

Snow Depth Retrieval from Wide Band Radar in Trail Valley Creek

by

Wei Wang

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Geography

Waterloo, Ontario, Canada, 2022

© Wei Wang 2022

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Snow is an essential component in hydrology systems, and more specifically, monitoring variations in snow cover can provide valuable information about water supply, wildlife habitats, and climate changes. In recent years, the potential of wide band snow radar (e.g. 2-8 GHz) has been discovered with a series of campaigns in Operation IceBridge (OIB). Due to data from OIB mainly focusing on sea ice, most of the algorithms were also developed for snow on sea ice. As a result, this thesis aimed to test the applicability of the interface-based pulse peakiness snow depth retrieval method to snow on land. In addition, due to the common usage of radar altimeter in sea ice classification, this thesis also explored the possibility of adapting some of the ideas in sea ice classification to develop another retrieval method. Both approaches were tested on the 6 major vegetation types (tree, tall shrub, riparian shrub, dwarf shrub, tussock, and lichen) in the study area. Snow depth derived from Airborne Laser Scanner (ALS) point clouds was used as the reference for snow depth retrieval. Running the recalibrated pulse peakiness algorithm yielded a Mean Absolute Error (MAE) of 12 cm, 27 cm, 29 cm, 13 cm, 9 cm, and 10 cm for tree, tall shrub, riparian shrub, dwarf shrub, tussock and lichen respectively.

It was concluded that the principles behind the pulse peakiness approach is valid for snow on land. The presence of surface vegetation and the hummocky terrain of the study area did present some considerable challenges. In comparison, the classification approach using K-means while produced some accurate results under specific situations, was not as robust as the existing pulse peakiness approach. Although, it was argued that the classification approach was in early stages and there were some potential for better results.

Acknowledgements

Firstly, I want to thank my supervisor, Dr. Richard Kelly for providing guidance and assistance throughout my Master's degree, especially throughout the second year when COVID hit. His patience and knowledge of the field are extremely beneficial for helping me navigate through the process of finding a topic that I am interested in, then working on the topic, and finally writing it up as a thesis. I am very grateful for having him as my mentor.

Secondly, I want to express my appreciation for all the help from Dr. Josh King at Environment and Climate Change Canada for access to the field data, and for his introduction of the work done using the wide band radar. His familiarity of the data and his knowledge of the radar system helped me get started with what was then a fairly new type of data. In addition, a sincere thank you goes to Dr. Julia Boike and Dr. Inge Grünberg who provided crucial information about their work on the vegetation map of the study area. Also, I would like to thank Dr. Christian Haas and Dr. Stephen Hendricks at the Alfred Wegener Institute for Polar and Marine Research for access to the airborne ultra-wideband radar data.

Thirdly, I would like to thank other members of my research group, especially Aaron who asked some great questions about my research during group meetings and offered some extremely useful advice. Also, I want to express my gratitude towards my friends and family, without whose support and encouragement, I would not be where I am today.

Table of Contents

List of Figures	ix
List of Tables	xii
1 Introduction	1
1.1 Context and Motivation	1
1.2 Aims and Objectives	3
1.3 Thesis Structure	4
2 Background	5
2.1 Snowpack Physical Properties	5
2.2 Wide Band Radar and Its Interaction with Snow	8
2.3 Characterizing Wide Band Radar Waveforms	13
2.4 Existing Algorithms for Wide Band Radar Snow Depth Retrieval	17

2.4.1	Kurtz et al. (2013) algorithm	17
2.4.2	Kurtz et al. (2014) algorithm	18
2.4.3	Koenig et al. (2016) algorithm	19
2.4.4	Newman et al. (2014) algorithm	20
2.4.5	Kwok and Maksym (2014) algorithm	21
2.4.6	Julita et al. (2021) algorithm	22
3	Snow Depth Retrieval from Wide Band Radar in Trail Valley Creek	23
3.1	Introduction	23
3.2	Study Area	27
3.3	Datasets	29
3.3.1	Snow Radar Data and Airborne Laser Scanning (ALS) Data from Icebird 2019 Winter	29
3.3.2	ALS Point Clouds From 2016	33
3.3.3	Field Snow Depth Probe Measurements	35
3.3.4	TVC Vegetation Map	38
3.4	Methods	40
3.4.1	Pre-processing	40
3.4.2	Evaluation of Lidar Snow Depth Product	42
3.4.3	Estimation of Snow Depth Using the Pulse Peakiness Approach	43

3.4.4	Estimation of Snow Depth Using a Classification Approach	47
3.4.5	Evaluation and Comparison of Both Approaches	51
3.5	Results	51
3.5.1	Evaluation of Lidar Snow Depth Product	51
3.5.2	Pulse Peakiness Approach Results	54
3.5.3	Classification Approach Results	59
3.5.4	Comparison Between the Two Approaches	62
3.6	Discussion and Conclusion	65
3.6.1	Underestimation and Overestimation	65
3.6.2	Error of Omission for the Pulse Peakiness Approach	70
3.6.3	Conclusion	72
4	Discussion and Conclusion	73
4.1	Summary	73
4.2	Methodology Limitations	75
4.3	Recommendations	77
4.4	Conclusion	79
	References	81
	References	81

APPENDICES	89
A Full Optimization Results for the Pulse Peakiness Approach	90
B Full Results for the K-means Approach	97

List of Figures

2.1	Main components of snowpack backscatter (Rott et al., 2010)	7
2.2	Sawtooth modulation (EuroMotor virtual college., n.d.)	9
2.3	Time delay between the transmitted and received signal (Wolff, n.d.) . . .	9
2.4	A simplified wide band radar system (Mostafa, 2019)	10
2.5	2-peak waveform demonstrating the interface impulse response (Farrell et al., 2012)	12
2.6	Ambiguity between air-snow and snow-ice interface (Galin et al., 2012) . .	12
2.7	SARAL radar altimeter waveforms for ocean, lead and sea ice (Müller et al., 2017)	13
2.8	Definitions of LeW and TeW (Zygmuntowska et al., 2013)	15
3.1	Landsat 8 images of study area; left is from 2019.06.18 and right is from 2019.04.15. Red triangle represents the location of TVC research station .	29
3.2	Location of the used radar data points colored by flight elevation (Landsat 8 from 2019.04.15 as background)	32

3.3	2016 snow off DTM, max vegetation height and mean vegetation height maps	34
3.4	Locations of used field measurement (Landsat 8 from 2019.04.15 as back-ground)	36
3.5	Histogram distribution of snow depth measurements (collected between 2019.03.21 and 2019.03.26)	37
3.6	TVC vegetation map (Boike & Grünberg, 2019).	39
3.7	Interference in the beginning and final bins	41
3.8	Matching field snow depth and lidar snow depth	43
3.9	Example target peak defined by the closet eligible log peak to the lidar reference air-snow interface (waveform shown is from tussock)	46
3.10	Workflow for every waveform in a cluster illustrating how lidar snow depth is used as the reference and how cluster statistics is calculated	50
3.11	Lidar snow depth product for the study area	52
3.12	Scatterplot comapring lidar and field snow depth	53
3.13	Pulse peakiness snow depth variation for each vegetation type.	55
3.14	Scatterplots bwteen pulse peakiness and lidar snow depth estimates for the following classes a) tree, b) tall shrub, c) riparian shrub, d) dwarf shrub e) tussock and f) lichen.	58
3.15	Scatterplots between Classification and Lidar snow depth estimates for the following classes a) tree, b) tall shrub, c) riparian shrub, d) dwarf shrub e) tussock and f) lichen.	61

3.16	Classification snow depth variation for each vegetation type.	62
3.17	Comparison between classification (K-means) and Pulse Peakiness snow depth	64
3.18	Comparison between classification (K-means) and Lidar snow depth for those points with a valid Pulse Peakiness output	65
3.19	Sample underestimating waveforms for pulse peakiness where the Lidar ref- erence air-snow interface falls below the logarithmic threshold	66
3.20	Before and after bias correction	68
3.21	Sample overestimation for dwarf shrubs where there are multiple peaks close to the maximum peak.	69
3.22	Sample overestimation for lichen where there are multiple peaks close to the maximum peak.	70
3.23	Comparison of snow depth distribution between Pulse Peakiness valid out- put and Lidar points	71

List of Tables

3.1	Details of three field campaigns	35
3.2	Mean and median snow depth and standard deviation for each vegetation type	37
3.3	Number of radar points and K-means clusters for each vegetation type . .	49
3.4	Evaluation of lidar snow depth	54
3.5	Optimal pulse peakiness parameters for each vegetation type	55
3.6	Best Snow Depth Retrieval Results	60
3.7	Comparison of valid points percentage between pulse peakiness and classi- fication approach	63
3.8	Distribution statistics for Lidar snow depth and pulse peakiness valid output	71
A.1	Pulse peakiness parameters optimization results for trees	91
A.2	Pulse peakiness parameters optimization results for tall shrubs	92
A.3	Pulse peakiness parameters optimization results for riparian shrubs	93

A.4	Pulse peakiness parameters optimization results for dwarf shrubs	94
A.5	Pulse peakiness parameters optimization results for tussocks	95
A.6	Pulse peakiness parameters optimization results for Lichens	96
B.1	K-means approach results for trees	97
B.2	K-means approach results for tall shrubs	98
B.3	K-means approach results for riparian shrubs	98
B.4	K-means approach results for dwarf shrubs	98
B.5	K-means approach results for tussocks	99
B.6	K-means approach results for lichens	99

Chapter 1

Introduction

1.1 Context and Motivation

Snow and its associated physical processes like snow melt play a significant role in global climate and can have serious environmental as well as economical impact (Sturm et al., 2017). In mid-latitude and polar regions most of the precipitation events are associated with the production of snow in clouds (Field and Heymsfield, 2015). In addition to being involved with precipitation, snow is capable of modulating river flow patterns and serves as a source of clean water (Biemans et al., 2019). Economically, Barnett et al. (2015) explained that snowpacks and glaciers provide water supply for more than one sixth of the planet's population, and recent hydrological changes could heavily affect water availability. In China, snow scarcity has caused numerous spring droughts leading to crop failure and affected the water flow in Yellow River, the second largest river in China (Dahe et al.,

2006).

The importance of snow and the necessity for snow monitoring has led to scientists using different approaches to quantify snow cover in a region through snow depth and snow water equivalent (SWE) estimates. Both active and passive microwave sensors have been used to develop SWE retrieval algorithms. While passive microwave remote sensing can offer near-real time daily data, active microwave sensors onboard satellites and aircraft often provide researchers with observations at a higher spatial resolution. Traditionally, active microwave remote sensors used in SWE retrieval are focused on a specific band like Ku or X band (Yueh et al., 2009; Oveisgharan et al., 2020). In comparison, methods for snow depth retrieval for snow on sea ice using data collected from airborne Frequency Modulated Continuous Wave (FMCW) radars have been experimented with over the years (Kwok et al., 2017). The shift from SWE to snow depth as the retrieval target for FMCW radars is mainly because these methods often rely on obtaining the air-snow and snow-ice interface while many traditional Synthetic Aperture Radar (SAR) based approaches are relying on the relationship between the backscatter and SWE. SWE and snow depth are convertible when the bulk density is provided, as demonstrated by Sturm et al. (2010).

$$SWE = h_s \frac{\rho_b}{\rho} \quad (1.1)$$

where h_s is snow depth; ρ_b is local bulk density; and ρ is the density of water.

A large effort in the development of these airborne FMCW radar related methods stems from a series of campaigns called Operation IceBridge (OIB) which ran from 2009-2020 (NASA., n.d.). The campaigns were aimed at providing observations of polar ice.

As a result, most of the methods for snow depth retrieval were developed and tested during OIB for snow on sea ice or glacial fern. There are only a very few studies that applied or adapted these methods for snow on land (yan et al., 2016), and this is the main motivation of this thesis. The field campaign Icebird 2019 Winter provided an opportunity to conduct experiments on the utility of using wide band radar data collected over the Trail Valley Creek research field site for snow depth estimation. In addition, this data also allows for developing a new snow depth retrieval method based on principles behind sea ice classification as radar altimeter (a similar FMCW radar system to the wide band radar) waveforms were characterized based on a series of parameters and then clustered into representing different objects like first-year ice and multi-year ice.

1.2 Aims and Objectives

The overall aim of this thesis is to test the performance of interface-based snow depth retrieval algorithms developed for snow on sea ice for snow on land, and to explore the potential of adapting the approach for improved snow depth retrievals. Therefore, three specific objectives are listed below:

1. Explain and summarize the basic principles behind using wide band radar data for snow depth retrievals as well as various existing algorithms.
2. Evaluate a recent snow depth retrieval algorithm used for snow on sea ice for snow on land.

3. Develop a terrestrial snow depth retrieval algorithm that is based on existing snow depth on sea ice retrieval methods.

1.3 Thesis Structure

This thesis follows the manuscript style where chapter three is a standalone paper. Chapter two provides the necessary background information about snowpack physical properties, basics of a wide band radar operating in FMCW mode and its physical interaction with a snowpack, a description of existing wide band radar snow depth retrieval algorithms, and the associated ways a waveform can be characterized. Chapter three is the core manuscript paper entitled "Snow depth retrieval from wide band radar in Trail Valley Creek". Chapter four summarizes the limitations and challenges facing the use of the interface-based snow depth retrieval on land, and the broader implications of the study along with suggestions on how some of the limitations can be addressed, and offers a conclusion of the content of the thesis.

Chapter 2

Background

This chapter covers necessary background information for the interface-based pulse peakiness snow depth retrieval algorithm used in this thesis including snowpack physical properties, characteristics of wide band radar systems and their interaction with snowpacks, existing snow depth retrieval algorithms developed for snow on sea ice and associated ways to characterize radar waveforms.

2.1 Snowpack Physical Properties

Two physical properties of snow will be discussed due to their significance in active radar observation of snow. First, a snowpack's response to incoming microwave radiation at a given frequency is largely affected by its dielectric properties, with the relative permittivity used to quantify such properties. Matzler et al. (1984) demonstrated that because snow is a mixture of air, ice and in case of wet snow, liquid water, the effective relative

permittivity of snow is a weighted average of all three components. According to Ulaby et al. (1977), liquid water content fraction which controls snow wetness, has the largest impact on the dielectric properties of the ice-air-water mixture. In addition, other physical parameters including snow density, temperature as well as microstructure properties such as snow crystal morphology and grain size can also affect the snow's dielectric constant.

Second, the backscatter of a snowpack is an essential measurement of imaging radar system such as SAR and non-imaging radar systems such as narrow or wide band radar altimeters. The backscatter from a snowpack is frequency dependent and when present is generated by two scattering types, surface scattering and volume scattering. Surface scattering happens at the interfaces between significant layers with different dielectric properties, such as the air-snow interface, snow-ground interface, and interfaces between snow layers within the pack. Volume scattering by comparison occurs within the snowpack, and as explained by Ulaby et al. (1984), is due to mie or rayleigh scattering processes acting on snowpack ice crystals that exist within an air background. Rott et al. (2010) provided a simplified equation to summarize all the backscatter components from a snowpack as is shown below:

$$\sigma_t = \sigma_{as} + \sigma_v + \sigma_{gv} + \sigma_{g'} \quad (2.1)$$

where the four components from left to right are scattering at air-snow interface (σ_{as}), direct snow volume scattering (σ_v), scattering from snow volume after the signal interacted with the snow-ground interface (σ_{gv}), and scattering from the snow-ground interface ($\sigma_{g'}$), as shown in Figure 2.1.

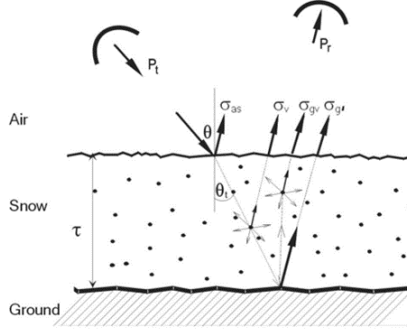


Figure 2.1: Main components of snowpack backscatter (Rott et al., 2010)

The penetration depth for a radar signal not only depends on the dielectric properties of the snowpack but also on the frequency of the radar signal itself. Therefore, different frequencies lead to different scattering interactions. More specifically, if the radar signal frequency is low enough it will completely penetrate the snowpack. For example, at C-band (5.85–6.43GHz), the scattering from the snowpack is generally considered negligible. However, the presence of water in the snowpack will change the relative permittivity of the snowpack so that C-band radar can be used for wet snow mapping and in some cases for wet snow depth mapping. As demonstrated by Lievens et al. (2019), an empirical relationship between wet snow depth and co-pol ratio of a C-band radar can be developed for deep mountain snow. By comparison, at higher frequencies, such as Ku-band (12–18 GHz), the incident radar signal interacts with the snow through either or both volume and surface scattering mechanisms within the snow. Snowpack structural properties, such as liquid water content, density and grain size, control the type of scattering interaction and influence the penetration depth of a radar signal at a given frequency.

2.2 Wide Band Radar and Its Interaction with Snow

A wide band radar in this thesis is specifically defined as an airborne FMCW radar capable of operating in a wide bandwidth similar to the ones deployed in OIB and Icebird 2019 Winter. However, this is a narrow definition and as demonstrated by Marshall and Koh (2008), various FMCW radar systems with different configurations and operating bandwidths have been used for snow research since 1977. Successful SWE estimates as well as snow stratigraphy measurements have been obtained using data from mostly ground-based FMCW radar systems, and it was shown that methods including developing empirical relationships between the two-way travel time of the radar signal and SWE as well as utilizing frequency spectrum were viable (Marshall et al., 2004; Marshall et al., 2005; Marshall et al., 2007).

A wide band radar differs from a normal Continuous Wave (CW) radar in that the emitted frequency of a wide band radar is modulated meaning that it changes overtime. More specifically, while a C-band SAR such as Sentinel-1 emits radar signals with a central frequency of 5.405 GHz (ESA, n.d.) and a narrow bandwidth (usually in scale of MHz), a wide band radar such as the one used for the Icebird 2019 experiment, emits radar signals ranging from 2 GHz to 18 GHz (Hendricks et al., 2019). There are several ways the frequency can be modulated including sawtooth, triangular, square-wave and stepped modulation, which all describe how the frequency changes over time. For example, a sawtooth modulation means that the frequency increases linearly overtime and returns to its original value at the end of a cycle, as is shown in Figure 2.2.

