travel time of a laser pulse is increased due to reflection off multiple surfaces before returning to the sensor, a pit is created below the true surface (Hopkinson et al. 2004). Unreasonable snow depth values can also occur when steep slopes and

horizontal DEM uncertainty are present (Hodgson et al. 2005), as a small shift in either the x or y direction on a steep slope can cause a high angle cell to change location between the DSSM and DEM.

Quality control approaches vary across LiDAR snow studies. Thresholding is one method, where a maximum possible depth is assigned. Measuring depth in a topographic depression to assign a threshold value was used by Revuelto and others (2014b). Other researchers (Kirchener et al. 2014; Grunewald et al. 2010) have employed long-term field depth records or canopy height information to determine a feasible threshold depth value. Alternatively, cells can be eliminated from dense datasets based on their outlying and unrepresentative properties within the overall sample population. Deems et al. (2013) and Bhardwaj et al. (2013) both provide further explanation of LiDAR as a cryospheric measurement tool and the inherent challenges of this technology, particularly in mountainous zones.

The potential to use LiDAR sampling (i.e. not covering the whole study area) to model the complete extent of a study area is an under-explored field of research and was first addressed in Alberta by Hopkinson et al. (2012). Hopkinson et al. (2012) adopted a snow accumulation unit approach to modelling, where terrain and land cover class means were determined with LiDAR transect sample datasets and extended over larger, unsampled areas based on the physical attributes of the terrain. This concept is the basis for the analyses in this chapter, which were conducted to provide a foundation for data mining and optimization of dense LiDAR snow depth data collection with a machine learning approach. Optimized LiDAR snow depth sampling has the

potential to reduce the costs of surveys and enable more widespread use of LiDAR for collecting high resolution data as water supply monitoring data inputs.

4.1.3 Snow Depth Modelling

Since the inception of snow depth and SWE modelling, primary methods used have been multiple linear regression (MLR; Golding, 1974; Elder et al., 1991) and binary regression trees (BRT; Elder et al. 1998; Balk & Elder, 2000).These techniques can determine the order of snow depth predictor variable importance (as measured by the contribution of the variable to overall model accuracy) and apply those statistical relationships in a predictive modelling context. This has been achieved primarily through using different combinations of variables, power functions of variables and/or interactions with variables to see which produces the best model, rather than determining a ranked order of predictor variables ability to explain snow depth distribution. The first documented regression model for predicting snow accumulation was developed in 1960 (Ku'zmin, 1960) using canopy cover in a Russian boreal forest study area. Early multiple linear regression (MLR) studies typically focused on SWE, instead of depth. Another early study that utilized MLR included multiple predictors, which could explain 48% of the variability in SWE using elevation, topographic position index, aspect, slope and forest density (Golding, 1974). A similar approach was applied by Lopez-Moreno and Nogues-Bravo (2006) who used elevation, elevation range, radiation and two location parameters to explain 50% of SWE variability. As a statistical method MLR is well established and model performance is generally limited by input data complexity and quality.

A variety of MLR-based studies have used different variable combinations with varying degrees of success (e.g. Hosang and Dettwiler, 1991; Elder et al. 1991. Chang and Li, 2000; Jost et al. 2007). Comparing results across studies can be challenging given the source and quality of input snow data, choice of predictor variables, as well as model workflow is often study specific (Table 4.1). Consequently, the results could also reflect methodological choices. Results (in terms of R^2) of other studies are provided as ranges as many of the authors used multiple study areas or input data types to explore model performance but maintained the goal of predicting snow depth or water equivalent. Multiple linear regression modelling techniques were recently used in tandem with Airborne LiDAR snow depth data from around the world (Grunewald et al. 2013). This paper represents what is still a small subset of publications which utilize LiDAR snow data for modelling instead of interpolated rasters from relatively small field data sets. With LSDMs from around the world, the authors compared snow depth drivers in mountainous zones. They concluded that global snow depth models do not exist and that predictor variables interact in such a complex manner that their importance is often study-site dependent, which is in line with other research suggesting some snow modelling may be site-specific (Molotch et al., 2005).

Table 4.1 Summary of snow modelling studies.

Author and Year	Data	Method	Variables	Results (r^2)
Elder et al., 1998	Field	BRT	solar radiation, slope angle, elevation	0.6-0.7 (snow depth)
Balk and Elder, 2000	Field	BRT	net solar radiation, elevation, slope, vegetation cover type	0.54-0.65; 0.60-0.85 (snow depth)
Erxleben et al., 2002	Field	BRT	elevation, slope, aspect, net solar radiation, vegetation	0.18-0.30 (snow depth)
Winstral et al., 2002	Field	BRT	upwind index, drift delineator, elevation, solar radiation, slope	<0.4; 0.50-0.60 (snow depth)
Marchand and Killingtveit, 2005	Field, Geodar	MLR	elevation, aspect, curvature, slope	0.5-0.48 (snow depth)
Lopez-Moreno and Nogues-Bravo, 2006	Field	BRT	elevation, elevation range, solar radiation, slope, location (to seas/oceans, elevational divide)	0.15-0.70 (snow depth)
Lopez-Moreno and Nogues-Bravo, 2006	Field	MLR	elevation, elevation range, radiation, slope, location (to seas/oceans, elevational divide)	0.51-0.58 (snow depth)
Lehning et al., 2011	ALS	MLR	elevation, fractal roughness	>0.70 (snow depth)
Grunewald et al., 2013	ALS	MLR	elevation, slope, northing, wind, surface roughness	0.27-0.90 (snow depth)
Golding, 1974	Field	MLR	elevation, TPI, aspect, slope, forest density	0.48 (SWE)
Elder et al., 1991	Field	MLR	elevation, slope, radiation	0.27-0.40 (SWE)
Molotch et al., 2005	Field	BRT	elevation, slope, aspect, northness	0.28-0.41 (SWE)
Plattner et al., 2006	Field	MLR	elevation, curvature, distance from ridge, shelter	0.41 (SWE)
Jost et al., 2007	Field	MLR	elevation, aspect, forest cover, solar radiation, temperature	0.83-0.88 (SWE)
Litaor et al., 2008	Field	BRT	upwind index, elevation, slope, aspect, slope-aspect topoclimatic index, solar radiation, plant biomass, species richness	0.85-0.91 (snow depth and SWE)
Grunewald et al., 2010	TLS	MLR	elevation, slope, aspect, radiation/elevation, slope, max SWE, wind speed	0.30-0.40 (daily ablation rates)

Binary regression trees (BRTs), a type of decision tree, are another common statistical method applied in snowpack modelling. Decision trees are a non-parametric means of recursively splitting predictor variable data to minimize each group's sum of squared residuals (Breiman et al. 1984). The non-linear, hierarchical means by which BRTs relate dependent and independent variables makes this approach highly desirable for analyzing a medium as complex as the cryosphere. As is common with classification/regression techniques, BRTs bin input data such that subsets become progressively homogenous as more splits are made (Kuhn and Johnson, 2013). At each node of a BRT, splits remain binary down a series of nested if/then statements.

Individual BRTs typically yield better results for distributed snow depth and water equivalence estimates than other methods (Molotch et al. 2005). An advantage of decision trees over MLR is their interpretability, as well as their capacity to handle more predictor variable distributions (e.g. skewed, continuous, categorical) than MLR, for which a primary assumption of the technique is that distributions are normal (Kuhn and Johnson, 2013). Details on BRTs and their use in snow depth and water equivalent modelling are further described in Molotch et al. 2005, Breiman et al. 1984, Elder et al. 1995 and 1998, Balk and Elder 2000. Across the previously mentioned studies, 18-90% of variation in snow depth or water equivalence has been explained with BRT or MLR approaches (Table 4.1). Both methods can provide information on variable importance, but this information has typically been reported in terms of which combination of variables in a model produce the best results rather than a ranking of individual predictor importance.

The complex nature of snow, terrain and microclimatology is perhaps one of the greatest challenges to establishing replicable snowpack modelling routines, despite impressive technological advances in data collection, processing and analysis over recent decades. Considering the logistical challenges associated with collecting robust datasets manually, as well as the cost of high resolution LiDAR datasets, optimized LiDAR sampling has the potential to provide the data, at a reduced cost, required for site (watershed) specific models to provide sufficient detail for meaningful water supply predictions.

4.1.4 Random Forest Modelling

In mountainous environments, there are many predictors which interact at a variety of spatiotemporal scales to produce an observed snow depth distribution. Relationships between snow depth and independent variables are often non-linear, which complicates predictive modelling (Anderton, 2000; Nogues-Bravo, 2003), especially when MLR is used. Random Forest, a statistical method similar to BRTs with classification and regression capabilities, is an algorithm which creates ensembles of decision trees or a forest. Variable order importance of snow depth controls in the Spanish Pyrenees has been determined with RF (Lopez-Moreno et al. 2017) and it has also been used to quantify snow volume and predict snow depth errors (Tinkham et al. 2014), but the combination of RF and snow depth remains relatively underexplored compared to MLR and BRT methods. Random forest has recently been utilized in studies of wetland mapping (Chignell et al. 2018; Wang et al. 2015) as well as crop yield

monitoring (Jeong et al. 2016) in attempts to reduce the costs and/or time of field or aerial photography-based monitoring.

Random forest (RF) employs a sophisticated data splitting process that gives it machine learning status over MLR or BRT workflows. BRT and RF methods are similar in that they both rely on decision trees with binary nodes that create branches (yes/no). A forest is created when multiple decision trees are used. RF splits data more intelligently than is possible through other techniques. First, a subset of the data is randomly selected, from which a portion is used to create the tree/model and the other portion is used to test and validate the model. As many bootstrap samples as there are trees are taken from the data, and a tree is fit to each sample (Cutler, 2007). As the data are split, a few randomly selected variables are used at each binary node of the tree. Unused predictors at a given split do not influence the results and make computation more efficient. Predictor variable importance and models based on ensembles rather than a single tree result in more stable models, making RF more desirable than BRTs (Kuhn and Johnson, 2013). RF's ability to detect relationships within extremely large datasets makes it an optimal choice for modelling snow depth using high resolution airborne LiDAR data, provided sufficient spatially coincident predictor variables are available.

RF differs from other statistical modelling techniques in that it does not rely on statistical significance (as in MLR) to determine variable importance. Importance in RF can be determined by the R statistical software program using the Random Forest package (Liaw and Wiener, 2002). Outputs of variable performance from this package are increase in mean square error (MSE). The package's MSE output is essentially a

comparison of how the MSE of the original bootstrap sample changed when variable values were permuted. When a variable is important to a model, permuting values for that variable over the dataset will negatively affect predictions. Output MSE values, given as "% Increase MSE", can sometimes be greater than 100%. This metric is derived first by finding MSE for the whole model before and after permutation of predictor values, then the percent change in MSE between the two values (Breiman, 2002). Larger MSE values indicate that when random values replace true predictor values the results are degraded, thus demonstrating the replaced predictor variable is important.