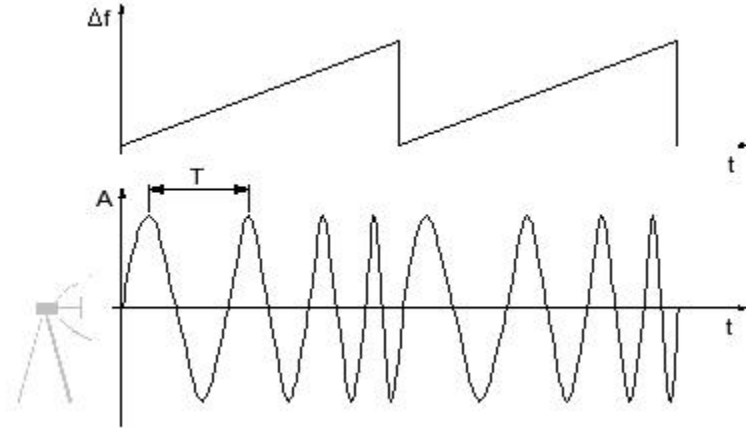


Figure 2.2: Sawtooth modulation (EuroMotor virtual college., n.d.)

The core characteristics of a wide band radar is the modulated frequency in which the constant change in frequency provides the system with a time delay to determine accurately when the signal was emitted and when its echoes are received as shown in Figure 2.3 (Griffiths, 1990; Mostafa, 2019; Wolff, n.d.). In contrast, a normal CW radar system cannot determine the time when a signal is transmitted, and its echo received to calculate target ranges due to the lack of time markers.

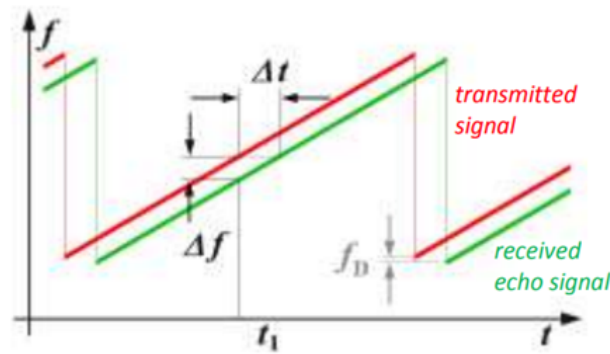


Figure 2.3: Time delay between the transmitted and received signal (Wolff, n.d.)

A simplified wide band radar system is shown in Figure 2.4, where a signal with modulated frequency, called a Chirp is transmitted at the chirp/FMCW generator. The received echo is enhanced and correlated with the original chirp to form the beat signal (Mostafa, 2019; Brooker, 2005). To convert the beat signal to detect ranges, a range Fast Fourier Transformation (FFT) is used where each range FFT bin equals a specific distance. After the transformation, the peaks in the waveform often represent the presence of targets at different distances.

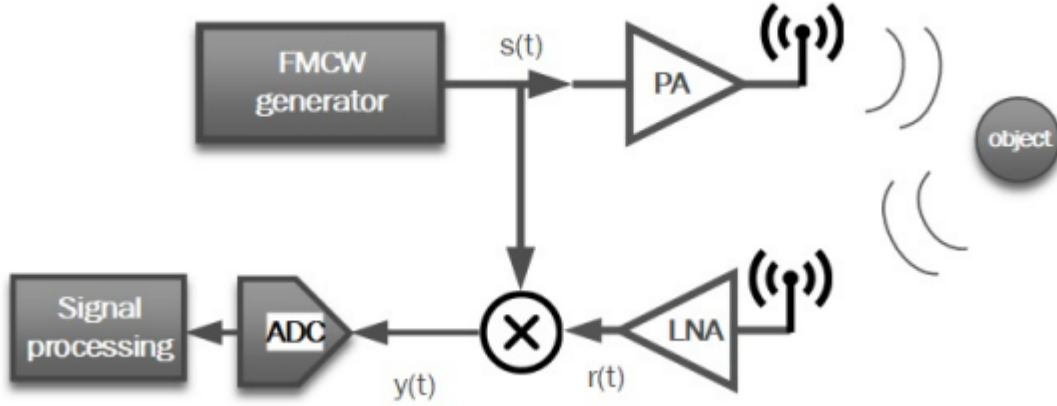


Figure 2.4: A simplified wide band radar system (Mostafa, 2019)

Stelzer et al. (2008) and Panzer et al. (2013) provided the following equation for calculating distance for objects in a vacuum or open air from a wide band radar system.

$$R = \frac{c_0 |\Delta t|}{2} = \frac{c_0 |\Delta f|}{2(d(f)/d(t))} \quad (2.2)$$

where c_0 is the speed of light; Δt is the time delay; Δf is the measured frequency difference; $d(f)/d(t)$ is the frequency shift per unit of time. And the range resolution of a specific

wide band radar is calculated using the following equation:

$$R_{min} = \frac{c_0}{2B} \quad (2.3)$$

where B is the bandwidth of the chirp.

There are a number of advantages for the wide band radar. The wide band radar provides a highly accurate measurement of range. It is also capable of measuring target ranges and their relative speed at the same time (Stove, 1992; Wolff, n.d.; Griffiths, 1990). In terms of the interaction between a wide band radar and a snowpack, similar to using a FMCW radar to identify objects at different distances, different interfaces, for example, air-snow interface, are often represented by peaks in the converted beat signal due to the drastic changes in dielectric properties. Figure 2.5 demonstrates an interaction where only two interfaces would be recorded in the beat signal (Farrell et al., 2012). Compared with a regular radar system with a fixed frequency where the backscatter components can often be determined by the radar parameters and snowpack parameters, due to the modulated frequency of a wide band radar, the backscatter from the snowpack would not only be from the two interfaces but also from within the snowpack, making the ideal two-peak waveform unlikely. In addition to waveforms containing other information than the two interfaces, there also exists the possible confusion between an air-snow interface return from the next cell and the snow-ice return from the current cell, as is shown in Figure 2.6 (Galin et al., 2012).

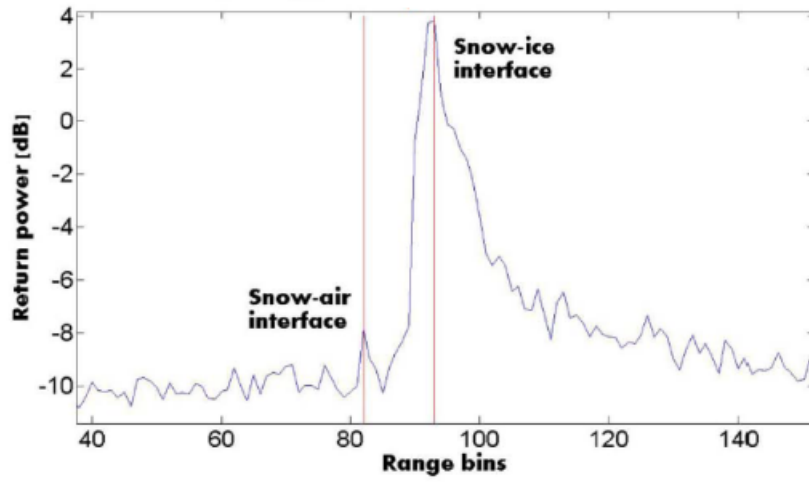


Figure 2.5: 2-peak waveform demonstrating the interface impulse response (Farrell et al., 2012)

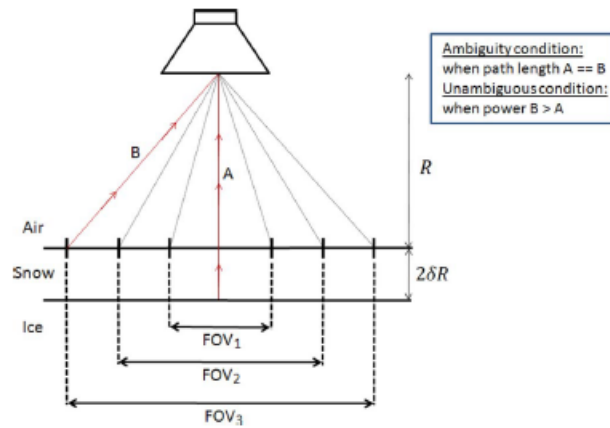


Figure 2.6: Ambiguity between air-snow and snow-ice interface (Galin et al., 2012)

2.3 Characterizing Wide Band Radar Waveforms

Multiple papers have described approaches to run classification algorithms on the radar waveform based on the fact that waveforms from a radar altimeter for different types of sea ice will have noticeable different shapes as is shown in Figure 2.7 (Müller et al., 2017; Zygmuntowska et al., 2013; Shen et al., 2017; Ricker et al., 2015; Aldenhoff et al., 2019). It is necessary to note that radar altimeters are similar to the wide band radar system but they normally have a relatively coarser range resolution. In addition to distinguishing first-year ice from multi-year ice, similar concept of comparing the characteristics of waveforms for different objects was also used for detecting leads in polar oceans (Dettmering et al., 2018). There are differences in details and how the classifications were conducted, but the initial step for these applications is similar which is characterizing the radar waveforms using several parameters.

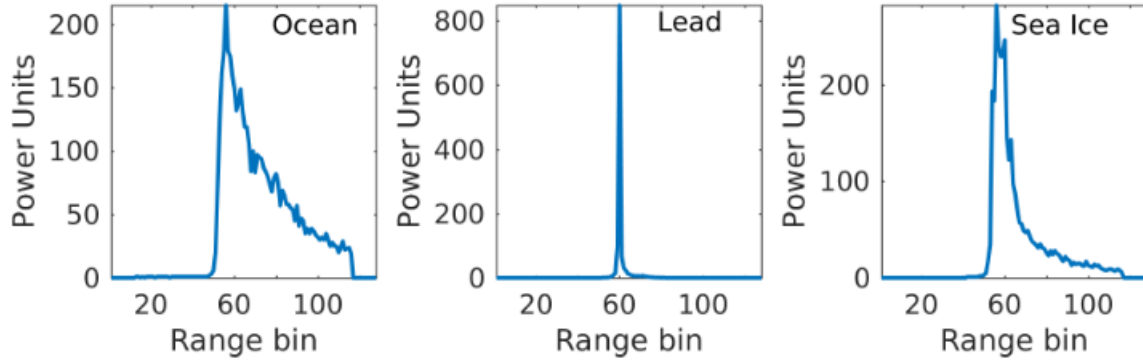


Figure 2.7: SARAL radar altimeter waveforms for ocean, lead and sea ice (Müller et al., 2017)

Common parameters used across different papers include the maximum return

power, stack standard deviation (SSD) which illustrates the variation in surface backscatter in relation to incident angles (Ricker et al., 2015, Wingham et al., 2006), pulse peakiness of the maximum return peak, leading edge width (LeW) and trailing edge width (TeW), as is shown in Figure 2.8. Zygmuntowska et al. (2013) defined LeW as the distance between 1% and 99 % of the maximum power return after fitting a Gaussian curve to the leading edge while TeW was defined as the distance between 99% and 1% of the maximum power return after fitting a Gaussian curve to the trailing edge. Additional characteristics are added based on specific research areas and goals. For lead detection, sea-ice concentration was added to distinguish between sea ice and open water (Dettmering et al., 2018). Ricker et al. (2015) utilized a second peakiness measurement parameters, stack kurtosis (K) to determine how susceptible the distribution is to an outlier.

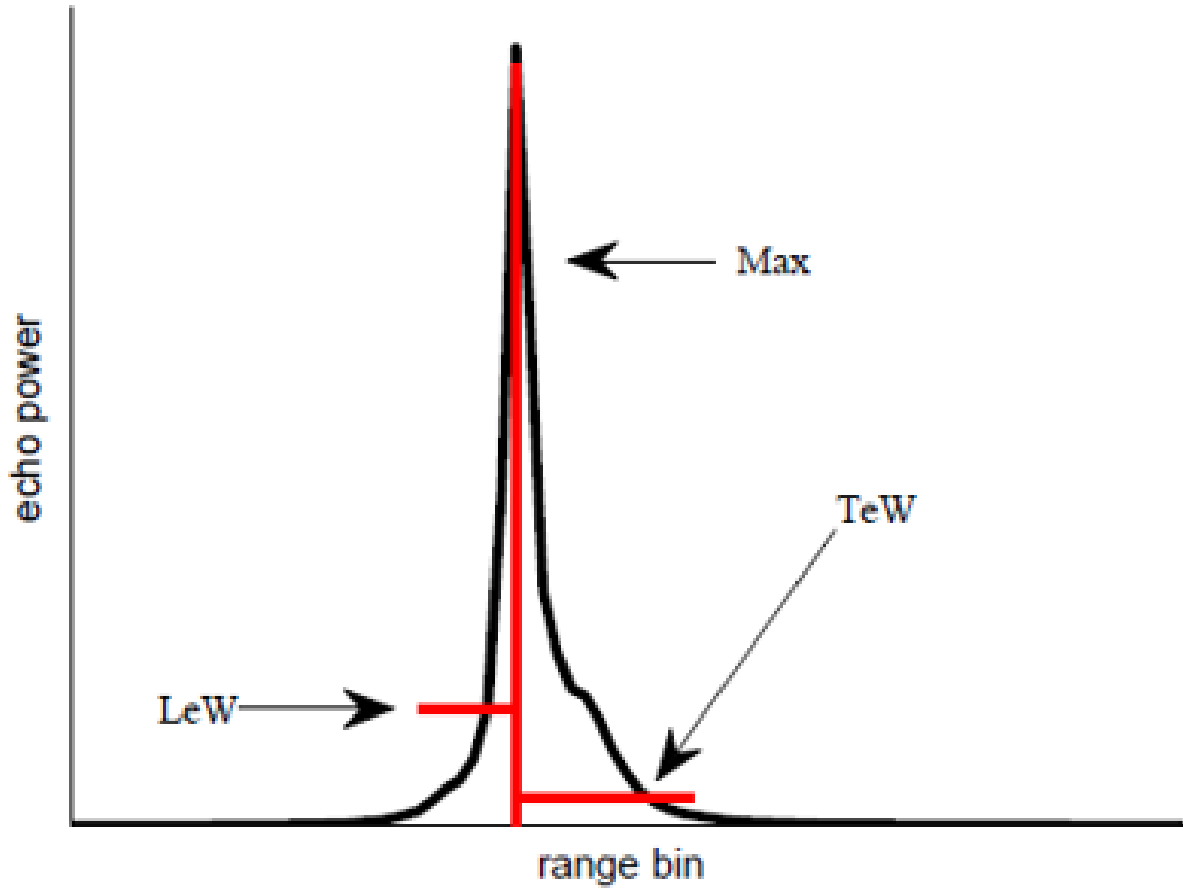


Figure 2.8: Definitions of LeW and TeW (Zygmuntowska et al., 2013)

The reason why characterizing waveform is important is that it offers a way to quantify the general shape of the waveform and represent a certain shape in a set of parameters that can be fed into various classification algorithms. The first approach was to classify the waveforms into as many clusters as was needed to achieve a desirable separation and then manually condense the clusters based on cluster waveform characteristics and prior knowledge into different surfaces depending on the study (Müller et al., 2017; Dettmer-

ing et al., 2018). In comparison, a second approach was taken by Shen et al. (2017) to feed the characterizing parameters of waveforms instead of the waveforms themselves into classifiers. A wide range of combinations of parameters were tested on numerous different classification algorithms. In a similar perspective of utilizing characterizing parameters in analyses is to set thresholds for those parameters to classify waveforms into specific surface types, as was explained by Ricker et al. (2015).

Similar to these applications of characterizing and classifying radar waveforms on sea ice, snowpacks with different snow depth would have wide band radar waveforms with different shapes, and therefore, it is possible to apply the same technique to snow depth retrieval. However, in terms of radar waveform shapes, the difference between a 20 cm snowpack and a 30 cm one is considerably less obvious than the difference between first-year ice and multi-year ice. Instead of focusing on the overall shape of the waveforms, peaks in the wide band radar waveforms that potentially represent interfaces between mediums with different dielectric properties are utilized for snow depth retrieval on sea ice. Despite the different perspective, some waveform characterizing parameters described above have been successfully scaled down to be applied to characterizing individual peaks, most notably the pulse peakiness concept.

2.4 Existing Algorithms for Wide Band Radar Snow Depth Retrieval

As is shown in Figure 2.5, all the retrieval methods described below are focused on obtaining two interfaces, the air-snow interface and snow-ice interface. Once the two interfaces are determined, snow depth was calculated by multiplying the time difference between the two by the speed of light in the snowpack based on an assumption of snow density.

2.4.1 Kurtz et al. (2013) algorithm

Kurtz and Farrell (2011) and Kurtz et al. (2013) described the snow depth retrieval algorithm used in NASA’s OIB. As it was an ongoing campaign, data was collected using slightly different radar systems. For Kurtz and Farrell (2011), the radar system used had a operating frequency range of 2.5-7.0 GHz and for Kurtz et al. (2013), the system had a frequency range of 2.0-6.5 GHz.

It was found that the strongest peak of the waveform corresponded to the snow-ice interface while the air-snow interface often was not represented by a specific peak (Kurtz et al., 2013; Farrell et al., 2012). The reason that air-snow interface was harder to match to a specific peak, as explained by Kurtz et al. (2013) was the combination of snow density variations, surface roughness as well as volume scattering, all of which resulted in a relatively more sophisticated waveform than the ideal two clear peak pattern.

As for the algorithm itself, there were slight differences between the ones used in

Kurtz and Farrell (2011) and Kurtz et al. (2013) due to the latter having an improved radar system. Therefore, the retrieval from the latter will be the focus here. First, a starting point of the search for the air-snow interface was defined as the first point whose return power in dB exceeded a set empirical threshold in addition to having the mean return power of its subsequent 6 range bins being above the same threshold (Kurtz et al., 2013). Second, after this starting point was found, the air-snow interface was defined as either the first point of continuous power increase or a local maximum that was greater than a set empirical threshold. Finally, snow depth was obtained using the aforementioned method of multiplying the time between two interfaces by the speed of light in the snowpack which was determined using snow density provided in previous studies.

During the retrieval, selected data were discarded due to various reasons including negligible snow cover on open water and newly frozen leads, insufficient signal power, and from physical confounding surface states such as the presence of liquid water (Kurtz et al., 2013).

2.4.2 Kurtz et al. (2014) algorithm

The principle behind GSFC-NK was similar to the NSIDC method where the key to snow depth retrieval was obtaining air-snow and snow-ice interface. Kwok et al., (2017) explained that the GSFC-NK algorithm was a waveform fitting method that matched a model waveform to the wide band snow radar data and the two interfaces were from the fitting results. A more detailed description could be found in Kurtz et al. (2014). A physical model for the radar altimeter onboard CryoSat-2 was established to simulate

waveforms over sea ice, after which a lookup table for a discrete set of cases for Arctic sea ice was created to enable fitting of individual waveforms. Discrete cases were later linearly interpolated using a fitting routine with calculated interpolation coefficients. According to Kurtz et al. (2014), to reduce the differences between the model fit and CryoSat-2 waveforms, a least-squares fitting routine utilizing a bounded trust region method was employed.

Kwok et al., (2017) noted that the fitting algorithm was extremely sensitive to the selected fitting parameters, and for data from operation Icebridge, some of the parameters were chosen based on existing NSIDC empirical algorithms.