4.2 Objectives

It is important to find a balance between the optimal data collection coverage, computer resources, total analysis time and quality of the outputs for it to become a viable component of operational water resource monitoring frameworks. To efficiently utilize LiDAR data and reduce input data requirements, machine learning approaches, such as Random Forest (RF), to model and impute snow depth over large areas offer great potential. Knowledge of snow depth predictor variable importance over multiple years can help guide decisions in sampling schemes if any terrain and landcover stratification is to be implemented. The temporal element of this analysis can also guide timing of data collection by illustrating which time during the winter yields better predictive results: mid-winter or melt onset. Mid-winter sampling might produce similar results, while melt onset may represent more homogeneous conditions in late season snowpacks (Lopez-Moreno et al. 2014). Data points extracted from the watershed and individual flightlines will be compared to determine if acceptable

results can be achieved with partial datasets. With the dense datasets provided by LiDAR, machine learning is a logical choice for analyses and modelling.

This research chapter addresses the following objectives:

1) Use of RF to model snow depth and assess prediction accuracy under mid-winter and melt onset conditions
2) Assess the consistency of inter-annual and seasonal predictor variable importance; i.e. do watershed-scale snowpack depth controls vary through time?
3) Determine the feasibility of using lidar snow depth sample datasets for high-resolution, spatial imputation
4) Optimise the RF snow depth modeling approach by exploring trade-offs between computation time and output quality as a result of training data sample sizes, input raster type, raster resolution and input data quality control methods.

Ultimately, this workflow is being developed and evaluated for operational monitoring purposes. Therefore, a variety of modelling trials will be executed to determine more effective ways to utilize computer resources and thus save time. These trials are meant to provide a foundation for the use of RF with airborne LiDAR snow depth data to predict snow depth in unsampled areas. Results of optimized steps are given, as well as further recommendations for potential efficiency gains. This chapter presents the first machine learning study that utilizes the Random Forest algorithm with LiDAR snow depth models (LSDMs) to explore an operational snow depth sampling and spatial imputation framework.

4.3 Methods

4.3.1 Data

The West Castle Watershed (WCW) study area is in southwestern Alberta, Canada, along the continental divide (Figure 4.1). The extent of the LSDMs from which input data were taken is shown in Figure 4.1. LSDMs are acquired by subtracting a snow-free DEM (digital elevation model) from a winter DSSM (digital snow surface model) (Eq 4.1). Three LSDMs covering 103 km^2 and a fourth dataset covering the extent of two flightlines (Figure 4.1) exist for the WCW. Larger scale regional context of the study area is provided in chapter 3. Two of the datasets were collected in February, corresponding to mid-winter (MW) conditions, and the third dataset with complete watershed coverage was surveyed at melt onset (MO). The flightlines (FL) used in this analysis are the result of an intended full survey in March 2017 being interrupted by a rapid weather change, resulting in one sweep of the area of interest on its east and west boundaries in late winter (LW). For analyses in this chapter, WCW trials refer to those utilizing data from the entire watershed whereas FL trials only utilized input data from within the flightline extent (Figure 4.1). Each 1m LSDM was processed as per Hopkinson et al. (2012) and the canopy cover raster was calculated as a ratio of canopy returns to all returns (Barilotti et al., 2006) (Eq 4.2).

$$LSDM = DSSM - DEM \quad \text{(Eq 4.1)}$$

$$\text{Fractional Cover} = \text{Canopy Returns} / \text{All Returns} \quad \text{(Eq 4.2)}$$

Multiple quality control options exist for LSDMs (Chapter 3 of this thesis) and one approach is to assume all data beyond the 99th percentile value are outliers and can be removed from the sample without substantively impacting the population

attributes. That approach was used in this study, as the objective is to develop a snow depth mapping and interpolation routine that is generally applicable to the watershed as a whole rather than accurately characterising the behaviour of outlying data points. With the upper 1% of data eliminated, we can avoid assigning a depth value to places where there may not be any snow, as would be done by thresholding, and modelling is restricted to areas where we are more confident in data quality.

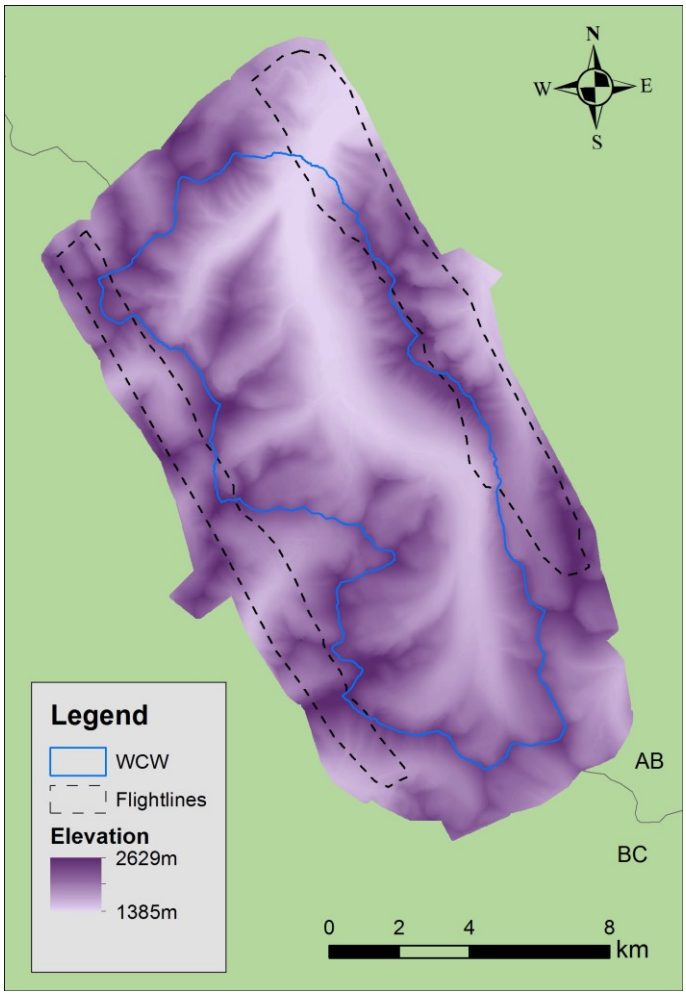

Figure 4.1 The WCW is southwest of Calgary, Alberta, with the west and south edges of the watershed following the continental divide between British Columbia and Alberta. All surveys covered the same extent except for the March 2017 flight lines.

Although the input data is available as 1m grid cells, the LSDMs, DEM and canopy cover layers were resampled to 3m x 3m using Bilinear interpolation after quality control. This resolution could integrate well into hydrological models, especially considering that resampling from high to low resolution is more feasible than low to high. Adopting such a high resolution for snow modelling is rare and not necessary for a watershed-scale operational estimate of snowpack volumes. Carrying out statistical modelling workflows at a 1m grid cell resolution, particularly when the end goal is to optimize the analysis to take as little time as possible without compromising results/performance, is not conducive to efficient computation. With higher resolution data, more grid cells exist and a larger sample of training data would be required to adequately represent the total number of grid cells across a 103 km^2 area. A great deal of computational power is required to manage large datasets and implement the random forest workflow, so carrying it out efficiently is important, especially in operational settings. With lower resolution input data and data volume, processing time can be reduced through use of less RAM than 1m input data would require.

Resampling rasters derived from neighbourhood functions such as aspect, slope and TPI can change z-values in the process and can potentially create surfaces that misrepresent true terrain attributes (Kienzle, 2004). For this reason, aspect, slope and TPI were calculated after a 3m DEM was available. Slope was calculated in degrees (Eq 4.3). Topographic position index (TPI) is a measure of terrain concavity (depressions), convexity (ridges) or transitional terrain, which compares each DEM grid cell to the elevation change of surrounding cells in a user-defined window. The optimal window size and shape to explain WCW snow depth is explained in Chapter 3 (Figure 3.2).

Slope ($^{\circ}$) = arctan(rise / run) (Eq. 4.3)

4.3.2 Random Forest Workflow

Five predictors were used for this study (elevation, aspect, slope, topographic position and fractional canopy cover), all of which have proven to be important predictors of snow accumulation in the Albertan Rockies (Golding, 1974b). Golding (1974b) determined this using multiple linear regression, which has also been used to demonstrate that elevation, aspect and slope are crucial to snow accumulation modelling (Chang and Li, 2000; Lopez-Moreno and Nogues-Bravo, 2006; Marchand and Killingtveit, 2005; Grunewald et al., 2013). Elevation has consistently performed well in BRT modelling workflows (Elder et al. 1998; Balk and Elder, 2000; Winstral et al., 2002).

Before random forest models were applied to each dataset and extent (WCW and FL), the impact on processing time and model statistics of training data format and sample size as well as forest size was assessed. An explanation of script order and location for the analyses in this chapter is available in Appendix A. Using the training data and predicted snow depths at the same X, Y locations, R^2 and RMSE (root mean square error) were both presented as R^2 is a relative measure of statistical correspondence whereas RMSE is an absolute measure. Using both metrics to assess model performance provides an opportunity for more thorough assessment in the workflow optimization. These trials were carried out with the Melt Onset 2016 data as later season snowpacks are more homogenous than earlier distributions, reducing variability in the input data (Lopez-Moreno, 2017). For random forest size analyses (100, 250, 500 and 1000 trees; Table 4.2), which was the first step in optimizing

workflow parameters (Green steps in Figure 4.2), integer format was used to reduce computation time (Bonham-Carter, 1994) with a 50,000 training point sample. Raster format influences the amount of storage space required (Bonham-Carter, 1994), making it important to assess how using integer-based rasters as modelling inputs, which occupy less memory than floating point data, could affect processing time and results. To create integer format rasters, predictors and the LSDMs were converted to a larger unit (i.e. The LSDM was multiplied by 10 to produce whole number snow depths in decimeters), then rounded and converted to integer. The maximum negative value of the TPI range was added to each TPI layer to avoid potential problems handling negative values during subsequent R script implementation. A trial with classified integer data, which further reduces the range of values and memory required (Bonham-Carter, 1994), was also included in the comparison of integer and floating point data. Predictions were not made over the whole watershed for exploratory optimization (number of trees, raster format and sample size) trials to save time, but rather just the model statistics between training and the corresponding predicted points are presented.