2.4.3 Koenig et al. (2016) algorithm

The Snow Radar Layer Detection (SRLD) algorithm was developed for detecting deeper layers below the snow-ice interface for land ice (Kwok et al., 2017 & Koenig et al., 2016). In the original algorithm, to determine air-snow interface, a median filter was applied vertically in the echogram to each radar trace to reduce noise and then the interface was obtained by setting a threshold of 4 times the background noise level (Koenig et al., 2016). After the snow surface was defined, the waveforms were aligned to the surface and layers below the surface were labeled by using a high-pass filter to detect different peaks.

To adapt the algorithm in Koenig et al. (2016) for snow depth retrieval on sea ice, a gradient between open air snow radar returns and the maximum return was established. Open air radar return was defined to be a sample median of the values above the surface (Kwok et al., 2017). Air-snow interface was obtained by locating where the radar return had

exceeded background noise level, i.e., open air return, near the beginning of the gradient. And snow-ice interface was set to be the place near the end of the gradient where a maximum peak was produced due to the dielectric difference in ice and snow. According to Kwok et al. (2017), the gradient threshold values vary from frame to frame.

2.4.4 Newman et al. (2014) algorithm

The Wavelet algorithm was introduced in Newman et al. (2014) and was made up of three components. First, two interfaces were estimated using the Haar wavelet-continuous wavelet transform (Haar-CWT). According to Newman et al. (2014) the advantage of using the Haar-CWT method is that it allows the developed snow depth retrieval to be independent from changes in transmitted power and receiver noise, both of which vary during different campaigns. Therefore, by using the Haar-CWT, the approach does not require fixed thresholds like other methods, such as the NSIDC method in Kurtz and Farrell (2011) and Kurtz et al. (2013).

More specifically, the Haar-CWT was developed and optimized to detect sudden transitions within the waveform (Kwok et al., 2017), and because of the dielectric contrasts between air, snow and ice, the sudden transitions in a waveform could be determined to be interface returns. To detect air-snow interface, the waveform was transformed into logarithmic scale while for snow-ice interface, the waveform was fed into the algorithm unchanged. CWT coefficients at each scale which started at scale 3 and ended at a manually set maximum scale were calculated using the two different inputs, normal for snow-ice interface and logarithmic for air-snow interface. A sum of all the CWT coefficients from odd scales

were obtained and the interfaces were defined as the range bin where the aforementioned sum reached its maximum (Newman et al., 2014).

Second, a topographic filter was applied to the retrieved snow depth where interfaces chosen in locations with heavily deformed ice topography were discarded. More specifically, a parameter h_{topo} which was calculated by subtracting the elevation of the 5th percentile of Airborne Topographic Mapper(ATM) measurements from the elevation of the 95th percentile of ATM measurements. A flat surface would have a h_{topo} of 0 and any interface chosen in places with a h_{topo} larger than 50cm was neglected (Newman et al., 2014; Kwok et al., 2017). The third and final component of the Wavelet algorithm was assigning snow depth precision to each retrieved snow depth independently (Kwok et al., 2017).

2.4.5 Kwok and Maksym (2014) algorithm

Kwok et al. (2017) explained that the JPL algorithm was a simplified version of the one purposed by Kwok and Maksym (2014). Similar to some of the other approaches, a baseline assumption was made that the maximum power return peak was the snow-ice interface. To determine whether a peak was significant enough to be considered for being a potential interface peak, instead of using a pre-determined intensity return threshold, Kwok and Maksym (2014) used how far a peak deviated from neighboring returns. More specifically, 2 conditions were to be met; first, a peak return needed to be above a signal level defined by neighbouring returns and the system noise level. And second, a peak return needed to be above a signal level defined by returns one range bin away from the peak. The farthest peak identified this way was designated as the air-snow interface. In

addition, Kwok and Maksym (2014) also provided improved criteria which utilized peak signal to noise ratio (PSNR) for detecting air-snow interface to decrease the effect of radar sidelobes. The algorithm used several parameters determined by the radar system, and therefore could adapt well to the improvement made during Operation Icebridge missions (Kwok et al., 2017).

2.4.6 Julita et al. (2021) algorithm

Julita et al. (2021) provided a different interface-choosing method based on the existing concept of pulse peakiness. As this will be further discussed in detail, only a brief introduction will be provided here. Instead of calculating the pulse peakiness of the entire waveform, which is based around the maximum return peak, this approach in turn calculated pulse peakiness for each individual peak that had the potential of being air-snow interface or snow-ice interface, after which, the pulse peakiness of these peaks was used in conjunction with manually set thresholds to determine the two interfaces.

Chapter 3

Snow Depth Retrieval from Wide Band Radar in Trail Valley Creek

3.1 Introduction

Seasonal snow is an essential component of the hydrologic cycle, where changes in snow cover and distribution can have significant impacts on precipitation, wildlife, and water supply. In addition, variations in snow cover can serve as an indicator for climate change (Croce et al., 2018). A timely accurate estimate of snow depth or snow water equivalent (SWE) can be used for hydrology and climate modeling, managing local water supply as well as presenting a quantifiable parameter that responds to climate change.

It is because of the necessity of snow monitoring that the methods for snow depth / SWE estimates have been developed throughout the years. While field measurements or

records from weather stations are known to be reliable, they are localised in representativeness and so to cover the vast areas around the globe, remote sensing technology including passive and active microwave sensors have been utilized along with numerical models.

More specifically for snow depth or SWE retrieval using active sensors, there have been experimental observation methods developed utilizing radar systems at Ku or X band (Yueh et al., 2009; Oveisgharan et al., 2020; Leinss et al., 2015). As explained by Leinss et al. (2015) and Oveisgharan et al. (2020), it is possible to retrieve SWE with the interferometric phase differences between two measurements at specific frequencies, especially at L-band. Cui et al. (2016) proposed using Synthetic Aperture Radar (SAR) observations at Ku and X band coupled with a backscatter forward model to estimate the optical thickness of the snowpack. By contrast, Operation Icebridge (OIB) deployed an experimental wide band radar with a frequency range of 2-8 GHz which differs from the radar systems used by Leinss et al. (2015) and Cui et al. (2016) in that the frequency of the transmitted signal is constantly modulated. In addition, various other FMCW radar systems have also been used for SWE retrieval and snow stratigraphy measurements (Marshall and Koh., 2008). However, as those mostly ground-based FMCW radar systems are not the focus of this paper, a wide band radar here is specifically defined as an airborne FMCW radar capable of operating in a wide bandwidth similar to the ones used in OIB and AWI Icebird campaigns.

Currently, most snow depth retrieval algorithms using data collected by a wide band radar are based on obtaining two crucial interfaces, air-snow interface, and snow-ice / snow-ground interface (Kwok et al., 2017). As OIB which provided a large percentage of the data used for developing various retrieval algorithms is focused on offering detailed information

about polar ice, most of these retrieval methods are for snow on sea ice. According to Kwok et al. (2017), the difference amongst these algorithms mainly lies in how the two interfaces are obtained and how they are optimized for a specific radar system. Kurtz and Farrell (2011) and Kurtz et al. (2013) proposed an interface identifying method that combined searching for local maximum and empirical thresholding, which was adopted by the National Snow and Ice Data Center (NSIDC). Later on, a waveform fitting method was developed by Kurtz et al. (2014) where a physical model was established to simulate wide band radar waveforms over sea ice and the two interfaces were obtained by fitting data to the model. In addition, a retrieval based on the gradient of radar intensity return was adapted from Koenig et al. (2016) and tested in Kwok et al. (2017). In comparison, Newman et al. (2014) took a slightly different approach that avoided the need for fixed sets of thresholds used by several previously mentioned methods. As demonstrated by Newman et al. (2014) and Kwok et al. (2017), a Haar-CWT was utilized to detect abrupt changes within the waveform and those identified abrupt transitions were used to determine the two interfaces under specific criteria. To cope with the gradual improvement of the radar system used in different campaigns in OIB, a simplified version of the algorithm from Kwok and Maksym (2014) was developed where interfaces were located by measuring their deviation from their neighbors. Kwok et al. (2017) concluded that all methods described above were capable of reproducing the high snow depth pattern expected in the study area, but the performance of each individual method when compared with field measurement differs. More specifically, according to Kwok et al. (2017), when comparing the retrieval snow depth from these five methods with field data from Eureka, their corresponding mean bias are -0.8 cm (kurtz et al., 2013), -5.7 cm (kurtz et al., 2014), 1.3 cm (Koenig et al.,

2016), 2.0 cm (newman et al., 2004) and -2.2 cm (kwok and Maksym, 2014)

In addition to its application in snow depth retrieval, FMCW radar has also been used for sea ice classification in the form of radar altimeter onboard satellites. For sea ice classification, instead of looking for interfaces, the shape of the waveform is of more importance (Müller et al., 2017; Zygmuntowska et al., 2013; Shen et al., 2017; Ricker et al., 2015; Aldenhoff et al., 2019). Different types of sea ice as well as open water present waveforms of various shapes and therefore can be classified by running numerous classification algorithms. While the difference of waveform shapes amongst snowpacks with different thickness may not be as prominent, it can be argued that it may be possible to apply the same principle to classify waveforms of snow on sea ice / land into clusters of different snow depths.

Since the start of OIB, most algorithms developed including the ones described above are for snow on sea ice because of several possible reasons. First, the difference between dielectric properties of snow and frozen soil is generally weaker compared with snow and sea ice; second, the presence of vegetation presents a sizable challenge. When there is vegetation buried in the snowpack, it would likely increase the complexity of the snowpack while if the vegetation is emerging from the snowpack, it could interfere with radar waveforms before they reach the snowpack. Thirdly, terrain features on land like steep slopes and hummocks could further increase the complexity of the waveforms. In addition, the frequency range of the radar systems used by these algorithms are often not wide enough to cover shallow snow. However, Icebird 2019 Winter collected data over the Trail Valley Creek research station using a wide band radar with a frequency range of 2-18 GHz and this provides an opportunity to examine waveforms for snow over various

land terrain features including different types of vegetation as well as hummocks and to determine whether principles behind snow on sea ice algorithms can be used for snow on land.

As a result, the goal of this study is to test the applicability of the interface-based snow depth retrieval methods developed for snow on sea ice to snow on land, and to explore the possibility of adapting sea ice waveform classification algorithms for snow depth retrievals. To achieve these goals, specific objectives are to:

1. Apply and adapt the pulse peakiness algorithms for snow on sea ice to snow on land.
2. Explore the effect of surface vegetation and topography on the wide band radar waveform retrieval performance.
3. Develop a classification based snow depth retrieval method based on sea ice classification approach.

3.2 Study Area

The study area for this experiment is based at the Trail Valley Creek (TVC) research station, located at the northern boundary of the tundra-boreal forest ecotone, roughly 50 km north of Inuvik, Northwest Territories (Pohl et al., 2005; Pomeroy et al., 1998). The TVC research basin is within the Inuvialuit Settlement Region near the Inuvik-Tuktoyaktuk Highway. The watershed drains approximately 58 square kilometers into the Husky Lakes Estuary network. In addition, the basin itself is underlain with a continuous ice-rich permafrost (Trail Valley Creek Arctic Research Station, n.d.).

Topographically, the area is dominated by gently rolling hills (Pohl et al., 2005). A major terrain feature of the study area is the hummocks, which are defined as sub-meter to meter-scale circular to oval mounds (Verret et al., 2019). The existence of hummocks leads to a relatively large amount of the micro terrain that is below the spatial resolution of most field measurements. More complex terrain is present in the southern margins of the study area, in proximity to forests.

According to Boike and Grünberg (2019) and Palmer et al. (2012), the study area contains 6 main vegetation types including trees (white and black spruce (*Picea glauca* and *Picea mariana*), tall shrubs, riparian shrubs, dwarf shrubs, tussocks, and lichens. Tussocks, lichens and dwarf shrubs are the dominant vegetation types while trees and tall shrubs have a relatively sparse distribution. For tall shrubs particularly, there are often other low height vegetation in between two tall shrubs (Boike and Grünberg, 2019). The landscape also features several lakes as shown in Figure 3.1 with a number of streams and rivers running through the study area.

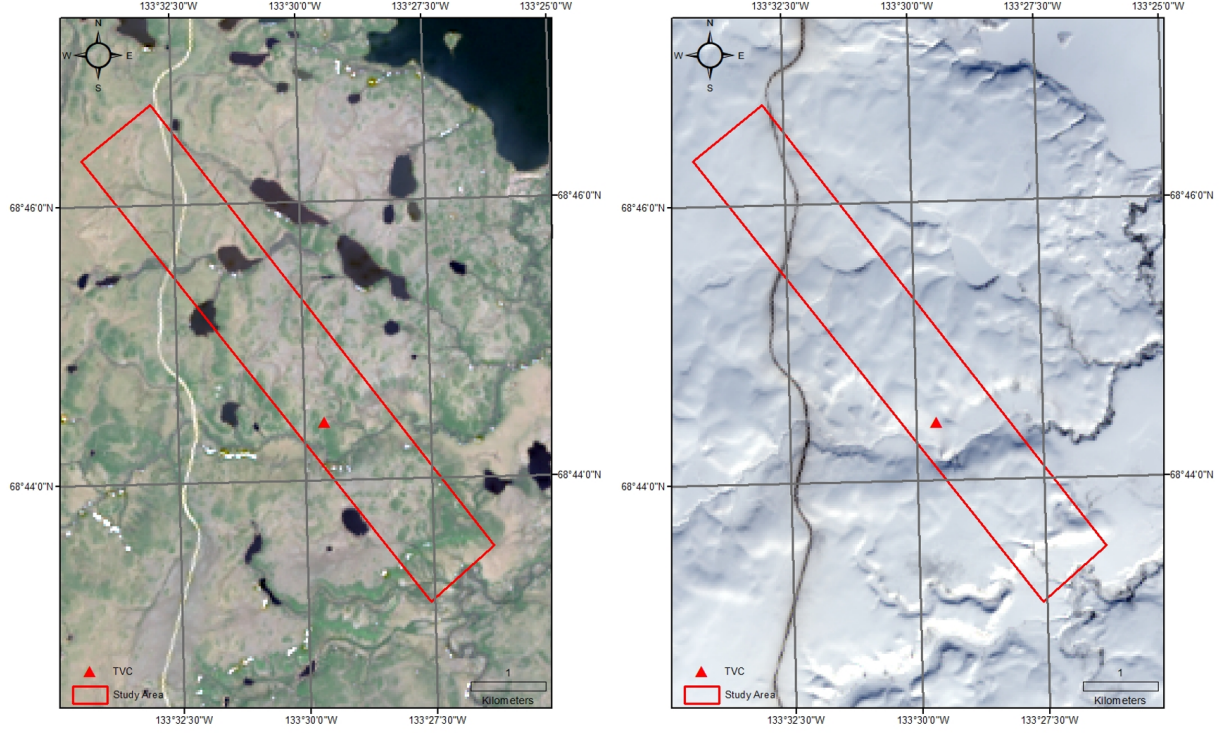


Figure 3.1: Landsat 8 images of study area; left is from 2019.06.18 and right is from 2019.04.15. Red triangle represents the location of TVC research station

3.3 Datasets

3.3.1 Snow Radar Data and Airborne Laser Scanning (ALS) Data from Icebird 2019 Winter

The Icebird 2019 Winter campaign was aimed at conducting sea ice surveys over different ice regimes to observe, analyze and estimate sea ice thickness in addition to

examining regional snow distribution (Hendricks et al., 2019). An airborne wide band snow radar system observed snow on sea ice and an airborne laser scanner (ALS) system observed surface roughness features (Hendricks et al., 2019). Five field observation periods were planned for the 2019 winter campaign, in which the fourth period, conducted between April 6th and April 11th, 2019, was when the snow radar and ALS data were collected and are used in this study. The particular study date that covered the TVC region was April 10th, 2019.

The onboard ALS system was a Riegl VQ-580 (Hendricks et al., 2019). The ALS point cloud data used in this study was level 1b with the included metadata consisting of geolocation and positioning information. With AWI toolbox (Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, 2019), it was possible to filter the large point clouds by time period and narrowed it down to point clouds that were within the TVC study area.

The snow radar system, mounted on the Polar-6 Basler BT-67, aircraft was set to operate in Frequency-Modulated and Continuous Wave (FMCW) mode for low-altitude missions. The aircraft flew at an elevation of approximately 61 m (200 ft) and 488 m (1600 ft) when using the radar for observation (Hendricks et al., 2019). The wide band radar operated at a frequency between 2 and 18 GHz with 2 receiving channels (horizontal and vertical). Four combinations of polarization observations were achieved, VV, VH, HV, HH. The theoretical range resolution within the snowpack, assuming a snow density of 300

kg/m^3 , was 1.14 cm, calculated using the following equation:

$$R_{min} = \frac{k \times c_0 \times 0.5}{B \times \sqrt{(1 + 0.51 \times \rho)^3}} \quad (3.1)$$

where c_0 is the speed of light in a vacuum, k is parameter that re-spaces the waveform array data, B is the radar bandwidth and ρ is the density of the snowpack (Julita et al., 2021).

At 61 m (200 ft) height, the radar had an across-track footprint of 2.1 m and an along-track footprint of 2.0 m while at 488 m (1600 ft) height, the two respective footprints were 6.0 m and 10.8 m. According to Hendricks et al. (2019), initial processing of the raw wide band radar was done using CRESIS Matlab Toolbox, followed by coherent noise removal and deconvolution of echograms (Julita et al., 2021).

After the pre-processing, the data was separated in to 24 continuous segments of the flight path in a Matlab file format. Within the TVC study area, elements from 6 different flight paths were used (path numbers 1, 2, 6, 7, 11, 12, at VV polarization). The reason for using VV polarization instead of HH is that VV waveforms are generally of higher quality, for example, less noise and interference, and are more suited for identifying peaks. Using PySnowRadar (King et al., 2020), the files were read into python and each waveform was presented as a 1x14050 numpy array, with each element of the array being an intensity return. Each waveform data record included the impulse data, longitude, latitude, and the elevation of the aircraft. Figure 3.2 shows the flight line locations. Waveforms over water bodies (lakes and rivers or streams) were removed from the analysis leaving a total of 3128 radar data locations.