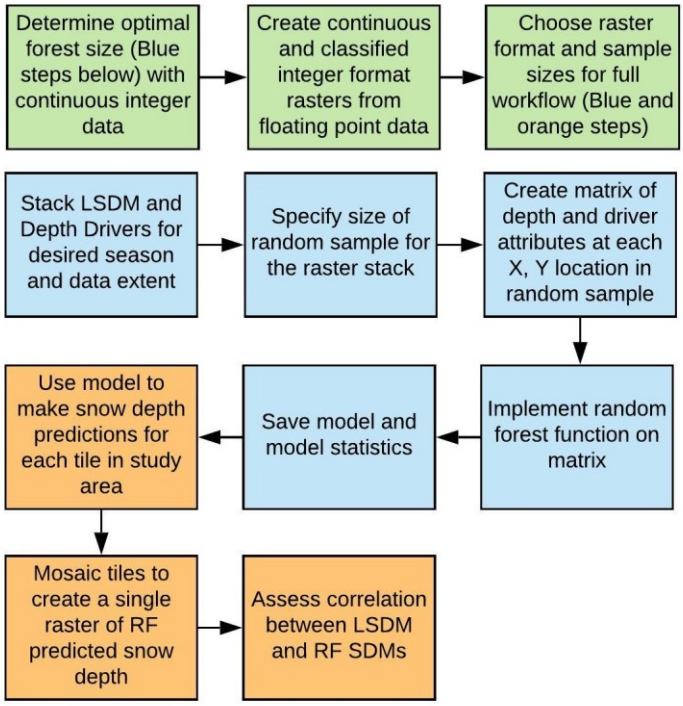

Figure 4.2 Random forest modelling workflow.

To determine an appropriate sample size of training points with respect to processing time and correspondence between training and predicted data, trials using 50,000 (0.44% of 3m grid cells in the WCW) and 100,000 (0.88% of 3m grid cells in the WCW) points from the WCW were completed. Once a suitable sample size for the WCW was determined, the same proportion of WCW training points was used to identify the sample size of training points for the flight lines. For sampling training data, rasters were stacked by XY coordinates (Figure 4.3) and a random sample of points was taken across the entire area (WCW or FL extent). This step of the code simply requires identifying a seed value to ensure the same XY locations are sampled for each data extent and does not involve any specification of sampling a certain amount of snow

depth driver attributes (e.g. X% of sample must contain north facing terrain) as it is purely random. Sampling a stack returns a matrix with the values of each layer at every X, Y location (Figure 4.3). This information was used to determine the proportion of each snow depth driver class (Figure 3.5) present in the random sample (Table 4.3). The matrix is also the input dataset for creating the final random forest models (Figure 4.2) with the Random Forest package for R statistical programming software (Liaw & Wiener, 2018). After the ideal numbers of trees and sample points as well as raster format was determined, the TuneRF function within the Random Forest package was added to the script, which recreates a random forest using only the most useful variables in the predictive model. Making predictions is computationally intensive, therefore predictions were made for subsets of the WCW (tiles) and multi-core capabilities were implemented in the prediction portion of the script using the DoSNOW (Simple Network of Workstations) package in R (Tierney et al., 2018). Tiles were mosaicked to create a single raster for which correlation results and summary statistics (mean, maximum depth etc.) could be determined (Figure 4.2). Model performance was assessed in three ways at the watershed scale: R^2, RMSE and Pearson Correlation Coefficients (PCC) using the same n-value as the training data sample size in each trial. While R^2 and RMSE is calculated from training data and corresponding predictions, PCC values are derived from independent samples of the LSDMs and RF SDMs (snow depth models), maintaining the same sample size.

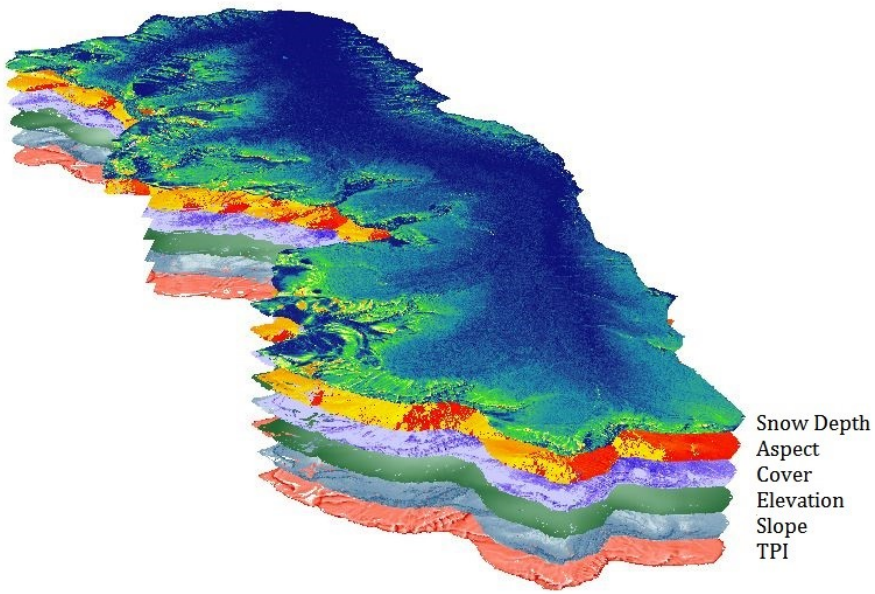

Figure 4.3 An example of a raster stack of the WCW, aligned by XY coordinates.

4.4 Results

4.4.1 Input Raster Types/RF Optimization

Results are presented in the order they are output by the workflow (Figure 4.2). For the green boxes which represent model optimization, forest size was the first parameter assessed. The R^2 and RMSE of the model (between training and corresponding predicted points) was unchanged by forest size (Table 4.2). As forests got larger, processing time to create a random forest model increased. RMSE is given in units of meters, and therefore rounded to a single decimal place.

Table 4.2 Forest size (ntree) optimization timing.

ntree	R^2	RMSE	Time (min)
100	0.61	0.5	83
250	0.61	0.5	159
500	0.61	0.5	202
1000	0.61	0.5	275

Random sampling of the two data extents (WCW and FL) resulted in relatively similar proportions of each snow depth driver class being present in the training data, compared to the proportion of grid cells each class occupies at the WCW scale (Table 4.3). Elevation, TPI and slope classes are well represented in the WCW training data. Notable deviations from the watershed proportions are present in the WCW seed (the value that dictates which XY locations are randomly sampled) for the intermediate and closed canopy cover classes, where the seed oversampled intermediate cover and undersampled closed canopies. The WCW seed also contains fewer grid cells that face northeast and a greater amount of northwest terrain compared to the percent areas of those classes for the entire watershed. The FL seed has a higher proportion of data in classes above 1700 m a.s.l. as well as below 1400 m a.s.l. Open canopies were oversampled by the FL seed compared to the other cover classes. Slope and aspect classes training data proportions corresponded the best with watershed class percent areas. TPI was poorly represented by the FL training data with 98.1% of the training data coming from uplands.

Table 4.3 Percent of each WCW snow depth driver class in the training data.

Driver	Class	% of WCW area	% of WCW Training Data	% of FL Training Data
Elevation (m a.s.l.)	<1400	1.4	1.4	3.0
	1400-1500	9.5	9.6	6.0
	1500-1600	11.5	11.8	4.8
	1600-1700	11.3	11.2	9.3
	1700-1800	11.7	11.7	12.1
	1800-1900	11.6	11.7	12.6
	1900-2000	12.6	12.5	12.4
	2000-2100	11.3	11.1	12.2
	2100-2200	8.7	8.6	10.6
	2200-2300	6.0	6.0	8.9
	2300-2400	3.3	3.2	5.5
	2400-2500	0.9	0.9	2.1
	2500-2600	0.2	0.2	0.4
Cover	Open	33.2	33.9	46.4
	Intermediate	33.2	40.9	36.9
	Closed	33.6	25.3	16.7
TPI	Depressions	25.3	25.5	1.5
	Transitional	38.7	38.3	0.4
	Uplands	36.0	36.2	98.1
Aspect	N	5.1	5.4	5.3
	NE	17.6	12.4	12.8
	E	17.7	17.7	14.3
	SE	15.4	15.3	12.9
	S	7.9	8.0	12.2
	SW	11.5	11.9	12.4
	W	14.1	13.5	14.7
	NW	10.8	15.7	15.5
Slope	0-30	58.5	58.6	53.0
	30-60	39.8	39.7	44.7
	60-90	1.7	1.7	2.3

Using 50,000 points from integer files performed best in terms of model R^2, RMSE and processing time compared to other trials (Table 4.4). Although classified integer data had a faster run time than continuous integer data, model R^2 and RMSE was

compromised. Minimal variation in RMSE and R^2 was observed for trials with continuous data, although floating point inputs and larger sample sizes increased processing time. 50,000 point sample sizes extracted from continuous, integer format rasters was therefore used in subsequent modelling (Section 4.4.2).

Table 4.4 Outputs and timing as a result of input data type and point count.

Raster Format	n	R^2	RMSE	Time (min)
Floating point	50,000	0.61	0.5	140
Integer	50,000	0.61	0.5	79
Integer classes	50,000	0.56	0.5	18
Floating point	100,000	0.63	0.5	606
Integer	100,000	0.62	0.6	602

4.4.2 Inter and Intra-annual Predictor Importance

Predictor importance is represented by % Increase in MSE, the change in model MSE when a predictor is removed from the permutation process. Under mid-winter conditions, aspect was the most important predictor for both MW 2014 and 2017 (Figure 4.4). TPI was among the top three as well, as 3rd most important for MW 2014 and 2nd for MW 2017. Elevation had the 2nd highest importance for MW 2014 whereas its was less important than slope for MW 2017, which was 3rd. Elevation and aspect's importance under melt onset conditions was much higher than other drivers.

For the flight line-based predictions, aspect and TPI were consistently among the top three drivers for each dataset (Figure 4.5). Slope displayed higher importance than cover or elevation under mid-winter conditions whereas late winter and melt onset importance did not have slope in the top three. Elevation was only in the top three

drivers for late winter and melt onset data when training data was extracted from the flightlines.

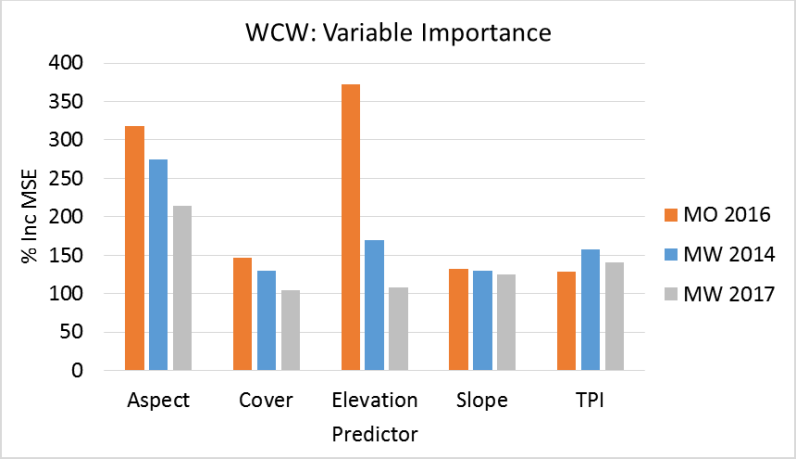

Figure 4.4 MSE values for input data from the watershed (WCW). Note: higher 'Inc MSE' represents better predictive capability (see text).