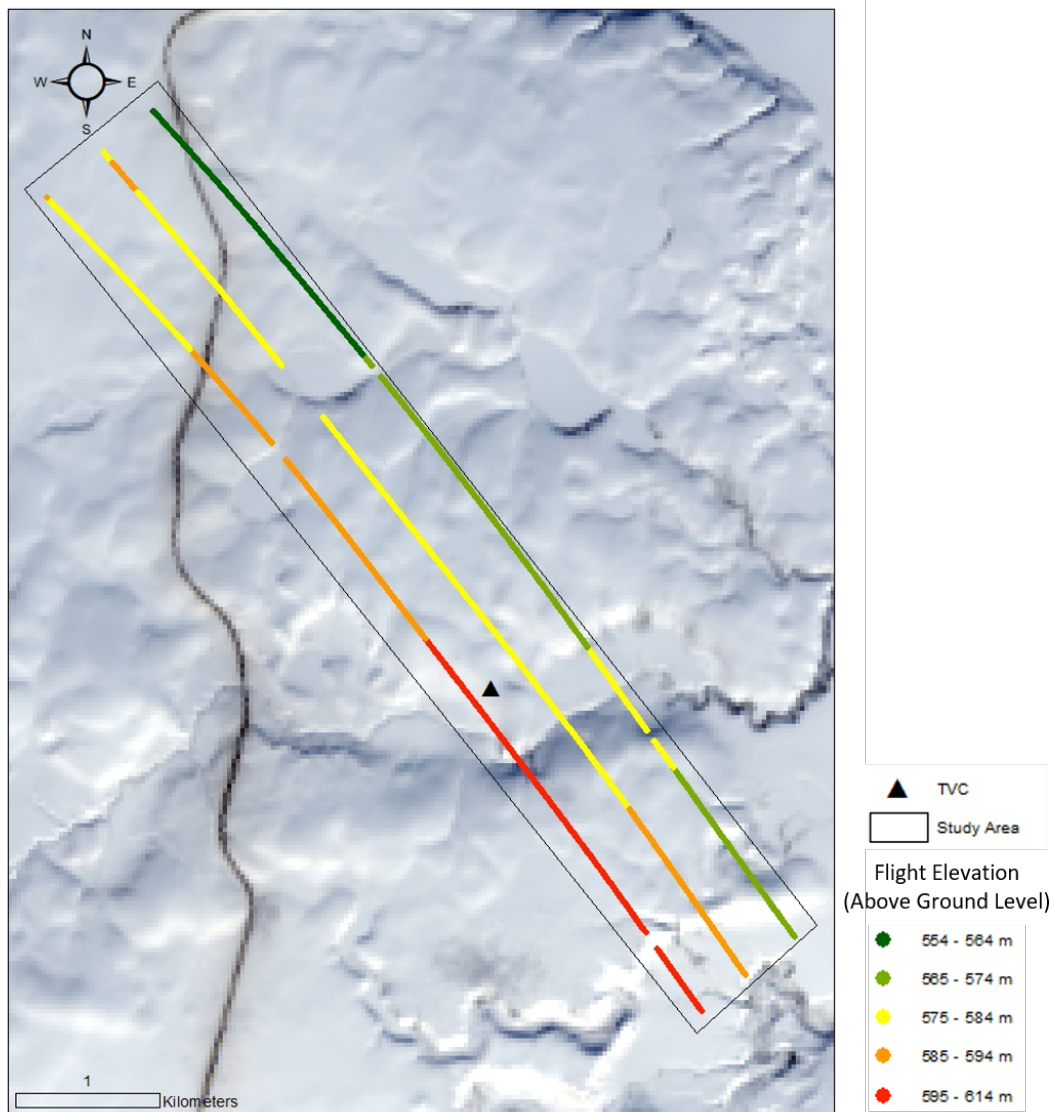


Figure 3.2: Location of the used radar data points colored by flight elevation (Landsat 8 from 2019.04.15 as background)

3.3.2 ALS Point Clouds From 2016

A snow-free ALS data set for the study region was acquired on September 13th, 2016, using a Riegl LMS-Q680i. The reasons for this flight was to provide a snow-off DEM for the development of a lidar snow depth product when combined with a snow-on DEM from 2019. Details of data acquisition can be found in Anders et al. (2016). Three products from Anders et al. (2016) were used in this study including a digital terrain model (DTM), maximum vegetation height map and mean vegetation height map, as is shown in Figure 3.3. All three data have a spatial resolution of 1 m and were later clipped to the study area. The 2016 DTM was used as a snow-off reference for the 2019 snow-on DTM to generate a snow depth map. Both vegetation height maps were used in conjunction with the TVC vegetation map to provide a guideline for typical heights of different vegetation types.

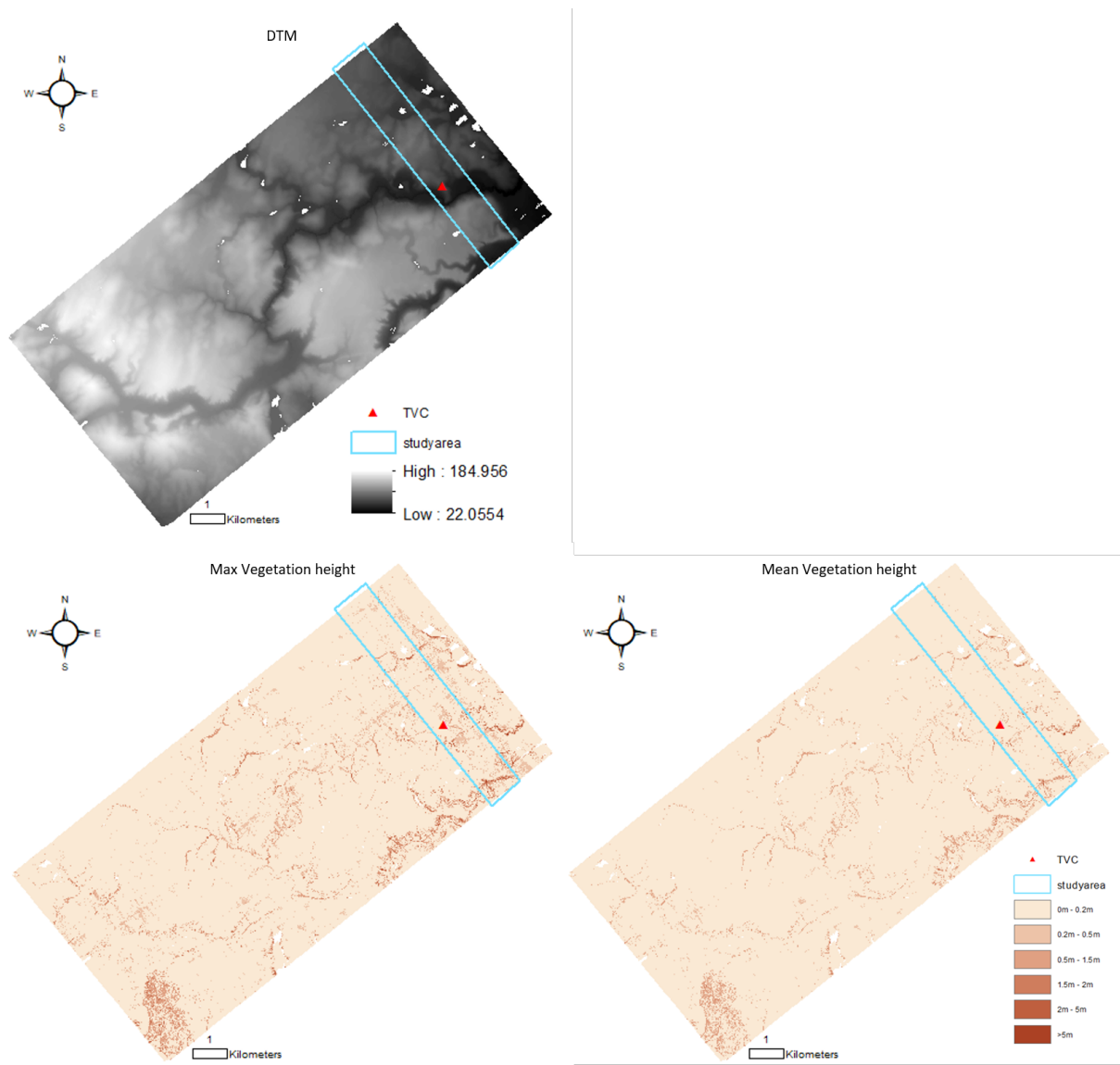


Figure 3.3: 2016 snow off DTM, max vegetation height and mean vegetation height maps

3.3.3 Field Snow Depth Probe Measurements

Field snow depth data were collected during three field campaigns in the winter of 2018-2019. Snow measurements focused on the open tundra area around the TVC research Station. Snow depth measurements were made using a GPS-enable depth measuring probe, called a Magnaprobe (SnowHydro, n.d.). Table 3.1 shows the dates of the three campaigns as well as the number of depths recorded. In this paper, only the data from the last campaign were used because this campaign was temporally closest to when the radar flights took place on April 10th. Due to the limited coverage of the 2019 ALS point clouds, only approximately half of the measurements from the third Campaign were used (4334 points), whose locations and snow depth distribution are shown in Figure 3.4 and Figure 3.5. Most points have a snow depth between 0.3 and 0.5 m. The mean snow depth for all the used measurements is 0.42 m with a standard deviation of 0.13 m. As the field measurements do not cover gulleys where deep snow is most likely to accumulate, there are few points with snow depth higher than 0.9 m.

Table 3.1: Details of three field campaigns

	Dates	Number of Depths Recorded
First Campaign	November 12th -18th 2018	6185
Second Campaign	January 11th-20th 2019	6740
Third Campaign	March 19th-27th 2019	8541

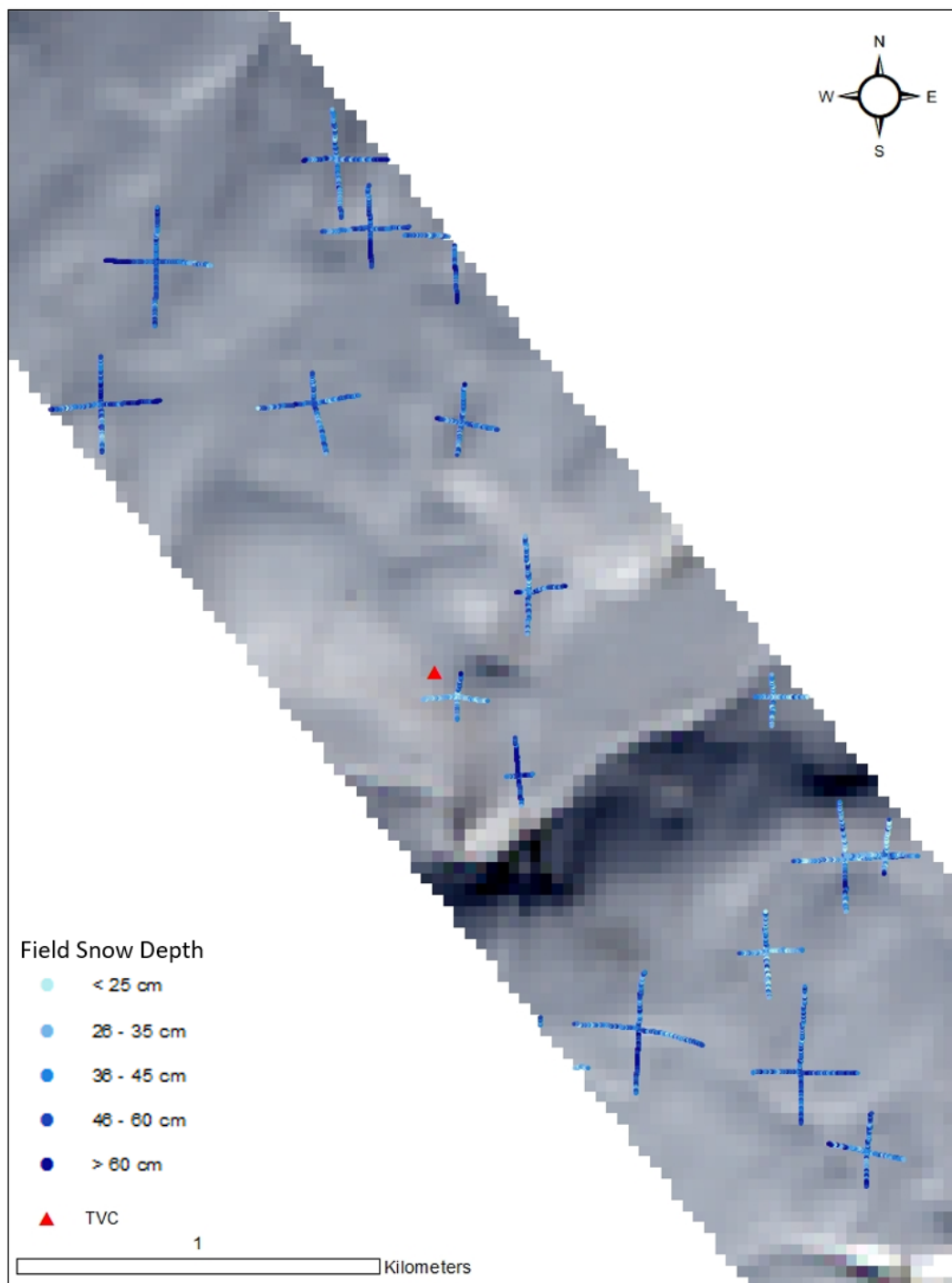


Figure 3.4: Locations of used field measurement (Landsat 8 from 2019.04.15 as background)

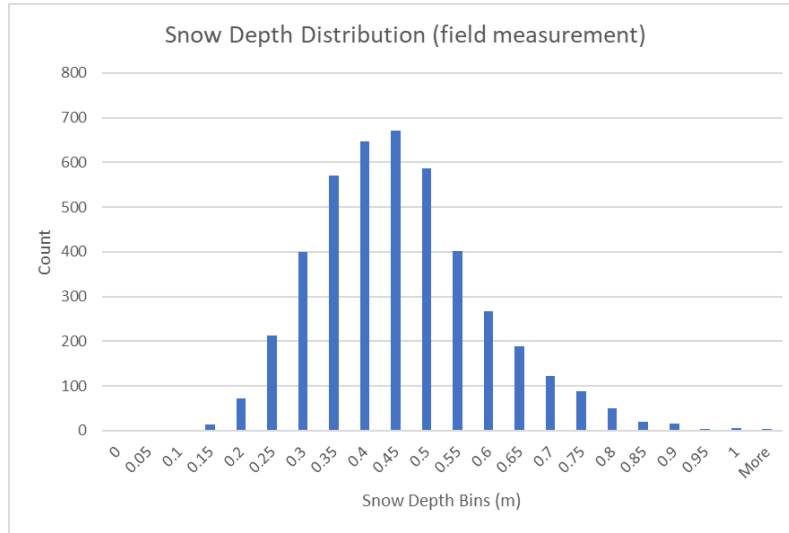


Figure 3.5: Histogram distribution of snow depth measurements (collected between 2019.03.21 and 2019.03.26)

In addition, Table 3.2 shows the mean and median snow depth and its standard deviation for each vegetation type from the field data. Among the six vegetation types, tree and tall shrub have the deeper snow but also a larger standard deviation while dwarf shrub, tussock and lichen are comparable in terms of mean, median snow depth and standard deviation.

Table 3.2: Mean and median snow depth and standard deviation for each vegetation type

	Tree	Tall Shrub	Riparian Shrub	Dwarf Shrub	Tussock	Lichen
mean (m)	0.56	0.63	0.49	0.42	0.42	0.43
median (m)	0.56	0.65	0.47	0.41	0.41	0.43
std (m)	0.21	0.16	0.17	0.13	0.13	0.12

3.3.4 TVC Vegetation Map

To account for the effect of different underlying vegetation on the radar waveform, a vegetation map was used to analyse spatial variability of vegetation type in the study region. A vegetation map of TVC was provided by Boike and Grünberg (2019), with a spatial resolution of 10 m. Vegetation is classified into 6 classes including tree, tall shrub, riparian shrub, dwarf shrub, tussock, as well as lichen as is shown in Figure 3.6. The map also contains a seventh class, water. The map covers the same area as the 2016 ALS point clouds data introduced above and is clipped to the study area.

According to Boike and Grünberg (2019), the vegetation map was generated through combining ALS point clouds and airborne orthophotos. Reference areas were chosen and vegetation types in these areas were classified manually, after which a set of parameters were calculated to characterize each vegetation type. Parameters include terrain attributes derived from ALS point clouds (e.g. slope, aspect, vegetation height, etc) as well as spectral attributes from orthophotos (e.g. average hue, average saturation, etc.).

In their study, Boike and Grünberg (2019) reported a final map accuracy of 87.7% for the validation dataset and 89.5% for the calibration dataset. However, it was noted that with a 10 m spatial resolution grid, mixing of different vegetation types was likely in a single cell, specifically for tall shrub, where the actual tall shrubs are sparse and between the tall shrubs are normally low vegetation.

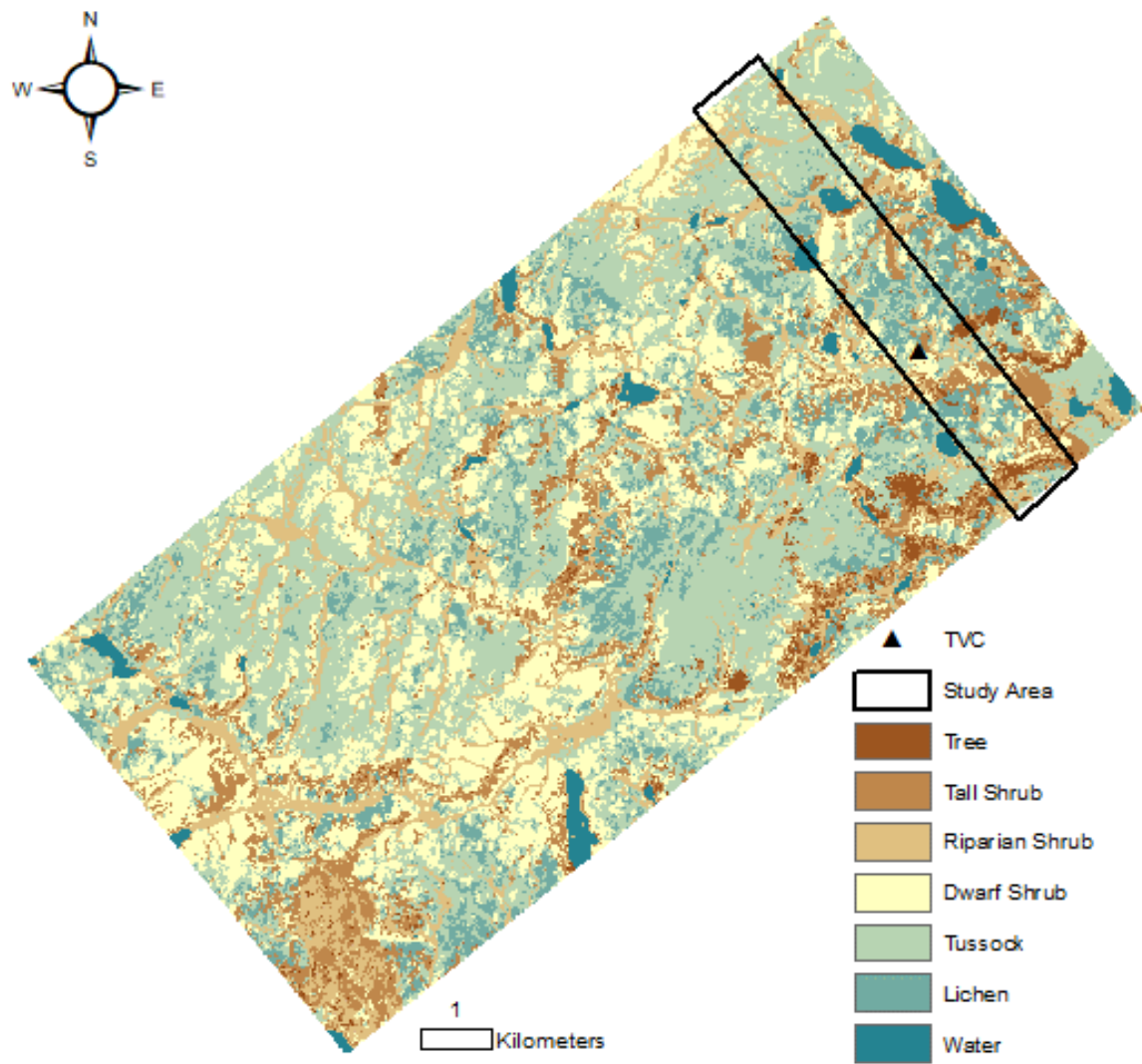


Figure 3.6: TVC vegetation map (Boike & Grünberg, 2019).

3.4 Methods

3.4.1 Pre-processing

Data pre-processing was divided into 2 components: the generation of a snow depth product for April 2019 and radar waveform pre-processing. Because the sparse field snow depth measurements did not spatially match the flight path of the airborne wide band radar, to provide snow depth reference for each radar data location, a snow depth product was generated by differencing the snow-off DTM from September 13th, 2016 from the snow-on DTM from April 10th 2019. In addition, based on the snow depth distribution from the field data, any snow depth that is above 2.5 m or a negative value is treated as invalid (NaN).

Four pre-processing steps were applied to radar waveform to prepare them for analysis. First, for each radar waveform (14050 range bins), the first 500 and the last 500 bins were removed to eliminate residual noise from these range bins, as is shown in Figure 3.7. More specifically, the maximum return of each waveform is assumed to be the snow-ground interface because the difference of dielectric properties between snow and soil is theoretically the largest amongst all different interfaces. However, for some waveforms, the maximum return is located in the beginning range bins or in the final range bins, and is an artefact of system processing rather than geophysical signal since both sets of range bins are located geometrically far from where the snow-ground interface is expected.

Second, a mean noise value was recorded for each waveform by averaging the first 100 returns after the first step. This was used as base reference to distinguish an interface

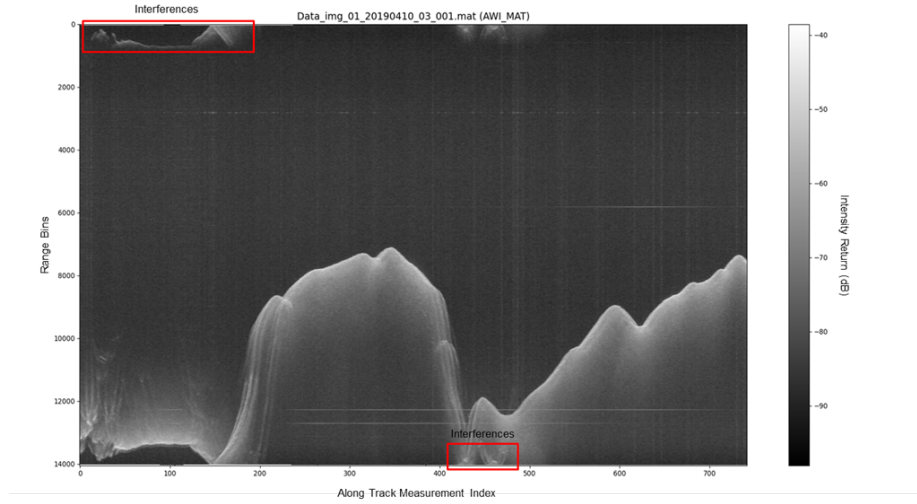


Figure 3.7: Interference in the beginning and final bins

peak in the waveform from noise. Third, the waveform was further reduced in length to 300 range bins, 150 bins before and 149 bins after the waveform maximum. The waveform maximum response is assumed to be the ground surface (snow-ground interface) which has the strongest power return. The reason for centering each waveform on its maximum was to provide a consistent reference when comparing different waveforms, regardless of ground elevation and flight elevation. This is especially significant and necessary step when conducting waveform classification because places that have the same snow depth may have a variety of waveform shapes due to differences in ground and flight elevation. However, the number of bins is arbitrary, and is largely influenced by the distribution of snow depth. Fourth, the waveforms were linearly normalized by dividing each return by its maximum response.

3.4.2 Evaluation of Lidar Snow Depth Product

Field snow depth data from King et al. (2020) collected between 2019.03.21 and 2019.03.26 were used as a ground reference to evaluate the lidar snow depth product obtained from differencing the 2016 and 2019 DTMs. The 1 m cells of the lidar product were converted into a point shapefile using the centroid of each cell. For each field snow depth measurement, its coincident lidar snow depth was obtained by selecting the closest converted lidar snow depth centroid within 1 m as is shown in Figure 3.8. Comparisons of field and lidar snow depths were conducted for each vegetation type. Distribution parameters for both data were calculated (mean, standard deviation, and median) in addition to obtaining mean absolute error (MAE).

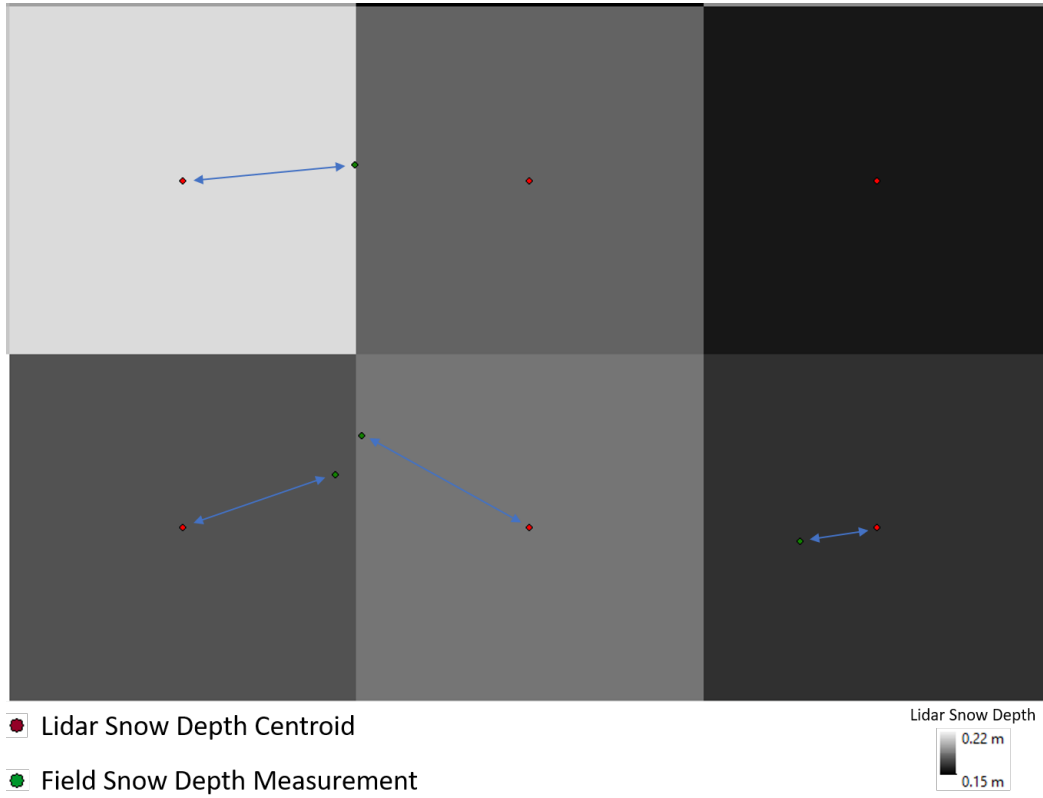


Figure 3.8: Matching field snow depth and lidar snow depth

3.4.3 Estimation of Snow Depth Using the Pulse Peakiness Approach

Julita et al. (2021) described the pulse peakiness approach as one of a range of peak picking methods based on the pulse peakiness of local peaks. Most methods have been applied for sea ice applications (Kwok et al., 2017; Newman et al., 2014; Kurtz et al., 2013). As a peak in the waveform can potentially indicate the boundary between two mediums with different dielectric properties, the approach was essentially searching for

peaks that represent the air-snow and the snow-ice interfaces. Logarithmic scale of the impulse return was used for locating the air-snow interface and linear scale of the intensity returns was used for obtaining the snow-ice interface. The reason for using linear for the snow-ice interface is that the difference of dielectric properties between snow and sea ice is large enough that using a linear scale will decrease the number of peaks detected while preserving the interface peak.

Eligible peaks were selected by applying a threshold over all detected peaks based on manually input parameters (log threshold & linear threshold) as well as the mean noise. A screening process was applied after identifying all the eligible peaks: if in linear scale, there were more than 5 eligible peaks, the waveform was considered too complicated to produce a robust result. For waveforms with fewer than 5 linear eligible peaks, two pulse peakiness values were calculated for each eligible peak. Unlike the pulse peakiness used in sea ice classification, Julita et al. (2021) defined these two pulse peakiness (left and right) as:

$$PP_l = \frac{S_{peak}}{\frac{1}{N} \sum_{i=1}^N S_{peak-i}} \times N \quad (3.2)$$

$$PP_r = \frac{S_{peak}}{\frac{1}{N} \sum_{i=1}^N S_{peak+i}} \times N \quad (3.3)$$

where S_{peak} is the linearly normalized intensity return, N is the bin range before and after the local peaks indicating how many bins are included into the calculation. It was recommended to use $N = 10$ because it was 2 times the range resolution of the radar (Julita et al., 2021). Finally, air-snow interface was defined to be the first eligible peak whose PP_l exceeded a set threshold while snow-ice interface was defined to be the last eligible peak

whose PP_r exceeded another set threshold.

Two changes were made to the pulse peakiness method that was developed for snow on sea ice. First, the process of defining the snow-ground interface using pulse peakiness "right" (equation 3.3) was removed and instead, it was assumed that the strongest return would be the snow-ground interface. For sea ice applications, there may be circumstances where the strongest return was the air-snow interface due to high brine concentrations in the lower part of the snow. Geldsetzer et al. (2019) has shown that the C-band radar response from snow on first year sea ice is complicated by the brine layer. The effect of the brine layer on other frequencies as wide band radar covers from 2GHz to 18GHz is less well described and possible future work is needed. Pulse peakiness right is introduced to address the situations where air-snow interface is the strongest return. However, for terrestrial snow, this is irrelevant. Second, additional screening process were added to overcome the complexity of the waveforms. Compared with snow on sea ice, it is expected that there are considerably more complex reflections in waveforms on land especially due to complicated vegetation interactions. Two kinds of complex interaction are expected: interactions between the wave, snow and emergent vegetation such as trees and tall shrubs, and interactions between snow and buried vegetation such as dwarf shrubs. Therefore, in addition to the original screening where the waveform must have no more than 5 eligible linear peaks to have a valid output, 3 more restrictions were added. To begin with, the normalized linear return of the air-snow interface peak must be less than 0.6. Moreover, the air-snow interface must be ahead of the strongest return (assumed snow-ground interface) in range bin sequence. Finally, a target peak was introduced to further filter out the waveforms that were too sophisticated. A target peak was defined to be the closest eligible

logarithmic peak to the reference air-snow interface, where the reference air-snow interface was obtained by converting the lidar snow depth into range bins using an assumed snow density. This is shown in Figure 3.9. For any waveform, if no target peaks were present within 20 range bins of the reference air-snow interface, then the waveform was considered to be ambiguous and was removed from the analysis.

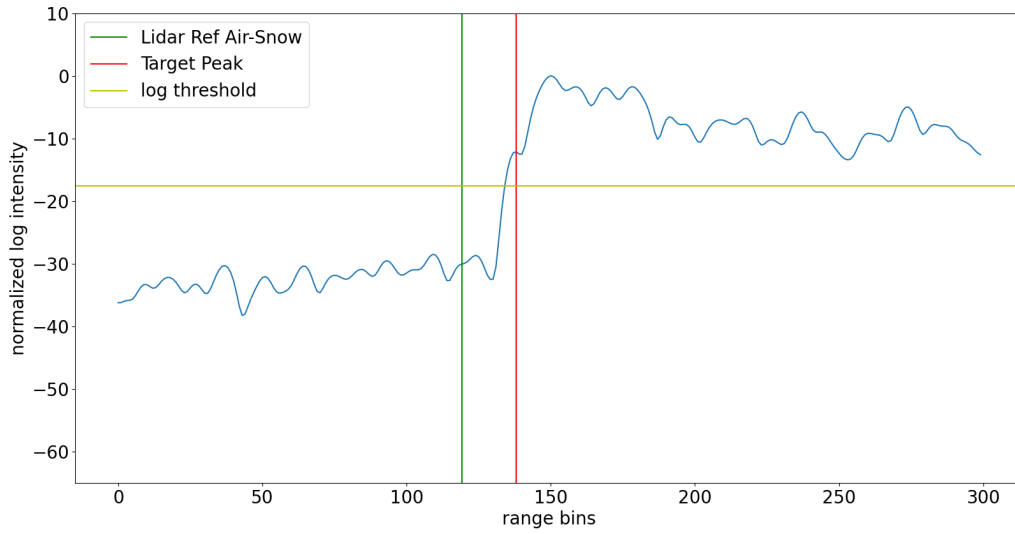


Figure 3.9: Example target peak defined by the closet eligible log peak to the lidar reference air-snow interface (waveform shown is from tussock)

In summary, the snow-ground interface was assumed to be the strongest return in the waveform. Therefore, the algorithm required 3 parameters instead of the 4 originally defined by Julita et al. (2021): 1) logarithmic threshold for determining eligible peaks in logarithmic scale; 2) linear threshold for determining eligible peaks in linear scale, and 3) the left pulse peakiness threshold for choosing the air-snow interface. All three parameters were optimized by using lidar snow depth data as reference. For both logarithmic and

linear threshold, the range was set to be between 0.1 and 0.9; for the left pulse peakiness threshold, the range was from 10 to 35. Optimization scoring for each set of parameters is shown below:

$$score = 0.75 \times MAE - 0.25 \times number\ of\ valid\ outputs \quad (3.4)$$

Both MAE and the number of valid outputs were normalized to ensure they were similarly scaled (from 0 to 1). In addition, to ensure there would be sufficient points for evaluation, a requirement for the number of valid outputs under any set of parameters must be above 10% of the total points was added. Optimization was performed for each vegetation type separately. Finally, the best set of parameters were defined as the ones with the lowest score (equation 3.4).

The optimization algorithm was run using 3 different divisions of the dataset. First, it was run on the entire dataset, and then the optimal parameters were applied to the entire dataset, which represented a best-case scenario. Second, it was run on an 80-20 split of the dataset; 80% of the data were used for optimization/training and 20% of the dataset for testing. Third, a 50-50 split of the dataset, 50% optimization/training and 50% testing.

3.4.4 Estimation of Snow Depth Using a Classification Approach

As snowpacks with different snow depths are expected to have different waveform shapes, traditional classification algorithms' capability of separating various waveform shapes into clusters as was seen in sea ice classification is the base principle of this ap-

proach. However, unlike sea ice classification where the end goal is different classes, snow depth retrieval presents a challenge where the separability of a 20 cm snowpack and a 25 cm snowpack in terms of waveform shapes is considerably less than that of sea ice and open ocean as is shown in 2.7. Consequently, a hybrid approach of combining classification and the pulse peakiness based interface method was developed and tested. Two basic assumptions were made: 1) there is a correlation between the location of the air-snow interface peak and two parameters, pulse peakiness left and logarithmic scale return of the interface peak; and 2) a classification algorithm is capable of clustering waveforms of similar snow depth together. When compared with a pure classification approach, utilizing a interface-based method allows the output of a specific snow depth and to some degrees overcome the low separability mentioned before while if compared with a pure pulse peakiness approach, the hybrid method does not need manually defined pulse peakiness left and logarithmic threshold values and instead uses the training data as knowledge to select the air-snow interface, similar to what was done in Kurtz et al. (2013).

The data were split equally into training and testing set. A classification was then performed on the training set. Theoretically, the classification algorithm could be any classification algorithm, an unsupervised K-means classification was chosen here to generate statistically unique clusters due to the lack of reliable training data. More specifically, to properly train a supervised classification algorithm or a neural network, a large quantity of waveforms similar to the two-peak shape (i.e. good quality waveforms with little interference) is needed. This approach was applied to each vegetation type separately. Therefore, the number of clusters for K-means was set based on the number of radar points in each vegetation type as well as several test runs to ensure there was good separability and no

empty clusters as is shown in Table 3.3. Within each cluster, for every waveform, a target peak was chosen, whose definition was the same as before, and the target peak's logarithmic scale return along with its left pulse peakiness as is shown in Figure 3.10. Then, for each cluster, an average logarithmic scale return and an average left pulse peakiness were calculated and recorded. Consequently, a representative target peak for each cluster was defined as a peak with the average logarithmic scale return and left pulse peakiness. Unlike the pulse peakiness approach which had multiple filters leading to an invalid output, the classification method had only the target peak filter as described before.

Table 3.3: Number of radar points and K-means clusters for each vegetation type

	Tree	Tall Shrub	Riparian Shrub	Dwarf Shrub	Tussock	Lichen
Number of Radar Points	83	439	334	504	1228	558
Number of K-means Clusters	16	32	32	32	64	32

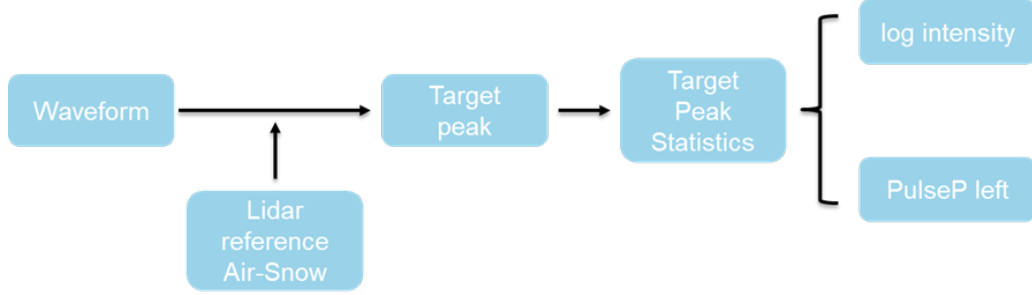


Figure 3.10: Workflow for every waveform in a cluster illustrating how lidar snow depth is used as the reference and how cluster statistics is calculated

After the clusters had been defined, all the waveforms in the testing set were assigned to the existing clusters based on their distance to the centroid of each cluster. Then, for each waveform in the testing set, eligible peaks were chosen the same way they were selected in the pulse peakiness approach. However, unlike the pulse peakiness approach where interfaces were defined by a threshold, a score to determine how similar each eligible peak was to the representative target peak of the cluster where waveform belonged was obtained using the following equation:

$$score = \omega_1 \times |log_{peak} - log_{target}| + \omega_2 \times |PP_{peak} - PP_{target}| \quad (3.5)$$

where ω_1 and ω_2 were the weight for each criterion. Currently, both weights are set at 0.5, but they can be adjusted. The reason for using equal weights is to assign equal importance to the shape of the peak and the impulse return strength of the peak, as is usually done in sea ice classification where multiple waveform characteristics are treated with equal importance (Shen et al. 2017). The air-snow interface was defined as the eligible peak with the lowest score.