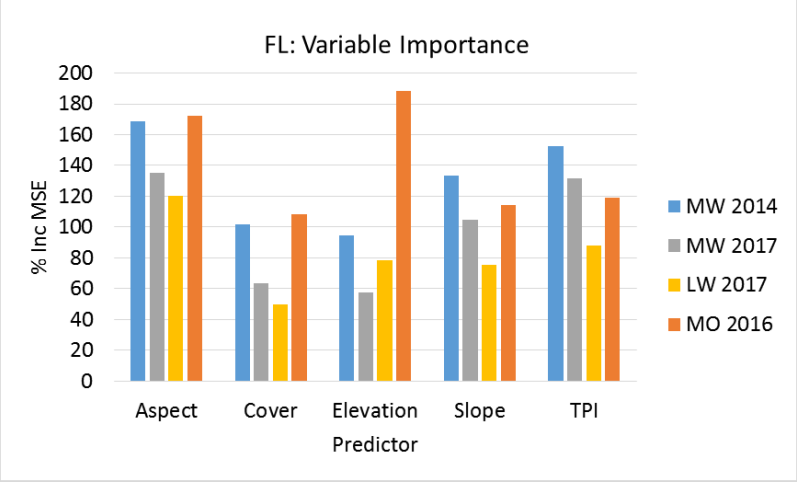

Figure 4.5 MSE values for input data from the flight line (FL) sample extent.

4.4.3 RF-Predicted Snow Depth Models

Model R^2 was consistently higher for WCW based training points compared to the flight lines (Table 4.5). PCC values between the original LSDM and predicted RF SDMs were also higher for WCW data, although all coefficients were significant. The best R^2 and PCC values occur under melt onset conditions. RMSE (in units of meters) values indicate the same absolute error with WCW data, and similar values for MW and MO conditions with FL based models. Although LW 2017 R^2 is low and RMSE is higher, the PCC between LiDAR and RF predicted data is slightly higher than for either mid-winter prediction using flightline training points.

*Table 4.5 Workflow timing, model performance and depth correlation data using Pearson Correlation Coefficients. * indicates $p < 0.05$.*

Data Extent	Season	R^2	RMSE	Time (min)	PCC
WCW	MO 2016	0.61	0.7	236	0.79*
	MW 2014	0.47	0.7	252	0.69*
	MW 2017	0.44	0.7	260	0.66*
Flight line	MO 2016	0.51	0.8	70	0.66*
	MW 2014	0.44	0.8	80	0.42*
	MW 2017	0.41	0.7	76	0.44*
	LW 2017	0.41	1.0	75	0.46*

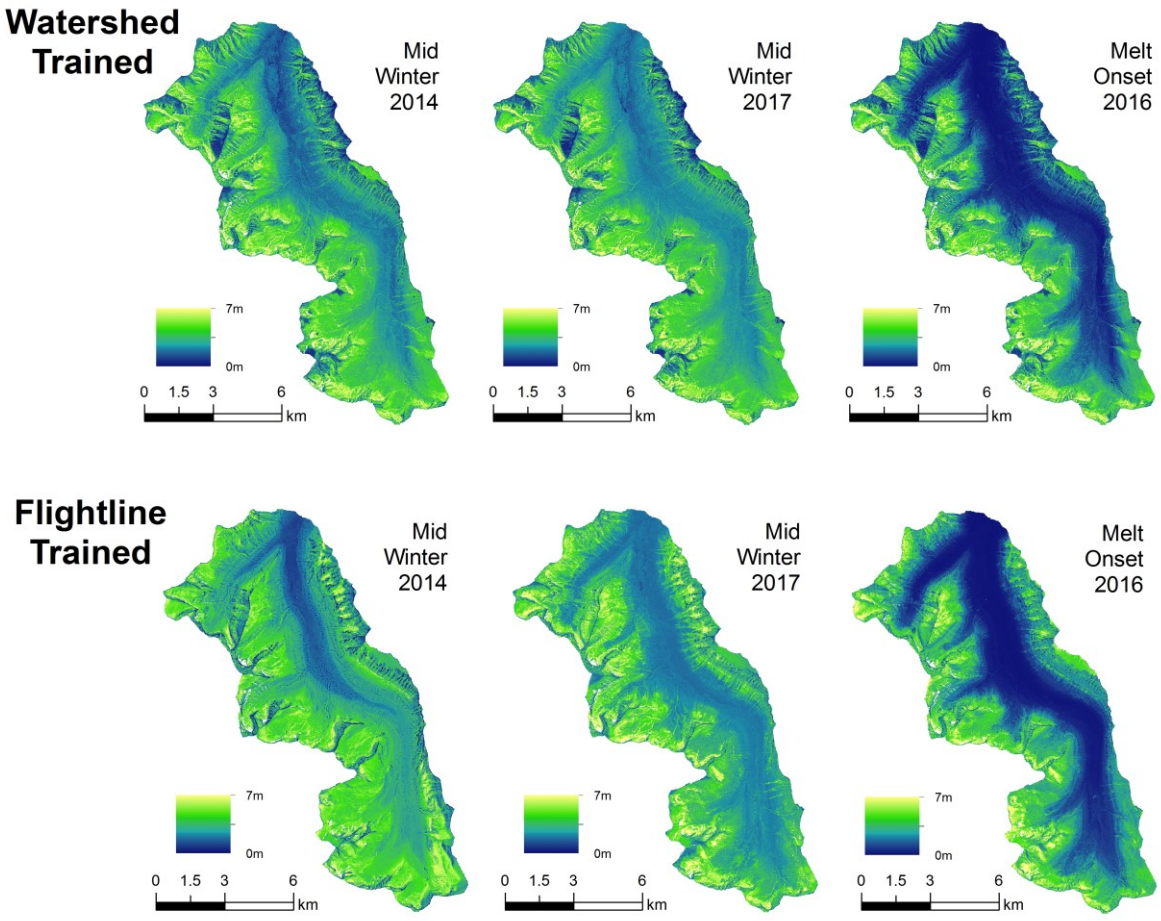

Figure 4.6 WCW and FL trained RF SDMs.

Summary statistics of each SDM (Table 4.6) illustrate that the LSDMs and RF SDMs had the same mean when training points were extracted from the whole watershed. Flightline based training points produced higher mean depths than observed in the LSDMs or WCW-based RF SDMs. RF predicted maximum depth values were consistently smaller than the LSDM value. Smaller standard deviations were observed for flightline based data compared to LSDMs and WCW RF SDM statistics.

Table 4.6 Summary statistics for LiDAR and Random Forest based snow depth models (SDMs), predicted using points from the watershed (WCW) or flight lines (FL).

Season	SDM	Mean (m)	Stdev (m)	Max (m)
MO 2016	LSDM	1.2	1.2	6.3
	WCW RF	1.2	0.9	5.5
	FL RF	1.5	1.1	5.3
MW 2014	LSDM	1.5	0.9	5.4
	WCW RF	1.5	0.7	4.7
	FL RF	2.2	0.7	4.5
MW 2017	LSDM	1.5	0.9	8.1
	WCW RF	1.5	0.7	7.2
	FL RF	1.9	0.8	6.2
LW 2017	LSDM	1.3	1.5	8.1
	FL RF	2.0	1.0	7.2

4.5 Discussion

4.5.1 Sampling and Model Optimization

The potential viability of flight line sampling as a basis for RF snow depth model training is supported by the results in Tables 4.5 and 4.6. However, the disparities in depth and driver class proportions between flight line training data and the whole WCW is likely a function of sub optimal sampling locations. The flight lines (Figure 4.1) were surveyed on the east and west edges of the N/S oriented watershed. By sampling the watershed boundaries, high elevation (>2100 m a.s.l.) and upland terrain is more abundant (Table 4.2) as it is indeed elevation that defines the watershed boundaries. With much of the FL training data falling within treeline and above, higher depths in these zones (Figure 3.7) could skew predictions if elevation is an important predictor. The WCW training data class proportions are more similar to the WCW class sizes,

although the WCW and FL training data included fewer closed canopy points than are present in the entire WCW (Table 4.2). Canopy cover decreases with elevation gains above treeline in the WCW (Figure 3.13), explaining why more cells with open and intermediate canopy attributes were present compared to closed canopy. Closed canopies had lower depth distributions compared to the other classes (Figure 3.10), and their undersampling in the training data may inflate depth values. With that said, mean depth values in Table 4.6 remained the same for WCW based predictions as they were in the original LSDM, suggesting that the training data was not influenced by the distribution of canopy characteristics in the input dataset. Upland terrain was represented by 98.1% of the FL training data, which could favour lower depth values (Figure 3.9; Hopkinson et al. 2004). With varying driver importance across seasons and data extents (Figure 4.4 & 4.5) as well as the potential for variables to interact, it is difficult to attribute the random forest modelling results solely to training data point location. The under or over representation of certain driver classes (Table 4.2) does however provide a basis to guide future flight line planning, especially when combined with RF-based driver importance rankings.

Optimized LiDAR sampling with strategic flight line positioning could increase the viability of using LiDAR to monitor snowpack. Efficiency is also a function of the data processing and analysis steps. One area to reduce computation time in the modelling workflow is observed with forest size (Table 4.2). Creating too many or too few trees in a random forest can lead to over or under-fitting of models (Kuhn and Johnson, 2013) so it was important to test a range of values. R^2 and RMSE were unchanged by forest size, but the computation times varied. It was decided that 250

trees would be a conservative choice when predictions were made over the 103 km^2 watershed, without expending the time required for 500 or 1000 trees. 1000 trees used in another RF snow study that did not test various forest sizes (Lopez-Moreno et al., 2017). The faster production of models with integer formats and fewer points was expected, as it is as an established concept in spatial and visual data processing to reduce storage space/memory with certain data formats or types of compression (Bonham-Carter, 1994). Table 4.4 illustrates a RAM limitation when 100,000 training points were used, as the difference in run time for integer and raster input data is much smaller compared to these data types in the 50,000 point trials. The RMSE value for the 100,000 point don't follow the expected correspondence of a higher R^2 with lower RMSE values, and more computational power might clarify this issue. Converting float data to integer format using a multiplier is an ideal approach to time-saving without compromising end results in term of both R^2 and RMSE. While a larger sample size may increase correspondence between the LSDMs and RF SDMs, the significant correlations between these datasets (Table 4.5) suggest that the current sample size is suitable. If a stronger correlation was desired, training data point placement (ie. Ensuring a representative variety of terrain and cover attributes are sampled) and large sample sizes should be considered.