3.4.5 Evaluation and Comparison of Both Approaches

For each vegetation type, snow depth retrieved by both pulse peakiness and classification approach were evaluated against the lidar snow depth data based on several criteria including number of points with valid output, MAE, Pearson correlation as well as p-value. Histograms were generated to determine how similar the distribution of snow depths was; in addition, scatterplots comparing retrieval and ground truth as well as comparing pulse peakiness retrieval and classification retrieval were produced to investigate whether there was any pattern in overestimation and underestimation.

3.5 Results

3.5.1 Evaluation of Lidar Snow Depth Product

Figure 3.11 shows the 1m spatial resolution lidar snow depth product from the 2016 snow-off DEM and 2019 snow-on DEM. The snow depths were around 30-40 cm in relatively flatter grounds mostly covered by tussocks, lichen and dwarf shrubs. This is similar to the distribution of field snow depth measurements despite there being a two week gap between when the field measurements and wideband radar measurements were made. In areas near streams or some valleys in the bottom part of the study area, the snow depth was greater than 60 cm, as indicated by the bright white in figure 3.11. This is due to the collection of snow in these gulleys and depressions.

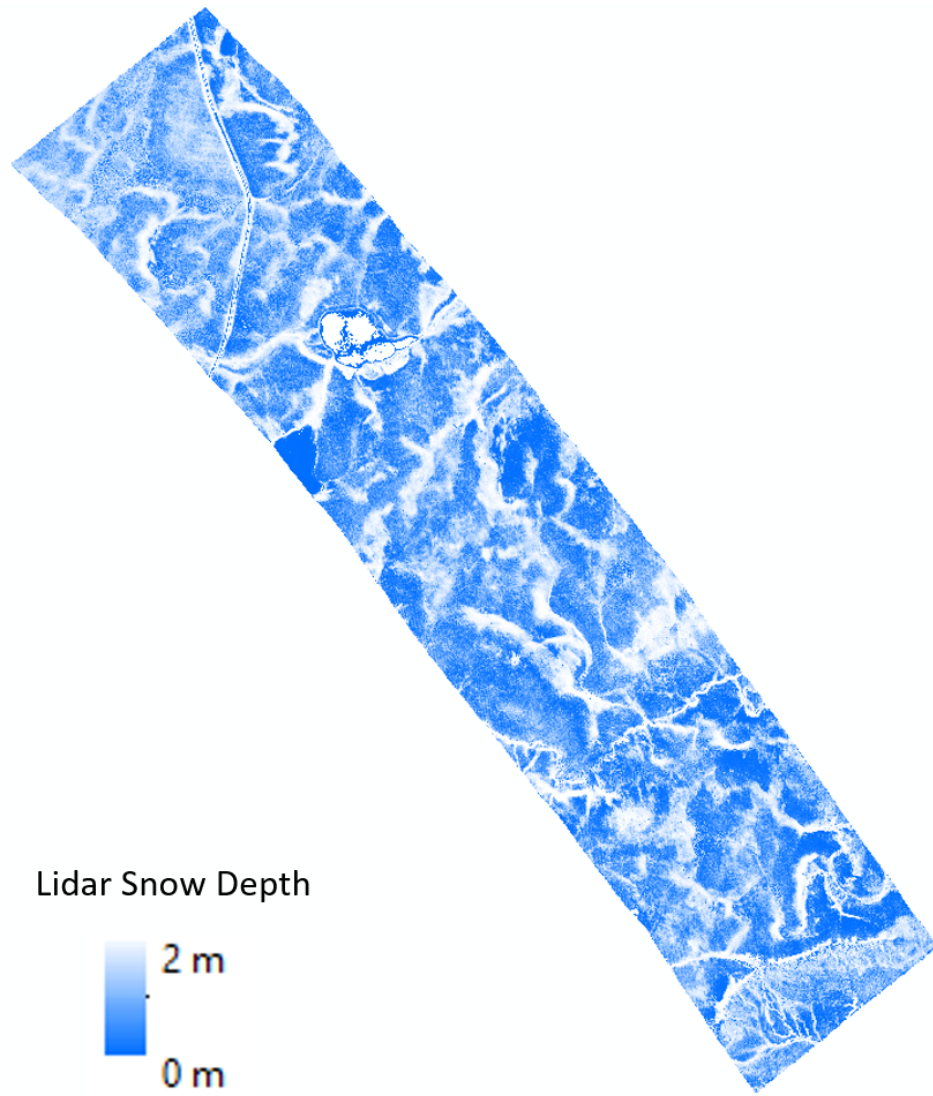


Figure 3.11: Lidar snow depth product for the study area

There were 4334 field snow depth measurements used for the evaluation of lidar snow depth product. Table 3.4 shows the results of comparing the lidar snow depth with field measurement. Treating field snow depth as the reference, the lidar estimates underestimate snow depth for all six vegetation types, especially for tree, riparian shrub, and dwarf shrub

classes as is shown in Figure 3.12. It is necessary to note that there was a 14-day difference between the time when these two data were collected. According to local weather station data, the temperature remained below zero with no significant precipitation event, however, some snowmelt might have been present in the field which could explain part of the reason for the underestimation. Further work to verify this possibility is ongoing. In addition, it is possible that the underestimation was the product of mismatch between snow-off and snow-on dates (the snow-off DEM used was from 2016); the underlying terrain or vegetation may have changed, leading to an inaccurate lidar snow depth. Equally, there could be an over-probe of up to 7.6 cm for the magnaprobe, as is mentioned in King et al. (2020), which would contribute to field snow depth being relatively greater.

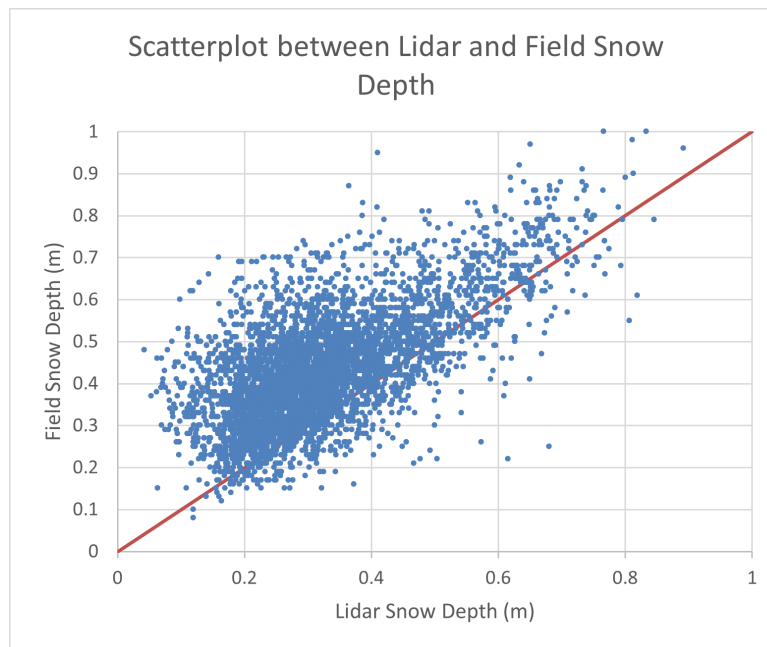


Figure 3.12: Scatterplot comparing lidar and field snow depth

Table 3.4: Evaluation of lidar snow depth

	Tree		Tall Shrub		Riparian Shrub		Dwarf Shrub		Tussock		Lichen	
	Field	Lidar	Field	Lidar	Field	Lidar	Field	Lidar	Field	Lidar	Field	Lidar
mean (m)	0.56	0.45	0.63	0.53	0.49	0.38	0.42	0.28	0.42	0.33	0.43	0.35
std (m)	0.21	0.20	0.16	0.17	0.17	0.16	0.13	0.11	0.13	0.11	0.12	0.12
median (m)	0.56	0.39	0.65	0.57	0.47	0.34	0.41	0.26	0.41	0.31	0.43	0.33
RMSE (m)	0.19		0.15		0.17		0.19		0.13		0.13	
bias (m)	-0.11		-0.10		-0.11		-0.15		-0.09		-0.08	
std.bias (m)	0.16		0.11		0.13		0.12		0.09		0.10	
count	13		151		231		1316		1456		1167	

3.5.2 Pulse Peakiness Approach Results

Table 3.5 shows the optimal pulse peakiness parameters for each vegetation type using a 100-0 split, and Figure 3.13 shows the retrieved snow depth variation for each vegetation classes. The Pulse Peakiness estimates used are from the optimal parameters obtained for a 100-0 split. In addition, parameter optimization results using a 80-20 and a 50-50 split can be found in Appendix A. The reason for using a 100-0 split is that the randomness in splitting the data has a significant effect on the optimization results, especially for tree, tall shrubs and riparian shrubs. Moreover, a 50-50 split may result in insufficient amount of the waveforms for the algorithms, which will enhance the effect of the randomness in splitting. It is necessary to note that this is the first attempt of developing the method and a stratified random split could be tested in the future.

Table 3.5: Optimal pulse peakiness parameters for each vegetation type

	Log	Lin	PP_L	MAE(m)	Bias(m)	RMSE(m)	Pearson	Valid / Total Points
Tree	0.6	0.5	25	0.12	-0.03	0.18	0.65	11 / 84
Tall Shrub	0.3	0.8	20	0.27	0.13	0.34	0.20*	226 / 441
Riparian Shrub	0.2	0.7	15	0.29	0.21	0.38	0.35*	209 / 321
Dwarf Shrub	0.4	0.5	25	0.13	0.04	0.19	0.29*	135 / 500
Tussock	0.5	0.5	30	0.09	-0.04	0.12	0.55*	141 / 1224
Lichen	0.3	0.5	30	0.10	-0.01	0.15	0.46*	74 / 558

*indicates significant P-value at $\alpha < .01$

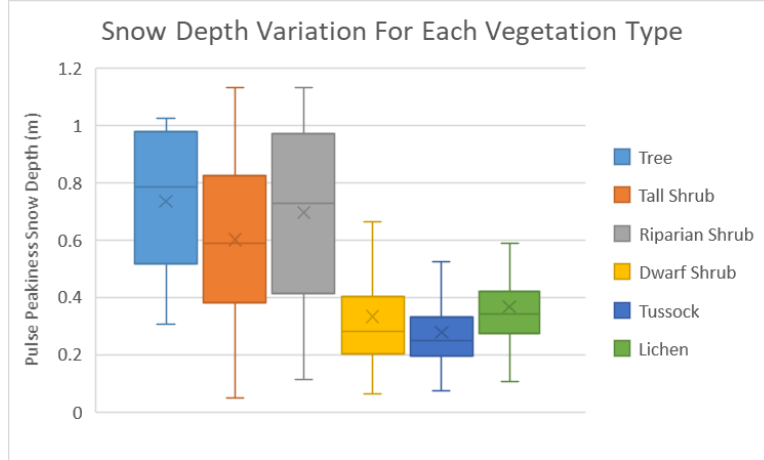


Figure 3.13: Pulse peakiness snow depth variation for each vegetation type.

Overall, from Figure 3.13, snow depth variation is considerably higher for trees, tall shrubs and riparian shrubs in addition to them having generally thicker snow, possibly due

to the higher vegetation.

For trees, despite a relatively high Pearson correlation and a low MAE, the P-value is not significant because there are few points. The scatterplot (Figure 3.14 a) for trees further proves the high correlation and low MAE where most points fall closely around the 1:1 line. While this is promising, due to the sparse distribution of trees in the study area, the valid points shown here are likely to be returns from spaces between trees instead of actual tree returns. This is further reinforced by the similarity of parameters between tree and three other low height vegetation (dwarf shrub, tussock, & lichen). The low linear threshold at 0.5 filtered out all waveforms from actual tree returns since these waveforms are highly complex with a large number of peaks. Both tall shrub and riparian shrub have low Pearson correlation coefficients (and high MAE values) of 0.2(0.27m) and 0.35(0.29m) respectively. The scatterplots (Figure 3.14 b and c) show the characteristics of relationships between radar and reference estimates with no discernible relationship for between tall shrub and lidar and a general overestimation of the radar estimates for tall shrubs. An additional observation here is the high linear threshold of these two vegetation types compared with low height vegetation like Tussock. As a high linear threshold leads to complex waveforms (ones often challenging for this retrieval approach) being more likely to be fed into the retrieval, the optimal linear threshold should be as low as possible provided that there will be sufficient valid output. For tall shrub and riparian shrub, a high optimal linear threshold indicates that fewer waveforms from these two vegetation types resembles the ideal two-peak shape and the optimization was forced to include more complex waveforms to satisfy the criteria of having enough valid outputs.

For dwarf shrub, the MAE is 0.13 m with a Pearson correlation of 0.29. Due to

the lower linear threshold, the increase in total points does not lead to an increase in valid points. From Figure 3.14d, the reason for a low Pearson correlation is the overestimation of snow depth by the Pulse Peakiness algorithm. More specifically, there are a series of points with increasing Pulse Peakiness snow depth, whose Lidar snow depths remain constant at around 0.2 and 0.3 m. This is most likely to be caused by the waveform having multiple peaks with close to maximum returns and will be further discussed later. In comparison, the underestimation of the retrieval retained the observable pattern from riparian shrub as most of the points above the 1:1 line still falls into a linear pattern.

According to Boike and Grünberg (2019) tussock and lichen classes are characteristically similar in classification terms. However, they are functionally different biophysically. They behave similarly in terms of their radar waveform. Compared with other vegetation types, they have the best retrieval results with MAE of 0.09 m for tussock and 0.1 m for lichen and a Pearson correlation of 0.55 and 0.46 respectively. This result is to some degree achieved by severely constraining the waveform selection for use with the algorithm, as indicated by the low valid / total points ratio. In other words, for tussock and lichen, the two thresholds are relatively more effective at filtering out waveforms that are difficult for the algorithm to accurately obtain interfaces. Both the decrease in MAE and the increase in Pearson correlation are illustrated in the scatterplot of Figure 3.14e and f, where the overestimation that results in the low correlation for dwarf shrubs has mostly disappeared. It is noted that the pattern for underestimation of the method persists in both of these vegetation type, indicating that this could possibly be addressed with empirical bias correction.

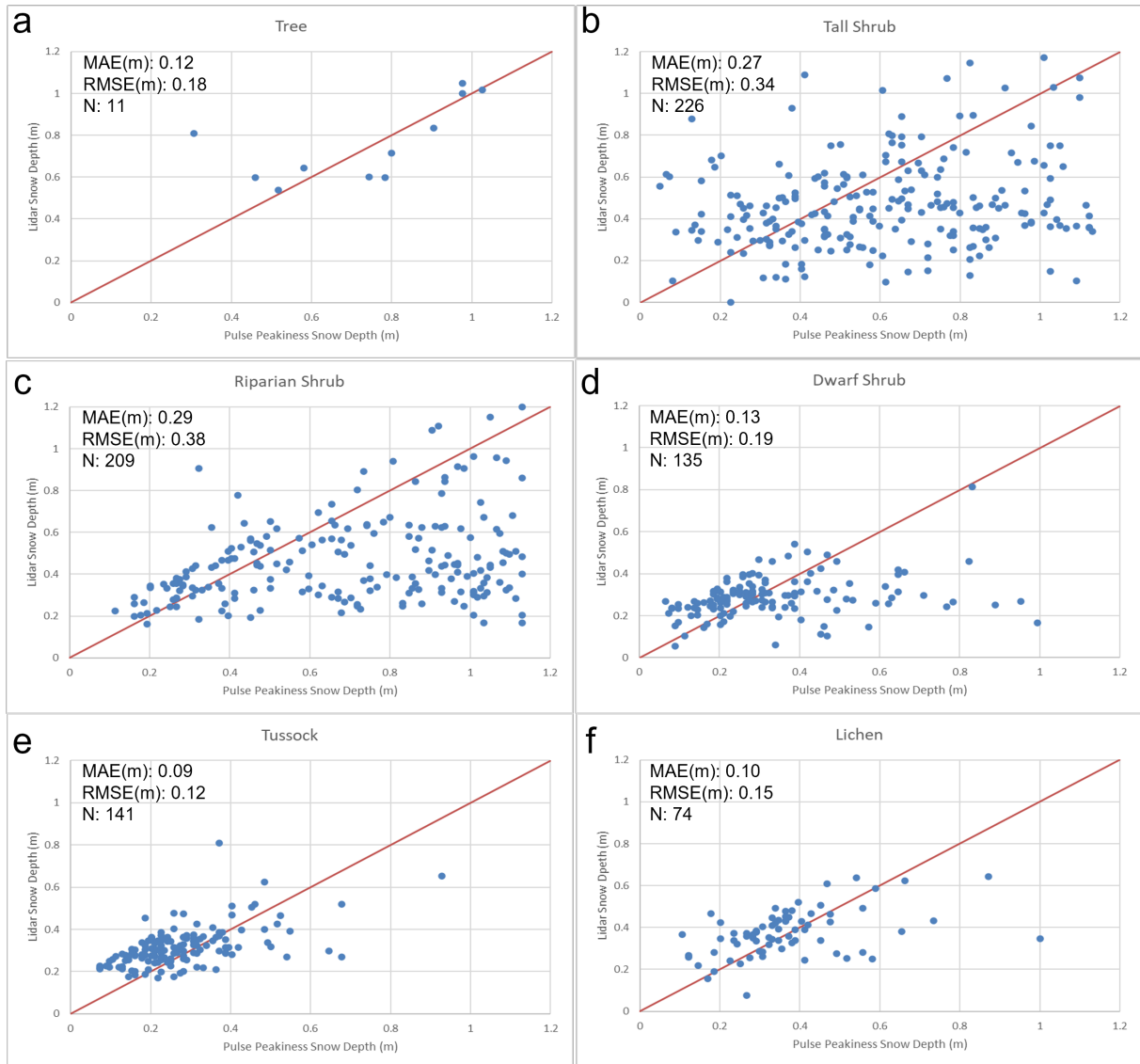


Figure 3.14: Scatterplots between pulse peakiness and lidar snow depth estimates for the following classes a) tree, b) tall shrub, c) riparian shrub, d) dwarf shrub e) tussock and f) lichen.

3.5.3 Classification Approach Results

Table 3.6 shows the snow depth retrieval results for the classification approach and Figure 3.15 demonstrates the corresponding scatterplots of classification snow depth compared with lidar snow depths. Results for all 5 runs can be found in Appendix A and only the best results are shown here. For the trees land cover category the MAE and Pearson correlation are 0.21 m 0.43 respectively, but like the pulse peakiness result, the correlation is not significant due to the low the number of paired measurements. From the scatterplot (Figure 3.15a), underestimation of the snow depth by the retrieval accounts for a large proportion of the error possibly due to tree canopy interfering with ALS resulting in either an underestimation of snow depth by radar or an overestimation by the Lidar due to erroneous returns from the canopy rather than under-story snow. Similar to the pulse peakiness results, the best estimates that have minimal error, i.e. sitting on the 1:1 line, are likely to be from areas between trees where low height tussock or lichen vegetation are located. The tall shrub land cover category has the lowest agreement with the lidar reference data of the six classes, with a MAE of 0.24 m and a Pearson correlation of 0.32. As shown in Figure 3.15b, the distribution of the points appear to be randomly distributed, especially those under the 1:1 line. Compared with the tall shrub class, the riparian shrub class has a lower MAE at 0.23 m and a higher Pearson correlation of 0.45. Once again, there are several estimates that underestimate the snow depth relative to the lidar estimates. For the dwarf shrub class (Figure 3.15d) there is a similar pattern to the pulse peakiness method in that there is an overestimation of snow depth by the radar classification approach relative to the lidar measurements. Consequently, despite a better

MAE of 0.19 m, the retrieval snow depth overestimation significantly affects the correlation. Both tussock and lichen are comparable in performance with a slightly improved MAE but similar Pearson correlation compared with dwarf shrub. More specifically, for those with a retrieval snow depth below 0.4 m, there is no observable relationship between Lidar and classification snow depth where for a Lidar snow depth between 0.2 and 0.3 m, the retrieval snow depth can vary from close to 0.05 m to 0.4 m. And this would suggest for a large number of these waveforms, the pulse peakiness left and the logarithmic scale return used for the classification approach can not uniquely identify the air-snow interface.

Table 3.6: Best Snow Depth Retrieval Results

	MAE(m)	Bias(m)	RMSE(m)	Pearson	Valid / Total
Tree	0.21	-0.08	0.27	0.43	29 / 42
Tall Shrub	0.24	0.07	0.31	0.32*	150 / 221
Riparian Shrub	0.23	0.03	0.30	0.45*	63 / 161
Dwarf Shrub	0.19	0.04	0.26	0.23*	167 / 250
Tussock	0.15	-0.04	0.21	0.26*	320 / 612
Lichen	0.18	-0.01	0.23	0.23*	185 / 279

*indicates significant P-value at $\alpha < .01$

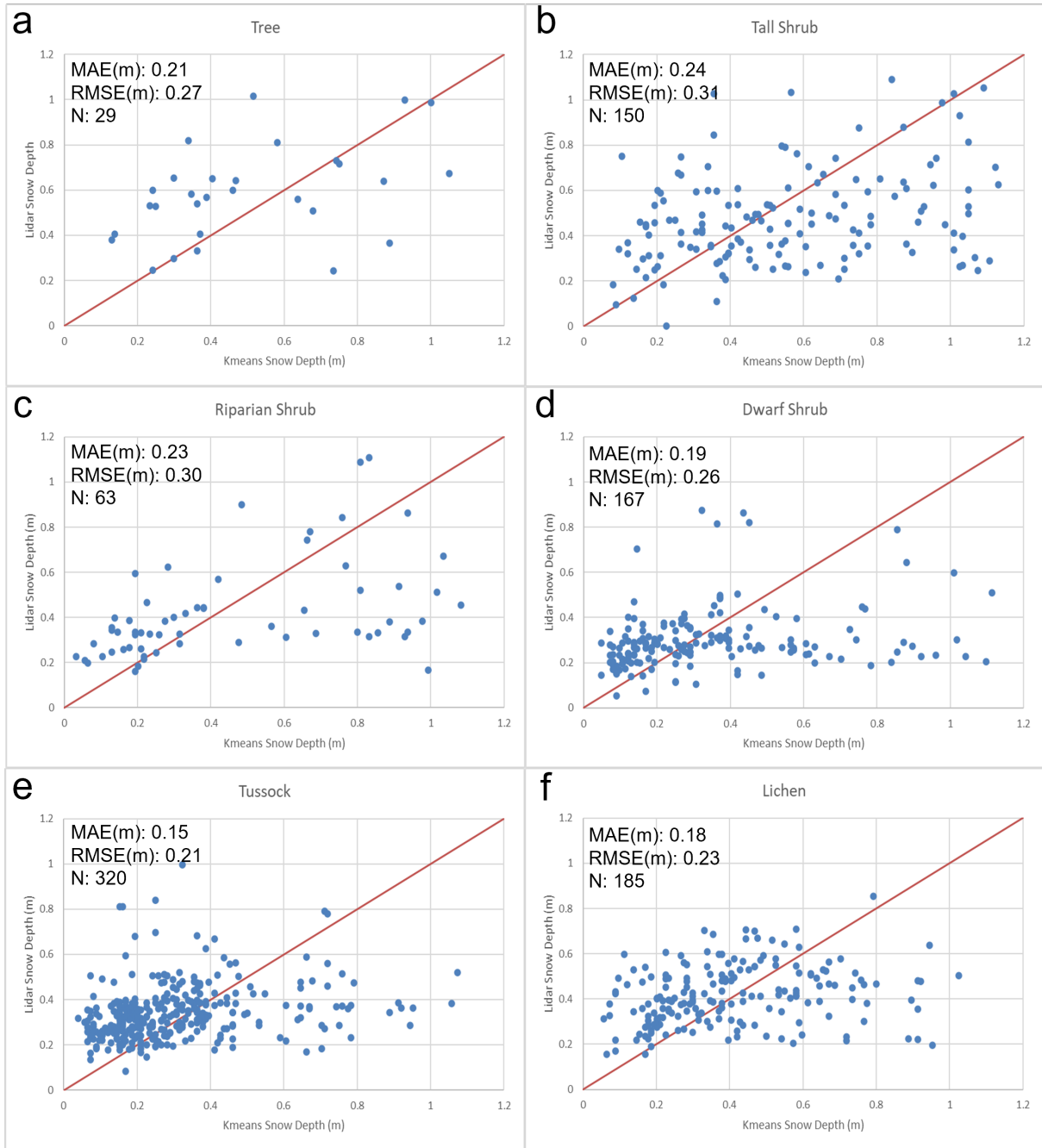


Figure 3.15: Scatterplots between Classification and Lidar snow depth estimates for the following classes a) tree, b) tall shrub, c) riparian shrub, d) dwarf shrub e) tussock and f) lichen.

As for snow depth variations for each vegetation type, Figure 3.16 shows the box plot for the classification retrieval results. A higher variation for trees, tall shrubs and riparian shrubs is also present, similar to that for the pulse peakiness retrieval. Although the three classes with higher vegetation (trees, tall shrubs and riparian shrubs) have a generally thicker snow, the maximum snow depths for the other classes are relatively high as well.

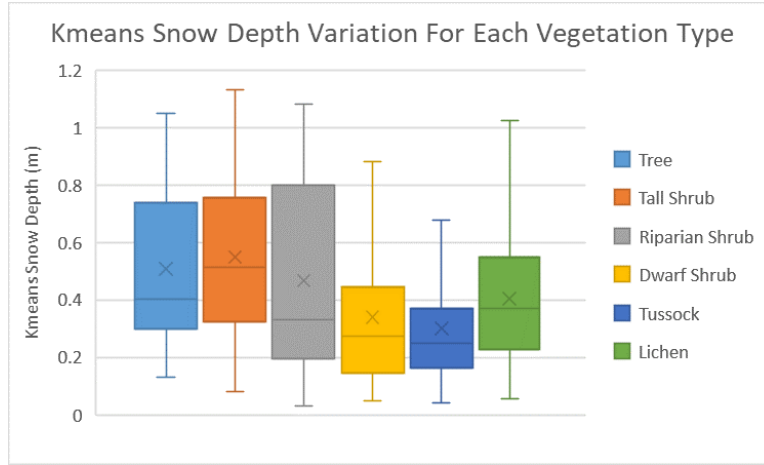


Figure 3.16: Classification snow depth variation for each vegetation type.

3.5.4 Comparison Between the Two Approaches

For the six vegetation types, it is clear that snow depth retrievals in trees, tall shrubs and riparian shrub classes is very challenging for both methods from the above section. Therefore, the comparison will focus on dwarf shrubs, tussock, and lichen.

An important consideration lies in the number of valid points and the number of waveform filters. Table 3.7 shows the summary for valid points / total points for all three

vegetation types. More specifically, the valid point number for the classification approach is an average of all 5 runs while for pulse peakiness approach, it is taken from the 50-50 split scenario and is also an average of all 5 runs.

Table 3.7: Comparison of valid points percentage between pulse peakiness and classification approach

	Pulse Peakiness	Classification
Dwarf Shrub	67 / 250	173 / 250
Tussock	188 / 612	316 / 612
Lichen	104 / 279	187 / 279

Even though the pulse peakiness approach outperforms the classification approach in all three vegetation types, this is partially due to the quality of waveforms fed into the retrieval algorithms. The two thresholds in the pulse peakiness approach together with other added filters lead to the algorithm accepting fewer waveforms suitable for identifying the interfaces. By comparison, these criteria are not applied in the classification thereby providing more valid points for the assessment although, perhaps at the expense of a degraded performance.

In addition, the way the classification approach identifies the air-snow interface is not as robust compared with the Pulse Peakiness approach despite the fact that they are using the same peak characteristics. Similar to the NSIDC algorithm developed in Kurtz and Farrell (2011) and Kurtz et al. (2013), the classification approach determines the interface based on pre-existing knowledge. While it is plausible in theory, using training data as existing knowledge is far less reliable and susceptible to the randomness of splitting the data. The flaws in identifying the interfaces can be further demonstrated by Figure [3.17](#)

and Figure 3.18. Figure 3.17 shows the comparison between classification snow depth and pulse peakiness snow depth wherein the classification method consistently underestimates the snow depth. This is reinforced by Figure 3.18 which shows the relationship between classification and Lidar snow depth for those with a valid Pulse Peakiness output (i.e. good quality waveforms). Even with suitable waveforms, the classification approach generally underestimates snow depth, indicating that the method itself is not robust.

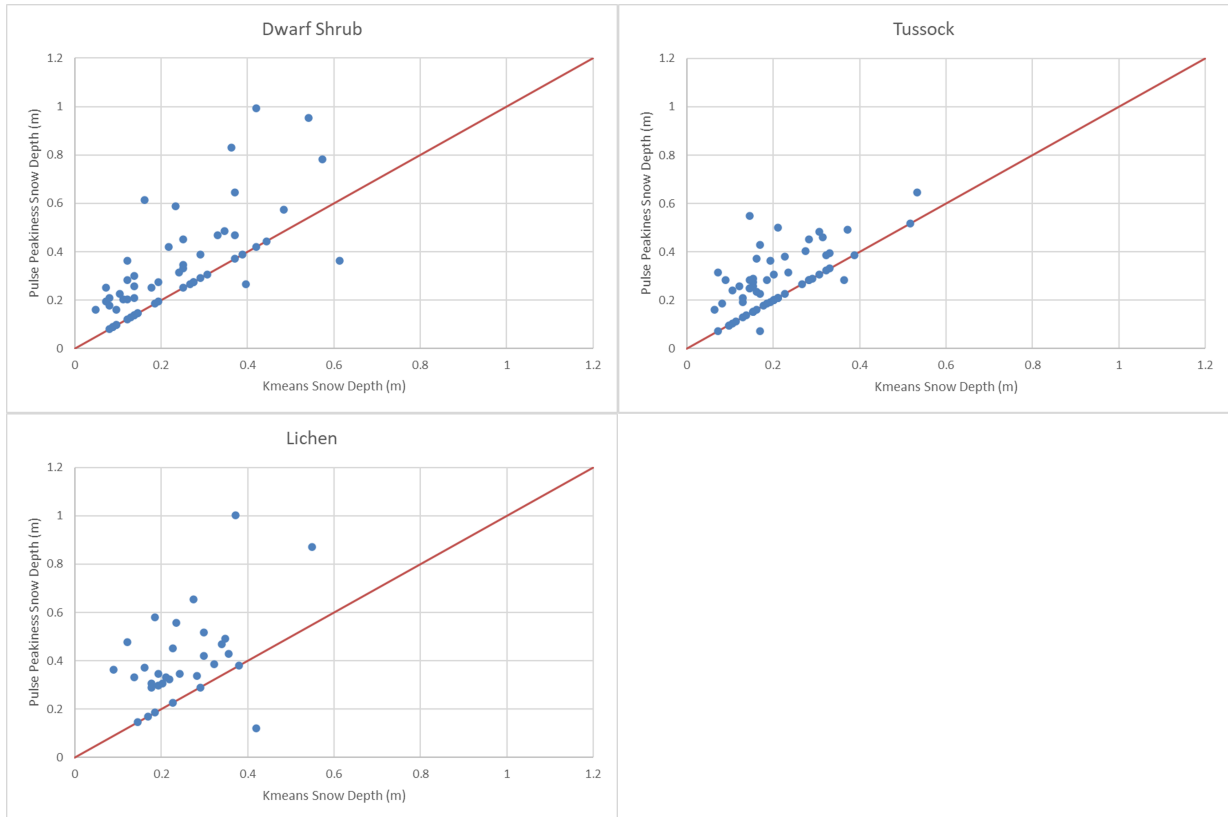


Figure 3.17: Comparison between classification (K-means) and Pulse Peakiness snow depth

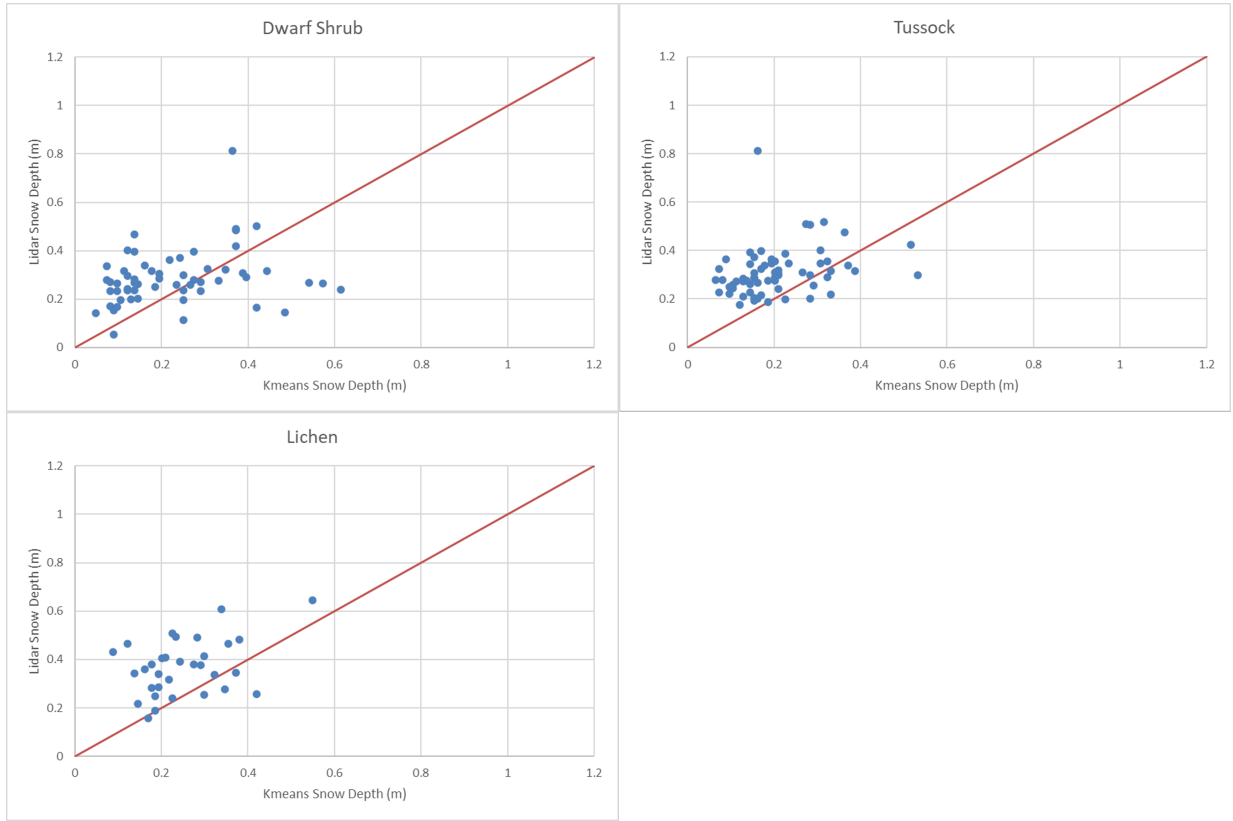


Figure 3.18: Comparison between classification (K-means) and Lidar snow depth for those points with a valid Pulse Peakiness output

3.6 Discussion and Conclusion

3.6.1 Underestimation and Overestimation

Both pulse peakiness and classification approaches have similar underestimation and overestimation characteristics but are the result from difference sources. With respect to underestimation, it is likely to not be from waveform / radar measurement and it follows

a clear pattern, meaning that it can be mitigated to a degree with a bias correction. On the contrary, the overestimation is mainly resulted from the waveforms and needs other more delicate solutions.

The tussock land cover class is used to illustrate the underestimation problem as this landcover class has the most radar points. When running the pulse peakiness approach on the 100-0 split scenario, out of 141 valid outputs, 57 has an absolute error of 10 cm or above. And out of these 57, 39 underestimated snow depth in a similar way as is shown in Figure 3.19. The Lidar reference air-snow interface falls into areas which could be considered noise in the logarithmic scale intensity return while the algorithm was successful in identifying the first peak that satisfied the logarithmic threshold.