4.5.2 Snow Depth Driver Variables

4.5.2.1 Elevation

Elevation has consistently proven to be a valuable predictor of snow depth distribution (Revuelto, 2014b; Elder et al. 1998; Erxleben et al. 2002; Hopkinson et al., 2012; Molotch and Bales, 2005; Grunewald et al. 2015). Lopez-Moreno et al (2010) found that elevation increases in importance as grid size increases. Elevation was often a useful variable as indicated by MSE (Figure 4.4 and Figure 4.5). Its utility was pronounced at melt onset, which is likely influenced by spring warming and environmental lapse rates (Pigeon and Jiskoot, 2008) creating a negative temperature gradient from valley bottoms to mountain summits. Lower elevations are also subject to rain on snow events in spring time. Rain adds thermal energy to the snowpack (Garvelmann et al. 2015), enhancing melt in areas that may already be prone to reduced snow inputs relative to upper elevations, where precipitation may still be adding to the snowpack.

Snowpacks become more homogenous as melt conditions progress (Lopez-Moreno et al. 2017), so a variable such as elevation which operates over larger scales (1000s of meters) rather than micro-features, such as changes in terrain or canopy cover, explains much of the overall watershed-scale snow depth distribution. Snow water equivalent records indicate that the mid and late winter datasets were surveyed while snow was still accumulating (Figure 3.4), meaning that snow from upper elevations may have limited gravitational redistribution down slope relative to the end of the season, making elevation a less important driver than it was at melt onset. Wind

erosion and reduced shading by vegetation promotes lower and/or non-linear snow accumulation with elevation gains above treeline (Grunewald et al. 2014; Hopkinson et al., 2012; Zheng et al. 2016), and it is possible that these effects were more pronounced in the February 2014 and April 2016 datasets, increasing elevation's importance as a snow depth control. These observations suggest LiDAR sampling and RF modeling is most successfully implemented late in the winter season, after a period of snowpack settling and homogenisation, several days after fresh snowfall.

4.5.2.2 Aspect

Aspect consistently performed well across the datasets. South facing terrain often accumulates the least amount of snow (Hopkinson et al. 2012; Kirchner et al. 2014; Zheng et al. 2016) primarily due to high solar radiation inputs relative to north-facing slopes. Higher accumulation on north facing terrain can be the result of reduced radiative inputs due to daytime shadowing and solar azimuth (Anderson and West, 1965; Haupt, 1951). In a watershed where strong wind gusts originate from south and south-westerly directions, the potential for complimentary wind and radiation loading effects exist. Wind erosion on the south and southwest slopes could enhance the ablation of snow that is already subject to more radiation-induced melt than on opposite facing terrain and therefore lower depth values result (Table 3.8). A combination of spatially variable radiation budgets as well as wind transport (Hiemstra et al. 2002) could explain why aspect ranked highly, as both are well-established drivers of snow depth distributions. With aspect consistently in the top three important variables for all seasons and extents, the proportion of training points for aspect

classes compared to the WCW class sizes may have more influence on the predicted depths than less important drivers.

4.5.2.3 TPI

Topographic position is comparable to curvature which is also used as a snow depth predictor in some studies (e.g. Marchand and Killingtveit, 2005; Lopez-Moreno and Nogues-Bravo, 2006; Plattner et al., 2006). These variables represent localized, potentially down to sub-meter terrain surface features where snow can either be trapped in depressions or scoured off convexities. Hopkinson et al. (2004) found that LSDM values were greater in valleys and shallower on ridge tops, with open area distributions closely related to topography. Topographic smoothing is the filling of depressions and the spaces between short vegetation (Schirmer and Lehning, 2011), and these processes increase with snow accumulation and persistence throughout the winter season. As discussed in the aspect section above, a consistent variable in terms of importance such as TPI may have more influence if it is not proportionately represented in the training data.

4.5.2.4 Canopy Cover

The burial of low vegetation by wind redistribution and trapped snow, coupled with the decline in cover as elevation increases, could have contributed to canopy cover's low performance as a depth predictor. A study by D'Eon (2004) that examined

snow accumulation in open and forested areas along an elevation gradient found that depth was significantly correlated with elevation, yet canopy and snow accumulation were only significantly correlated at lower elevations. The author attributed this to greater accumulation at higher sites with less canopy cover as a function of lower temperatures reducing ablation, suggesting that the importance of canopy cover varies with elevation (D'Eon, 2004).

Other recent LiDAR-based research, using regression, has illustrated that up to 50% of snow depth distributions can be explained by canopy metrics (Zheng et al. 2018). The development of consistent relationships between cover and snow depth in this analysis may have been complicated by parameters that could not be adequately represented at the watershed scale. A likely contender in this environment is the effect of variable wind fields across complex terrain and canopies, as this is a source region for strong Chinook winds where speeds frequently exceed 100km/h (Pigeon & Jiskoot, 2008). Considering that high elevation relationships between depth and canopy are difficult to establish (D'Eon 2004), as well as the large amount of training data that falls within open and intermediate classes where depth distributions were more variable in the WCW (Figure 3.10), these concepts may explain the higher mean depths produced by FL-trained models.

4.5.2.5 Slope

The West Castle Watershed is mostly forested below treeline, with a variety of slopes throughout the entire ~1200m elevation range. Steep slopes are often observed

in alpine zones, where snow depth in the WCW is more variable than areas below treeline (Figure 3.7). Depending on the surrounding terrain and wind vectors, the top of a slope or cliff could be the location of a depth outlier due the formation of a cornice (Schweizer et al. 2003) or an area of little snow due to wind erosion and exposure to solar radiation (Varhola et al. 2010). Areas where wind loading and/or cornice formation occur can eventually lose the built-up snow due to downward redistribution of snow by natural avalanches and sloughing (Schweizer et al. 2003). While an open area at the top of a steep slope may be bare due to these effects, the same level of canopy cover at a lower elevation may translate to enhanced accumulation (Varhola et al. 2010). If a study area is limited to open canopies and gentle/moderate slopes, slope can be an important predictor of snow depth (Grunewald et al. 2013). These interacting effects as well as the influence of wind and gravity make it difficult to establish consistent terrain-depth relationships in steep areas throughout the winter season.

When comparing slope surfaces derived from airborne LiDAR it is also important to consider the amplification of vertical uncertainty over steeper slopes, as horizontal uncertainty is higher than in other areas and must be considered during quality control and analysis (Hodgson, 2004; Hodgson et al. 2005). The study area does not contain many grid cells with slopes >60˚ and most of these cells were eliminated by quality control. Future work may benefit from finding another study area with a different distribution of terrain slope values to determine if this driver should be kept in future modelling exercises. Including a wind parameter in modelling workflows may be another way to establish reliable snow depth trends in steep areas.

4.5.3 Watershed Snow Depth Predictions

Correlations between LSDMs and RF-predicted rasters (Table 4.5), indicate that using LiDAR flight line sample datasets to impute watershed-scale snow depth is viable. Mid and late winter trials did not perform as well as melt onset in terms of PCC or R^2 values. However, the PCC values were all significant and correspondence in watershed-scale mean depths is more important from a water resources inventory perspective (Table 4.5). Model performance results such as R^2 simply indicate that models are imperfect. Absolute error, as indicated by RMSE, was unchanged for the WCW trials though. With mid-winter datasets, more complex snow depth distribution patterns produce weaker correlations (Lopez-Moreno et al. 2017). Our lower R^2 mid-winter results as well as the conclusions of Lopez-Moreno et al (2017) could be influenced by snow that had yet to undergo further metamorphosis and redistribution relative to more settled snowpacks (Figure 3.4). By melt onset, much of the low-lying vegetation and topographic depressions are buried (Schirmer and Lehning, 2013), creating smoother surfaces for new snow to settle on top of rather than being influenced by surface features. Once new snow stops accumulating, aspect and canopy-influenced ablation processes, such as radiation and sensible heat (Golding and Swanson 1986; Anderson et al. 1958a), might be more pronounced over a settled snowpack relative to wind distribution processes on a freshly accumulated snow surface. Static (or grid-cell level) features such as elevation, aspect and canopy cover, which were the top three WCW trained variables at melt onset, can therefore be more easily separated at melt onset than during mid-winter when localized turbulence and wind redistribution of fresh snow may mask their influence.

Under mid-winter conditions, WCW-based training points resulted in a RF SDM with the same mean depth as the LSDM (Table 4.6), despite regression results that are not strong ($R^2 < 0.5$; Table 4.5). Late-winter R^2 was the lowest but the PCC for this season was still significant. Standard deviation values for the FL SDMs also suggest more variability from these models in most cases. RMSE is an absolute measure of error and of the variance that isn't explained by the model, whereas relative fit of the model is provided by R^2. RMSE was provided to one decimal place as two decimal places for units of meters of snow depth is beyond the expected accuracy of most LiDAR surveys (~10cm). Although the RMSE values would be expected to show an inverse relationship with R^2, they were not always in correspondence which illustrates the need for further analysis of sample sizes and the best seed to use for all datasets. In order to objectively assess variable order importance to guide future sampling design, the X, Y locations of input data was constant across the datasets. This compromises the representativeness of the training sample compared to the depth distribution of the original LSDM as the locations of data to produce a good sample are not necessarily constant year to year. Statistical power and appropriate sample sizes could benefit from further exploration as more variety in depths might be captured from larger samples but increase the chance of type 2 statistical errors (Kaplan et al., 2014). R^2 is more sensitive to oddities in the dataset therefore it is possible that with a different seed and a more representative sample, R^2 values would be more reliable and possibly increase (Grace-Smith, 2019). However, direct comparisons of variable order importance would be less valid if varying depth driver attributes were present in the various training datasets as a result of using different seeds to sample the raster stack, as we would be unable to separate the effects of varying terrain attributes from seasonality when assessing

variable order. For the purposes of this exploratory analysis, assessment of watershed scale means provides a simple metric that illustrates how a dense dataset with millions of data points can be modelled within reasonable accuracy of the original LSDMs.

It is clear from the mean depth results (Table 4.6) that the FL RF SDM outputs for the watershed are higher relative to the RF SDM or LSDM means. Given no such systematic over-estimation occurs for the watershed-scale RF SDM (Table 4.6), this bias must be the result of the sampling configuration. While all landcover and terrain types are represented in the training datasets (Table 4.2), the centre of the valley, where elevations <1700 m a.s.l. and closed canopy covers are greatest, is under-represented by FL samples which likely explains the higher mean depths predicted from FL training data. Improved FL model results would be facilitated through more strategic flight line sampling to represent the range of land surface attributes and snow depth driver classes experienced within the watershed.

While many studies have included a radiation parameter (e.g. see Table 4.1), researchers have found that using a canopy cover or forest density parameter is an acceptable proxy for transmittance and better represents radiative inputs than parameters calculated from a DEM that excludes forest cover (Davis et al 1997; Hardy et al 1997). Erxleben et al (2000) found using the BRT methodology that model performance was improved when the radiation parameter was excluded, so only vegetation, elevation, aspect and slope were employed in model development. Considering these findings, and the ultimate goal of an efficient modelling workflow, we limited the predictors to the five chosen in situ (grid cell-level) properties. This is a simplification of reality, as other external (beyond the grid cell) snowpack controls,

such as wind or radiation load are ignored due to such spatial and temporal continuous data not being readily available. Dynamic external drivers that vary in time and space can be simulated, however (Winstral et al. 2002; Molotch et al. 2005) but are expected to have variable predictive power and thus add non-deterministic complexity that may reduce the effectiveness of a machine learning snow depth extrapolation framework. Additional predictors may improve modelling results in future analyses.

Although Lopez-Moreno 2017 also used LiDAR and random forest, they applied the algorithm to determine variable order importance, not to predict depths. With R^2 values ranging from 0.41-0.61 (Table 4.4), the RF results presented here exceed others' model performance in some cases (Table 4.1). Grunewald et al. 2013 reported R^2 values of 0.27-0.90 for MLR based models using LiDAR data in various mountainous study areas. It is important to recognize that the grid cell resolution in this modelling study, 3m, is high and the results in Grunewald et al. (2013) include resolutions of 100m, 200m and 400m. With a high spatial resolution, it can be more difficult to draw a representative sample from the population because the population size is larger for a high-resolution dataset than it is with a coarser grid cell size. <1% of the available depth values were utilized in the 3m RF modelling process. Random Forest is advantageous for such dense datasets compared to MLR or other methods primarily because the algorithm splits the data and introduces randomness by variable permutation at tree nodes. The re-creation of watershed scale snow maps from spatially sampled datasets is a novelty of this work, and the results presented demonstrate there is potential for further refinement.

4.6 Conclusion

Random forest is a promising machine learning routine for modelling snow depth from airborne LiDAR sample data, especially if strategic flight paths are chosen. Integration of this algorithm into snowpack monitoring frameworks may require an annually adaptive approach. Although the variables aspect, TPI and elevation were often chosen for random sampling at node splits, their inconsistent ranking suggests that one variable cannot be permanently prioritized over others when designing sampling routines using random forest. Other studies have suggested that importance measures are site-specific (Grunewald et al. 2013), but importance appears to be specific to individual datasets/seasons as well based on our results in the West Castle Watershed. While melt onset yields the best model performance metrics (PCC and R^2), similar watershed depth means at mid-winter suggest the ideal time for collection of partial LiDAR datasets as inputs to integrated hydrological monitoring frameworks is not necessarily limited to a certain season if other metrics of model performance are considered. Modeling late-season snowpack is desirable as it is this snow, not what is present earlier in the accumulation season, which represents the stored freshwater resource that has the potential to influence subsequent downstream flood or drought hazards. However, if and when early or mid-winter watershed-scale snowpack conditions are desired, then LiDAR sampling and RF modeling should be implemented after snowpack settling and homogenisation, and not immediately following fresh snowfall.

Significant correlations between FL and WCW RF SDMs and LSDMS (Table 4.4) demonstrate that Random Forest has great potential for snow depth imputation as part

of an overall operational LiDAR snow depth monitoring framework. Collection of data over an optimally selected flight path that considers variable importance of drivers, as presented in this thesis (Section 4.6.2), should improve model performance when trained from spatial sample data. This is an objective that needs to be addressed to operationalize LiDAR-based snow monitoring in a cost- and time-effective manner. Further analysis is required to develop such a flight line sampling strategy. Using spatial sampling to train the RF routine is a viable approach to integrated snowpack/hydrological monitoring with airborne LiDAR, but more research is needed to optimize landcover and terrain sampling. This analysis shows that airborne LiDAR in tandem with random forest produces results that are comparable to, or in some cases better than, previous work using multiple linear regression and binary regression trees to model snow depth over large spatial scales. Overall, LiDAR continues to be a valuable snowpack monitoring tool and this research confirms that the ability to use it in a more time- and cost-efficient manner is feasible.

5 RESEARCH CONCLUSION

5.1 Summary of Research Purpose

The need for enhanced headwater snowpack monitoring in the face of a changing, intensifying climate provided the basis for this thesis research. Western Canada is heavily reliant on seasonal snow for runoff, especially considering the degree to which agriculture in southern Alberta is dependent on irrigation. The use of spatially coarse public monitoring sites as indices for water supply forecasts may not hold into the future due to the dynamic nature of the Earth's climate. In a watershed such as the South Saskatchewan, inter-provincial and international water sharing agreements on top of high water demand within semi-arid southern Alberta make a precise quantification of stored freshwater resources valuable for mitigation of risk associated with floods and drought, which have cost the province billions of dollars in the last decade alone.

5.2 Research Findings and Future Research

With a detailed index of snow depth distributions (Chapter 3) and random forest modelling (Chapter 4), snow depth driver consistency and importance was determined in the West Castle Watershed. Relationships between snow depth and elevation to treeline are consistent regardless of the time of year when data was collected. Within and above treeline, wind-redistributed snow from upper elevations to areas of closed canopy cover is evident. Aspect and TPI, drivers that exhibited consistent distributions inter-annually, were also identified among the top three most important variables in

Random Forest modelling. The reliability of these depth drivers for predicting snow depth distributions can guide future flight planning and sampling schemes for inter-annual and seasonal integrated monitoring frameworks. An increased spatiotemporal availability of hydrometeorological data in the study area would help determine if public monitoring sites could be integrated with LiDAR and Random Forest data. Future analyses may benefit from thorough field campaigns and integration of continuous hydrometeorological data. Further research on the effect of grid cell resolution and quality control options could potentially improve random forest model performance. The most important aspect of this workflow to improvement upon are matters of sample size and statistical power.

A novelty of this work is that a partial dataset (two flightlines) was used to predict snow depth at the watershed scale with a reasonable degree of accuracy and statistically significant correlation to the parent LSDM. Smaller standard deviations from RF-predicted SDMs compared to the LSDMs indicate RF under-estimated the variability of snow depth distributions, as is common with sample-based modeling of class means. Considering the size of the Rocky Mountains and the Canadian land base dependant on headwater regions for seasonal snow melt runoff, strategically planned LiDAR sampling and random forest modelling could provide spatially extensive datasets for water supply forecasts in a much more time and cost-efficient manner than a full survey. The methods presented here could be utilized strategically as part of an integrated framework that utilizes optimized data collection and modelling to produce high resolution snow depth models to aid water supply forecasts and risk management.

6 References

AEP, 2017. Data obtained directly from Alberta Environment and Parks.

Alili, Y., Kuras, P., Bewley, D., Teti, P. and Blair, T. (2009). Effects of pine beetle infestations and treatments on hydrology, integrating stand-level data into mesoscale watershed functions. Department of Forest Resources Management, University of British Columbia. Retrieved from, http://www.cfs.nrcan.gc.ca/pubwarehouse/pdfs/31182.pdf

Anderson, H.W. (1956). Forest-cover effects on snowpack accumulation and melt. Central Sierra Snow Laboratory. Transactions American Geophysical Union, 37(3), 307–312.

Anderson, H.W., Rice, R.M. and West, A.J. (1958). Snow in forest openings and forest stands. Proceedings of the Society of American Foresters, 46–50.

Anderson, H.W. and West, A.J. (1965). Snow accumulation and melt in relation to terrain in wet and dry years. Proceedings of the 33rd Western Snow Conference, 73–82.

Anderton, S.P. (2000). An analysis of spatial variability in snow processes in a high mountain catchment. PhD Thesis, University of Durham, UK, 251 pp.

Arnold, N. S. and W. G. Rees. (2003). Self-similarity in glacier surface characteristics. Journal of Glaciology, 49(8), 547–554.

Balk, B., and Elder, K. (2000). Combining binary decision tree and geostatistical methods to estimate snow distribution in a mountain watershed. Water Resources Research, 36(1), 13-26.

Baños, I. M., Garcia, A. R., Alavedra, J. M., Figueras, P. O., Iglesias, J. P., Figueras P. M. and Lopez, J. T. (2011). Assessment of Airborne LIDAR for Snowpack Depth Modelling. Boletin de la Sociedad Geologica Mexicana 63(1), 95–107.

Barilotti, A., Turco, S., and Alberti, G. 2006. LAI determination in forestry ecosystems by LIDAR analysis. Workshop on 3D Remote Sensing in Forestry, BOKU Vienna.

Breiman, L., Friedman, J., Stone, C. J., and Olshen, R.A. (1984). Classification and Regression Trees. Oxford, Taylor and Francis.

Berndt, H. (1965). Snow accumulation and disappearance in lodgepole pine clearcut blocks in Wyoming. Journal of Foresty, 63(2), 88–91.

Bhardwaj, A., Sam, L., Bhardwaj, A. and Javier Martin-Torres, F. (2016). LiDAR remote sensing of the cryosphere, Present applications and future prospects. Remote Sensing of the Environment, 177, 125-143.

Bonham-Carter, D. (1994). Geographic Information Systems for Geoscientists, Volume 13. Ottawa, Pergamon. 416 pp.

Boon, S. (2007). Snow accumulation and ablation in a beetle-killed pine stand in northern interior British Columbia. BC Journal of Ecosystem Management, 8(3), 1–13.

Boon, S. (2009). Snow ablation energy balance in a dead forest stand. Hydrological Processes, 23(18), 2600-2610.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

Brooks, K.N., Folliott, P.F., Gregersen, H.M., and DeBano, L.F., (2003). Hydrology and the Management of Watersheds, 3rd Edition. IA, Iowa State Press Ames.

Butler, 2018. Retrieved from: https://www.studyblue.com/notes/note/n/butler-summer-exam-ii/deck/17578227

Byrne, J. M., Kienzle, S. Johnson, D. G., Duke, Gannon, V., Selinger, B. and Thomas, J. (2006). Current and future water issues in the Oldman River Basin, Alberta, Canada. Water Science and Technology, 53, 327-334.

Cartwright, 2018. Retrieved from: https://github.com/KelCbells/MSc_Thesis

CBC News, 2015. Retrieved from: http://www.cbc.ca/news/canada/calgary/agriculture-disaster-declared-in-alberta-after-losses-from-extreme-weather-1.3200420

Calgary Herald, 2013. Retrieved from: http://www.calgaryherald.com/news/Province+boosts+cost+Alberta+floods+billion/8952392/story.html

Chang, K. T. and Li, Z. (2000). Modelling snow accumulation with a geographic information system. International Journal of Geographical Information science, 14(7), 693-707.

Chignell, S. M., Luizza, M. W., Skach, S., Young, N. E and Evangelista, P. H. An integrative modeling approach to mamping wetlands and riparian areas in a heterogeneous Rocky Mountain Watershed. Remove Sensing in Ecology and Conservation, 2, 150-165.

Currier, W. R. and Lundquist, J. D. 2018. Snow depth variability at the forest edge in multiple climates in the western United States. Water Resources Research 22553, 49pp.

Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., and Lawler, J. T. (2007). Random Forests for Classification in Ecology. Ecology, 88(11), 2783-2792.

Dickinson, W.T., and H.R. Whiteley, (1972). A sampling scheme for shallow snowpacks. IAHS Bulletin, 17(3): 247-258.

Daugharty, D., and Dickinson, B. (1982). Snow cover distribution in forested and deforested landscapes in New Brunswick, Canada. Proceedings of the 39th Eastern Snow Conference, 10–19.

DeBeer, C. M. and Pomeroy, J. W. (2010) Simulation of the snowmelt runoff contributing area in a small alpine basin. Hydrology and Earth Systems Science, 14, 1205-1219.

Deems, J.S., Fassnacht, S. and Elder, K. (2006). Fractcal distribution of snow depth from LiDAR data. Journal of Hydrometerology, 7(2), 285-297.

Deems, J., Painter, T. and Finnegan, D. (2013). LiDAR measurement of snow depth, a review. Journal of Glaciology 59(215), 467-479.

D'Eon, R. (2004). Snow depth as a function of canopy cover and other site attributes in a forested ungulate winter range in southeast British Columbia. BC Journal of Ecosystem Management, 3, 1–9.

Dunford, E.G. and Niederhof, C.H. (1944). Influence of aspen, young lodgepole pine, and open grassland types upon factors affecting water yield. Journal of Forestry, 42(9), 673– 677.

Elder, K., Dozier, J. and Michaelsen, J. (1991). Snow accumulation and distribution in an alpine watershed. Water Resources Research, 27, 1541-1552.

Elder, K., Michaelsen, J., and Dozier, J. (1995). Small basin modelling of snow water equivalence using binary regression tree methods. Symposium on Biogeochemistry of Seasonally Snow-Covered Catchments. IAHS Publication 228, 129–139.

Elder, K., Rosenthal, W., and Davis, R. E. (1998). Estimating the spatial distribution of snow water equivalence in a montane watershed. Hydrological Processes, 12, 1793–1808.

Essery, R.L., Pomeroy, J.W., Parvianen, J., and Storck, P., (2003). Sublimation of snow from coniferous forests in a climate model. Journal of Climate 16, 1855–1864.

Essery, R.L., Bunting, P., Hardy, J., Link, T., Marks, D., Melloh, R., Pomeroy, J., Rowlands, A., and Rutter, N. (2008). Radiative transfer modelling of a coniferous canopy characterized by airborne remote sensing. Journal of Hydrometeorology 9, 228– 241.

Erxleben, J., Elder, K. and Davis, R. (2002). Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains. Hydrological Processes 16(18), 3627-3649.

Fang, X. and Pomeroy, J.W. (2016). Impact of antecedent conditions on simulations of a flood in a mountain headwater basin. Hydrological Processes 30, 2754-2772.

Figliuzzi, P. (2002). South Saskatchewan River Sub-Basin Contributions To International and Interprovincial Water-Sharing Agreements. Retrieved from: http://aep.alberta.ca/water/programs-and-services/south-saskatchewan-river-basin-water-information/studies/documents/Water-SharingAgreements-SSRB-Oct2002.pdf

Fritze, H., Stewart, I. T., and Pebesma, E. Shifts in Western North American Snowmelt Runoff Regimes for the Recent Warm Decades. Journal of Hydrometeorology, 12, 989-

1006.

Geist, T. and Stotter, J. (2008). Documentation of glacier surface elevation change with multi-temporal airborne LiDAR laser scanner data – Case Study: Hintereisferner and Kesselwandferner, Tyrol, Austria. Z. Gletscherkd. Glazialgeol, 41, 77-106.

Garvelmann, J., Pohl, S. and Weiler, M. (2015). Spatio-temporal controls of snowmelt and runoff generation during rain-on-snow events in a mid-latitude mountain catchment. Hydrological Processes, 29, 3649-3664.

Golding, D. L. (1974). The correlation of snowpack with topography and snowmelt runoff on Marmot Creek Basin, Alberta. Atmosphere, 12, 31–38.

Golding, D., and Swanson, R. (1986). Snow distribution patterns in clearings and adjacent forest. Water Resources Research, 22(13), 1931–1940.

Government of Alberta. (2015). Retrieved from: http://aep.alberta.ca/water/programs-and-services/south-saskatchewan-river-basin-water-information/default.aspx

Grace-Smith, K. (2019). Retrieved from: https://www.theanalysisfactor.com/assessing-the-fit-of-regression-models/

Greene, E. M., Liston, G. E., and Pielke, R. A. (1999). Simulation of above treeline snowdrift formation using a numerical snow-transport model. Cold Regions Science and Technology, 30, 135-144.

Grunewald, T., Buhler, Y. and Lehning, M. 2014. Elevation dependency of mountain snow depth. The Cryosphere 8, 2381-2394.

Grünewald, T., Schirmer, M., Mott R. and Lehning, M. (2010). Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment. The Cryosphere, 4, 215-225.

Grunewald, T., Stotter, J., Pomeroy, J.W., Dadic, R., Banos, I. M., Marturia, J., and Lehning, M. (2013). Statistical modelling of the snow depth distribution in open alpine terrain. Hydrology and Earth Systems Science, 17(8), 3005-3021.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). The Elements of Statistical Learning. New York, Springer.

Haupt, H.E. (1951). Snow accumulation and retention on ponderosa pine lands in Idaho. Journal of Foresty, 49(12), 869–871.

Hedstrom, N. and Pomeroy, J. (1998). Measurements and modelling of snow interception in the boreal forest. Hydrologic Processes, 12, 1611–1625.

Hendrick, R.L., Filgate, B.D., and Adams, W.M. (1971). Application of environment analysis to watershed snowmelt. Journal of Applied Meteorology, 10, 418–429.

Hiemstra, C.A., Liston, G.E., and Reiners, W.A. (2002). Snow redistribution by wind and interactions with vegetation at upper treeline in the Medicine Bow Mountains, Wyoming, USA. Arctic and Antarctic Alpine Research, 34, 262–273.

Hodgson, M. E., and Bresnahan, P. (2004). Accuracy of Airborne LiDAR-Derived Elevation. Photogrammetric Engineering and Remote Sensing, 70(3), 331-339.

Hodgson, M.E., Jensen, J., Raber, G., Tullis, J., Davis, B.A., Thompson, G., and Schuckman, K. (2005). An evaluation of LiDAR-derived elevation and terrain slope in leaf-off conditions. Photogrammetric Engineering and Remote Sensing, 71(7), 817- 823.

Hopkinson, C., Pomeroy, J., DeBeer, C., Ellis, C., and Anderson, A. (2011), Relationships between snowpack depth and primary LiDAR point cloud derivatives in a mountainous environment. Proceedings of Remote Sensing and Hydrology Symposium, 352.

Hopkinson, C., Sitar, M., Chasmer, L.E. and Treitz, P. (2004). Mapping snowpack depth beneath forest canopies using airborne LIDAR. Photogrammetric Engineering and Remote Sensing, 70, 323-330.

Hopkinson, C., and Demuth, M.N. (2006). Using airborne LiDAR to assess the influence of glacier downwasting on water resources in the Canadian Rocky Mountains. Canadian Journal of Remote Sensing, 32(2), 212–222.

Hopkinson, C., Collins, T., Anderson, A., Pomeroy, J. and Spooner, I. (2012). Spatial Snow Depth Assessment Using LiDAR Transect Samples and Public GIS Data Layers in the Elbow River Watershed, Alberta. Canadian Water Resources Journal, 37(2), 69-87.

Hosang, J., and Dettwiler, K. (1991). Evaluation of water equivalent of snow cover map in a small catchment area using a geostatistical approach. Hydrological Processes 5(3): 283-290.

Ishwaran, H. (2007). Variable importance in binary regression trees and forests. Electronic Journal of Statistics, 1, 519-537. DOI: 10.1214/07-EJS039.

Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., Timlin, D. J., Shim, K., Gerber, J. S., Reddy, V. R. and Kim, S. (2016). Random Forests for Global and Regional Crop Yield Predictions. PlosOne, 11(6), 15 pp.

Jonas, T., Marty, C., and Magnusson, J. (2009). Estimating the snow water equivalent from snow depth measurements in the swiss alps. Journal of Hydrology, 378, 161-167.

Jost, G., Weiler, M., Gluns, D. and Alila, Y. (2007). The influence of forest and topography on snow accumulation and melt at the watershed-scale. Journal of Hydrology 347:101–115.

Kaplan, R., Chambers, D. and Glasgow, R. Big Data and Large Sample Size: A Cautionary Note on the Potential for Bias. Clinical and Translational Science, 7(4), 342-346.

Kienzle, S. (2004). The effect of DEM raster resolution on first order, second order and compound terrain derivatives. Transactions in GIS, 8(1), 83-111.

Kienzle, S. W., and Mueller, M. (2013). Mapping Alberta's surface water resources for the period 1971-2000. Canadian Geographer, 57(4), 506–518. https://doi.org/10.1111/j.1541-0064.2013.12050.x

Kirchner, P. B., Bales, R. C., Molotch, N. P., Flanagan, J. and Guo, Q. (2014). LiDAR measurement of seasonal snow accumulation along an elevation gradient in the southern Sierra Nevada, California. Hydrology and Earth System Sciences 18(10), 4161-4275.

Kuhn, M. and Johnson, K. (2013). Applied Predictive Modelling. New York, Springer.

Kuz'min, P. P. (1960). Snowcover and snow reserves. Gidrometeorologicheskoe, Izdatelsko, Leningrad. [Translation], U.S. National Science Foundation, Washington, D.C.

Lehning, M., Grünewald, T. and Schirmer, (2011). M. Mountain snow distribution governed by an altitudinal gradient and terrain roughness. Geophysical Research Letters, 38, 5 pp.

Liaw, A. and Wiener, M. (2018). Retrieved from: https://cran.r-project.org/web/packages/randomForest/randomForest.pdf

Litaor, M. I., Williams, M., and Seastedt, T. R. (2008). Topographic controls on snow distribution, soil moisture, and species diversity of herbaceous alpine Vegetation, Netwot Ridge, Colorado. Journal of Geophysical Research: Biogeosciences, 113(2), 1–10.

Lopez-Moreno, J.I. (2005). Recent variations of snowpack depth in the central Spanish Pyrenees. Arctic, Antarctic and Alpine Research, 37, 253-260.

López-Moreno, J. I. and Nogués-Bravo, D. (2005). Mapping the spatial distribution of snow pack in the Spanish central Pyrenees. Hydrological Processes, 19, 3167–3176.

Lopez-Moreno, J. I. and Nogues-Bravo, D. (2006). Glacier development and topographic context. Earth Surface Processes and Landforms 31(12), 1585-1594.

López-Moreno, J.I., Latron J., and Lehmann, A. (2010). Effects of sample and grid size on the accuracy and stability of regression-based snow interpolation methods. Hydrological Processes 24(14), 1914–1928.

Lundquist, J.D., Cayan, D.R., and Dettinger, M.D. (2004). Spring onset in the Sierra Nevada, when is snowmelt independent of elevation? Journal of Hydrometeorology, 5, 327–334.

Lundquist, J.D., and Flint, A.L. (2006). Onset of snowmelt and streamflow in 2004 in the western United States, how shading may affect spring streamflow timing in a warmer world. Journal of Hydrometeorology, 7, 1199–1271.

Marchand, W. D. and Killingtveit, A. (2005). Statistical probability distribution of snow depth at the model sub-grid cell spatial scale. Hydrological Processes, 19, 355–369.

Marshall, H.P., Conway, H., and Rasmussen, L.A. (1999). Snow densification during rain.

Cold Regions Science and Technology, 30, 35-41.

Moeser, D., Stahli, M., and Jonas, T. (2015). Improved snow interception modelling using canopy parameters derived from airborne LiDAR data. Water Resources Research, 51, 5041-5059.

Molotch, N. P. and Bales, R.C. (2005). Scaling snow observations from the point to the grid element, Implications for observation network design. Water Resources Research 41(11), W11421, doi: 10.1029/2005WR004229.

Molotch, N. P., Colee, M.T., and Bales, R.C. (2005). Estimating the spatial distribution of snow water equivalent in an alpine basin using binary regression tree models, the impact of digital elevation data and independent variable selection. Hydrological Processes 19(7), 145901479.

Moore, R. D., and Wondzell, S. M. (2005). Physical Hydrology and the Effects of Forest Harvesting in the Pacific northwest, A Review. Journal of the American Water Resources Association, 41(4), 763-784.

Moran-Tejeda, E., Lopez-Moreno, J.I., and Beniston, M. (2013a). The changing roles of temperature and precipitation on snowpack variability in Switzerland as a function of altitude. Geophysical Research Letters, 40, 2131-2136.

Moran-Tejeda, E.S, Herrera, J. I., Lopez-Moreno, J.I., Revuelto, J., Lehmann, A., and Beniston, M. (2013b). Evolution and frequency (1970-2007) of combined temperature-precipitation modes in the Spanish mountains and sensitivity of snow cover. Regional Environmental Change, 134, 873-885.

Mote, P., Snover, A. K., Capalbo, S., Eigenbrode, S. D., Glick, P., Littell, J., Raymondi, R., and Reeder, S. (2014). Chapter 21, Northwest. Climate Change Impacts in the United States, The Third National Climate Assessment. U.S. Global Change Research Program, 487-513. doi:10.7930/J04Q7RWX.

Mott, R. Scipion, D., Schneebeli, M., Dawes, N., Berne, A. and Lehning, M. (2015). Orographic effects on snow deposition patterns in mountainous terrain. Journal of Geophysical Research, Atmospheres, 119, 1419-1439.

Murray, C., and Buttle, J. (2003). Impacts of clearcut harvesting on snow accumulation and melt in a northern hardwood forest. Journal of Hydrology, 3, 197–212.

Nogues-Bravo, D. (2003) El estudio de la distribución espacial de la biodiversidad, conceptos y métodos. Cuadernos de Investigación Geográfica, 29, 67–82

Oke., T.R. (1987). Boundary layer climates (Second edition). London, Methuen.

Oldman Watershed Council. (2010). State of the Watershed Report. Retrieved from, http://oldmanwatershed.ca/publications-list/state-of-the-watershed

Pelt, W., Phjola, V., and Reigmer, C. (2016). The changing impact of snow conditions and refreezing on the mass balance of an idealized Svalbard glacier. Frontiers in Earth Science, 25, doi.org/10.3389/feart.2016.00102.

Pigeon, K. E., and Jiskoot, H. (2008). Meteorological Controls on Snowpack Formation and Dynamics in the Rocky Mountains. Arctic, Antarctic and Alpine Research, 40(4), 716-730.

Plattner, C., Braun, L., and Brenning, A. (2006). Spatial variability of snow accumulation on Vernagtferner, Austrian Alps, in winter 2003/04, Z. Gletscherk. Glazialgeol., 39, 43–57.

Pomeroy, J.W., Parviainen, J., Hedstrom, N., and Gray, D.M. (1998). Coupled modelling of forest snow interception and sublimation. Hydrologic Processes, 12, 2317–2337.

Pomeroy, J. W. and Brun, E. (2001). Physical properties of snow. Snow Ecology: An Interdisciplinary Examination of Snow-covered Ecosystems, Cambridge UK, 45-118.

Pomeroy, J.W., Marsh, P., and Gray, D.M. (1997). Application of a distributed blowing snow model to the Arctic. Hydrological Processes, 11(98), 1451-1464.

Pomeroy, J. W., Gray, D. M., Hedstrom, N. R., and Janowicz, J. R. (2002). Prediction of seasonal snow accumulation in cold climate forests. Hydrological Processes, 16(18), 3543-3558.

Pugh, E. and Gordon, E. (2012). A conceptual model of water yield effects from beetle-induced tree death in snow-dominated lodgepole pine forests, Hydrological Processes, DOI: 10.1002/hyp.9312

Lopez-Moreno, J. I., Revuelto, J., Alonso-Gonzalez, E., Sanmiguel-Vallelado, A., Fassnacht, S. R., Deems, J., and Moran-Tejeda, E. (2017) Using very long-range terrestrial laser scanner to analyze the temporal consistency of the snowpack distribution in a high mountain environment. Journal of Mountain Science, 14(5), 823-842.

Revuelto, J., Lopez-Moreno, J. I., Azorin-Molina, C., and Vicente-Serrano, S. M. (2014b). Topographic control of snowpack distribution in a small catchment in the central Spanish Pyrenees, Intra- and inter- annual persistence. Cryosphere Discussions, 8(2), 1937–1972.

Roe, G.H., and Baker, M. B. (2006) Microphysical and Geometrical Controls on the Pattern of Orographic Precipitation. Journal of the Atmospheric Sciences, 63, 861-880.

Rood, S.B., Pan, J., Gill, K.M., Franks, C.G., Samuelson, G.M., and Shepherd, A. (2008). Declining summer flows of Rocky Mountain rivers, changing seasonal hydrology and probable impacts on floodplain forests. Journal of Hydrology, 349, 397-410.

Rowland, J.D., and Moore, R.D. (1992). Modelling solar irradiance on sloping surfaces under leafless deciduous forests. Agricultural and Forest Meteorology, 60, 111–132.

Schweizer, J., Jamieson, J. B., and Schneebeli, M. (2003). Snow avalanche formation. Revised Geoghysics, 40(4), 1016, doi:10.1029/2002RG000123.

Schindler, D.W, and Donahue, W.F. (2006). An impending water crisis in Canada's

prairie provinces. PNAS, 103(19), 7210-7216.

Schirmer, M., and Lehning, M. (2011). Persistence in intra-annual snow depth distribution, Fractal analysis of snow depth development. Water Resources Research, 47, doi, 10.1029/2010WR009426.

Shi, X., and Durran, D.R. (2013). The Response of Orographic Precipitation over Idealized Midlatitude Mountains Due to Increases in CO2. Journal of Climate, 27, 3938-3956.

Shook, K., and Gray, D. (1994). Determining the snow water equivalent of shallow prairie snow covers. Presented at 51st Eastern Snow Conference, 89-95.

Shook, K., and Gray, D. (1996). Small-scale spatial structure of shallow snowcovers. Hydrological Processes, 10(10), 1283-1292.

Sospedra-Alfonso, R., Melton, J.R., and Merryfield, W.J. (2015). Effects of temperature and precipitation on snowpack variability in the Central Rocky Mountains as a function of elevation. Geophysical Research Letters, 42, 4429-4438.

Steppuhn, H.W. (1976). Areal water equivalents for prairie snowcovers by centralized sampling. Proceedings of the Western Snow Conference, 44, 63-68.

Stewart, I. T., Cayan, D. R., and Dettinger, M. D. (2004). Changes in snowmelt runoff timing in western North America under a “business as usual” climate change scenario. Climate Change, 62, 217–232.

Strobl, C., Boulesteiz, A., Zeileis, A., and Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics, 8(25). doi: 10.1186/1471-2105-8-25.

Teti, P. (2008). Effects of Overstory Mortality on Snow Accumulation and Ablation. Ministry of Forests and Range, Williams Lake, British Columbia. 41 pp.

Tierney, L. T., Rossini, A. J., Li, N., and Sevcikova, H. (2018). Retrieved from: https://cran.r-project.org/web/packages/snow/snow.pdf

Tinkham, W.T., Smith, A, Marshall, H. P., Link, T. E., Falkowski, M. J., and Winstral, A. H., (2014). Quantifying Spatial Distribution of Snow Depth Errors from LiDAR Using Random Forest. Remote Sensing of Environment, 141, 105-115.

Toews, D.A., and Gluns, D.R. (1986). Snow accumulation and ablation on adjacent forested and clearcut sites in southeastern British Columbia. In, Proceedings of the 54th Western Snow Conference, 101–111.

Trujilo, E., Ramierez, J.A. and Elder, K.J. (2007). Topographic, meteorologic, and canopy controls on the scaling characteristics of the spatial distribution of snow depth fields. Water Resources Research, 43, 17pp.

Varhola., A., Coops, N.C., Weiler, M., and Moore, R.D. (2010). Forest canopy effects on snow accumulation and ablation, An integrative review of empirical results. Journal of Hydrometeorology 392, 219-233.

Veatch, W., Brooks, P.D., Gustafson, J.R., and Molotch, N.P. (2009). Quantifying the effects of forest canopy cover on net snow accumulation at a continental, mid-latitude site. Ecohydrology, 2(2), 115-128.

Wang, S., Deng, J., Chen, M., Weatherford, M. and Paugh, L. (2015). Random Forest Classification and Automation for Wetland Identification based on DEM Derivatives.

Winkler, R., and Roach, J. (2005). Snow accumulation in BC's southern interior forests. Stream Water Management Bulletin, 9(1), 1–5.

Winstral, A., Elder, K., and Davis, R.E. (2002). Spatial Snow Modelling Of Wind-Redistributed Snow Using Terrain-Based Parameters, 3, 524-538.

Woods, S., Ahl, R., Sappington, J., and McCaughey, W. (2006). Snow accumulation in thinned lodgepole pine stands, Montana, USA. Forest Ecology Management, 235, 202–211.

Zheng, Z., Kirchner, P.B., and Bales, R. C. (2016). Topographic and vegetation effects on snow accumulation in the southern Sierra Nevada, a statistical summary from LiDAR data. The Cryosphere, 10, 257-269.

Zheng, Z., Qin, M., Qian, K. and Bales, R. C. (2018). Canopy effects on snow accumulation: observations from lidar, canonical-view photos, and continuous ground measurements from sensor networks. Remote Sensing, 10(11), 1796.

APPENDIX A – SCRIPTS

Chapter three explains LSDM quality control and snow depth driver class creation using python scripts. The resulting classes were summarized and statistical analyses were employed in R. Chapter four provides a random forest analysis to determine variable order importance, make snow depth predictions and calculate correlation statistics in R. The scripts for both programs are available on Github, where files are numbered in the order they were utilized.

Scripts are maintained here:

https://github.com/KelCbells/MSc_Thesis