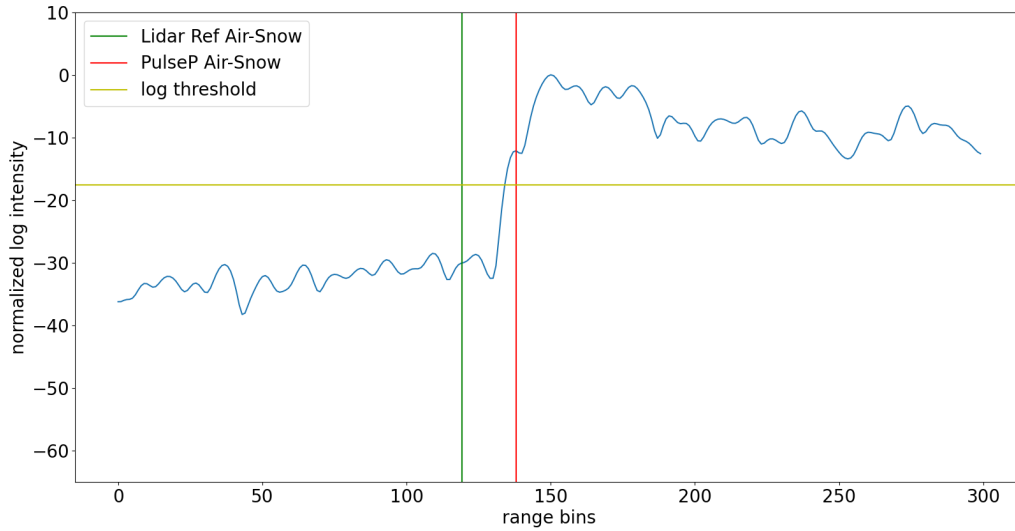


Figure 3.19: Sample underestimating waveforms for pulse peakiness where the Lidar reference air-snow interface falls below the logarithmic threshold

As pulse peakiness left value alone normally cannot uniquely identify a peak as the air-snow interface, it is likely that before the interface peak, there would be numerous peaks with similar pulse peakiness left values. Consequently, it is extremely important to exclude as many non-interface peaks before the interface peak as possible to increase the chance of it being the first that exceeds the pulse peakiness left value threshold. This is achieved through the application of a logarithmic threshold. By lowering the threshold to accommodate the type of underestimation shown in Figure 3.19 the result would most likely lead to an overestimation of snow depth. An alternative is to apply a post algorithm bias correction based on this dataset.

To achieve this, the data were split in half into training and testing subsets, for which an average of underestimation was calculated for all the waveforms in the training subset. This average was used as the bias correction for all the waveforms in the testing set. Figure 3.20 shows the comparison between the un-corrected scatterplot and the corrected scatterplot. The average bias for the training set in this case is approximately -4 cm and the MAE after the correction decreased approximately 1 cm. The reason for such a small improvement is that the bias correction is applied to all the waveforms in the testing sets including those that are grossly overestimating. More intricate solutions to this bias correction, such as the identification and character generalization of underestimating waveforms should further improve the results.

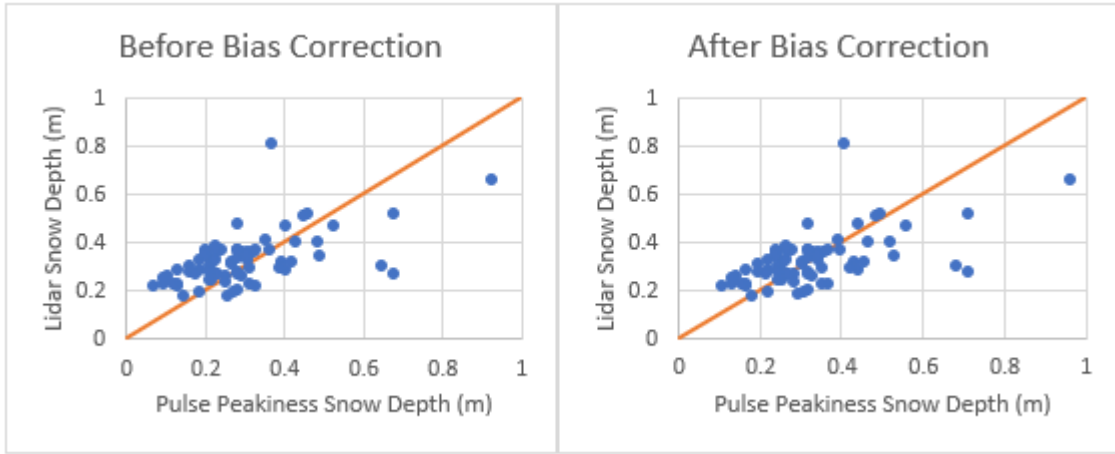


Figure 3.20: Before and after bias correction

With respect to the overestimation problem, the main reason for the overestimation in both pulse peakiness as well as classification approach is the complexity of the waveform, especially for the classification approach. More specifically, the scoring and matching system in the classification approach is considerably less robust when encountering a complex waveform. A complex waveform here is defined as a waveform that has more than one peak that is close to the maximum return, in other words, a waveform that has no clear single maximum return. There can be a number of reasons why a waveform may have multiple “maximum” returns, but one plausible possibility is that a feature in the snowpack has a similar dielectric response to the ground surface.

This kind of waveform is inconsistent with the baseline assumption of a clear single maximum return peak being the snow-ground interface, and therefore disrupts the snow depth calculation which is the distance between air-snow peak and the maximum return. Figure 3.21 and 3.22 shows examples of overestimation in dwarf shrubs and lichen.

To counteract this problem, one possible way is to apply a signal-to-noise ratio (SNR) threshold that excludes waveforms like the one shown in figure 3.21. For such waveforms, adding another screening criterion that requires the waveform to have only 1 peak whose return is above a certain value should help filter them out. The challenge would be to set the correct threshold so that it does not filter out too many of the waveforms leaving the algorithm with insufficient valid outputs.

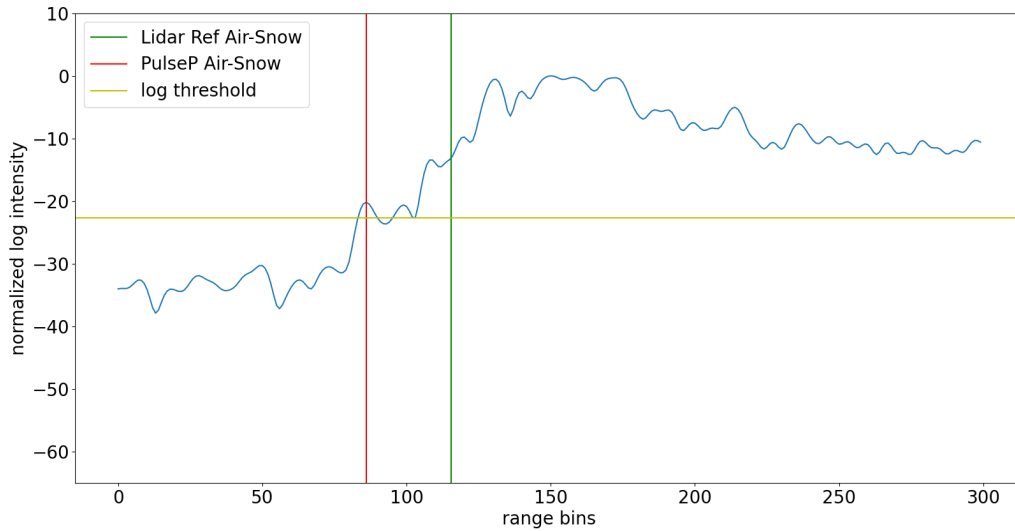


Figure 3.21: Sample overestimation for dwarf shrubs where there are multiple peaks close to the maximum peak.

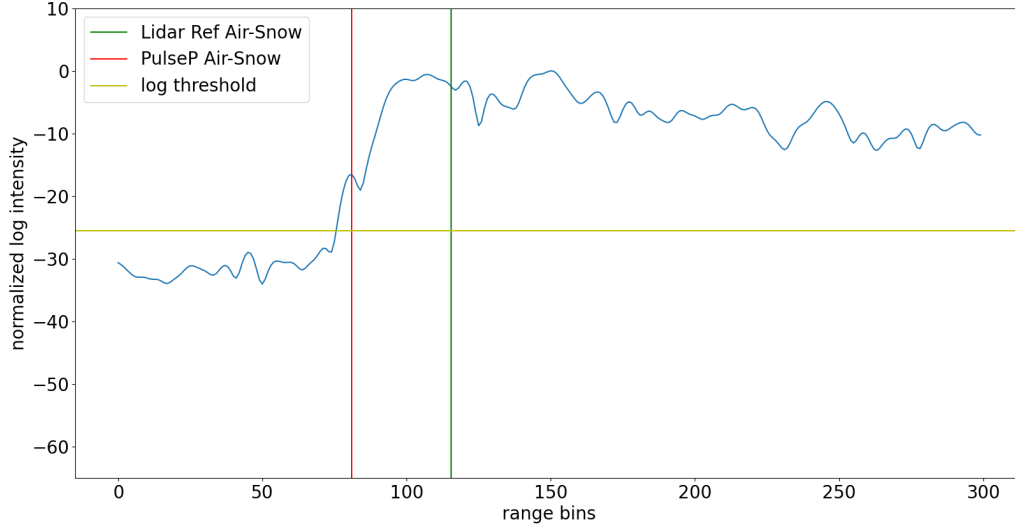


Figure 3.22: Sample overestimation for lichen where there are multiple peaks close to the maximum peak.

3.6.2 Error of Omission for the Pulse Peakiness Approach

Due to the limited number of valid outputs of the pulse peakiness approach, an analysis was done on the error of omission for dwarf shrub, tussock and lichen. The reason for excluding trees, tall shrub and riparian shrub is that these landcover types have more complex waveform responses and therefore are omitted from the algorithm.

Table 3.8 shows the distribution statistics for Lidar points belonging to dwarf shrub (500), tussock (1224), and Lichen (559) as well as those with a valid pulse peakiness output (135, 141, & 74). It demonstrates that for dwarf shrub and lichen, even though pulse peakiness filters excluded parts of the data, the distributions are still similar. However, for

Tussock, the pulse peakinesss distribution has shifted slightly to lower snow depth, which is further reinforced by the histogram shown in Figure 3.23. The pulse peakiness output is lacking representation of snow depth that is above 0.4m.

Table 3.8: Distribution statistics for Lidar snow depth and pulse peakiness valid output

	Mean (m)	Median (m)	Standard Deviation (m)
Dwarf Shrub			
Lidar	0.32	0.30	0.14
Pulse Peakiness	0.33	0.28	0.19
Tussock			
Lidar	0.38	0.36	0.14
Pulse Peakiness	0.28	0.25	0.13
Lichen			
Lidar	0.44	0.42	0.16
Pulse Peakiness	0.37	0.34	0.16

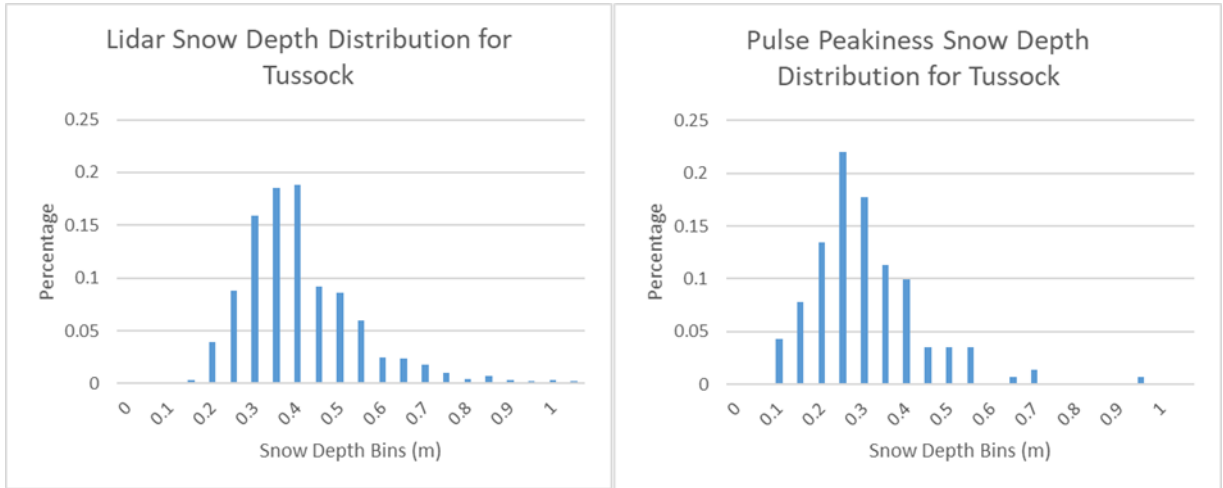


Figure 3.23: Comparison of snow depth distribution between Pulse Peakiness valid output and Lidar points

3.6.3 Conclusion

Overall, using a wide band radar system for snow depth retrieval on land can work relatively well for dwarf shrub, tussock and lichen classes. More specifically for the two approaches tested, both pulse peakiness and the classification approach are not suitable for areas with complex vegetation responses characterized by multiple peaks in the waveform, specifically trees, tall shrubs, and riparian shrubs in this case. For these classes, because the vegetation interacts with radar wavefront before and after it is attenuated by the snow, the waveforms are more complex in characteristic. For short height vegetation types, specifically shrub, tussock, and lichen, the pulse peakiness method tended to outperform the classification approach. However, a more constrained screening process was applied to the pulse peakiness method to ensure that waveforms better matched the ideal two interface peaks pattern. Consequently there were fewer waveforms that could be used to estimate snow depth but the quality of the pulse peakiness retrieval was more robust than the classification one.

Chapter 4

Discussion and Conclusion

4.1 Summary

Wide band radar has been used for snow depth retrieval on sea ice, mainly using data from Operation Icebridge. Core principle behind various retrieval algorithms is based on the ability of a wide band / FMCW radar to detect different objects at different distances due to its constantly changing frequency. More specifically, the principle in estimating snow depth is to identify two interfaces, the air-snow interface and the snow-ice interface which when differenced provide an estimate of snow thickness. Where algorithms differ is how the two interfaces are acquired and how the algorithm is optimized for a specific radar system.

Using the data from Icebird 2019 over the Trail Valley Creek research station and the pulse peakiness methods from Julita et al. (2021), it was possible to retrieve snow depth

over land. However, two key factors explain the differences between the wide band radar observations over land that are not present in observations over sea ice: a) the presence of surface vegetation impacts the waveform response, and b) the presence of micro-terrain feature like hummocks. In addition, it is believed that the off-nadir influences usually resulted from slopes of the ground is less likely because of the presence of hummocks. Both factors lead to terrestrial waveforms being generally more complex than those for sea ice. Consequently, additional criteria to determine whether a waveform was too ambiguous for the algorithm to identify the two interfaces was added to the method developed by Julita et al. (2021). Amongst the six vegetation types classified by Boike and Grünberg (2019), trees, tall shrubs, and riparian shrubs were relatively less suitable for estimating snow depth from the pulse peakiness method while dwarf shrubs, tussocks and lichens were more favourable.

An additional method inspired by the classification of radar altimeter data from sea ice classification (Müller et al., 2017; Zygmuntowska et al., 2013; Shen et al., 2017; Ricker et al., 2015; Aldenhoff et al., 2019) was also developed and tested with this dataset acquired over land. The difference between different snow depths were considerably less separable than the difference between for example first-year ice and multi-year ice. Moreover, a number of waveform characterizing parameters utilized in sea ice classification were not applicable due to the significantly more complex (i.e. larger number of peaks) waveforms. As a result, the pulse peakiness concept was combined with the classification method to develop a hybrid approach. Similar to the pure pulse peakiness approach, estimation of snow depth in tussocks and lichens produced the best performance. However, compared with the pulse peakiness approach, the proposed classification approach underperformed

as a result of the inadequacy of how well a linear k-means classification can generalize complex waveforms. A more advanced machine learning approach might be applicable if adequate training data were acquired.

Overall, this study found that it is possible to use an interface-based snow depth retrieval, such as the pulse peakiness algorithm, for wide band radar observations on land, particularly when applied to landscapes dominated by low stand vegetation and low relief, flat surface topography. Adapting the sea ice classification approach provided good results only after development of the method.

4.2 Methodology Limitations

Several limitations to the method are recognized from this study. First, both methods were operated under the same assumption that the strongest return peak was from the snow-ground interface. While several studies made the same assumption for snow on sea ice, the relatively more complicated ground condition can lead to a probability in some instances that this assumption might not always be valid for snow on land. In the study by Julita et al. (2021), the original pulse peakiness algorithm accounted for cases when the strongest return was not always the snow-ice interface but the top of a brine layer in the snow. In this study it was not clear why there were multiple peaks close to the strongest return unlike the situation generally found on sea ice. Possible explanations include suggestion that the dielectric contrast between snow and ground may not be as strong as the case for snow on sea ice, possibly buried vegetation in the snowpack being picked up by the radar, and signals reflected by the sides of the hummocks being added. Without

detailed field data to verify any of the above assumptions, it is difficult to fully explain the signal response. And as shown in the results section, waveforms that contradicted this base assumption were a major source of overestimation for both approaches.

Second, due to the large number of parameters involved in the pulse peakiness approach, the three main parameters were systematically optimized while the others were chosen after limiter testing. For example, in one of the added waveform filters where the air-snow interface peak logarithmic scale return has to be below 0.6, only three values were tested (0.5, 0.6, 0.7) and 0.6 seemed to have yielded the best balance between MAE and number of valid outputs. However, it is possible that other values could have provided a better result. In terms of parameters for the K-means approach, the number of pre-determined classes is another example where only limited testing was done. As too many classes would lead to clusters with no waveforms while too few classes would result in waveforms not similar to each other being clustered together, this could affect the final result.

Third, there are some inherent challenges in the wide band radar data itself, namely the requirement for flat surface and in some cases the presence of sidelobes (Newman et al., 2014; Kwok et al., 2011). Sidelobes specifically can be interpreted as air-snow interface peaks for shallower snowpacks as they are peaks close to the maximum return. Furthermore, Newman et al. (2014) demonstrated that even on sea ice, an ice surface topographic filter was needed to discard any waveforms in areas with significant topographic variability. While the 1m spatial resolution DEM used in this study would enable the use of a similar filter, the hummocks in the area are often smaller than a 1m x 1m cell. So a higher spatial resolution DEM would be needed. Also, the threshold for discarding waveform was not

straightforward to define due to the distribution of deep snow: deep snow in the study area were usually around areas with relatively steeper terrain and discarding them would severely change the distribution of snow depth. While Kwok et al. (2011) noted the existence of sidelobes, it was explained that in their dataset, the sidelobes appeared after the strongest return, which did not interfere with their interface picking. More specifically, Kwok et al. (2011) were able to identify the location and return strength of sidelobes and recommended that the sidelobes were reduced in future radar systems.

Finally, snow density across the entire study area was assumed to be $300\text{kg}/\text{m}^3$. It is expected that although bulk snow density likely was conservative in variability, at the local spatial scale and vertically through the pack, this was likely not the case for the entire study area. Analysis of snow density spatial variability could provide insight into waveform complexity.

4.3 Recommendations

Three recommendations are made to improve the approach taken in this study.

Recommendation 1: Improvement in identifying the snow-ground interface in the waveform from the strongest peak return. Kwok and Maksym (2014) used a peak signal to noise ratio threshold could to filter out complex waveforms like the ones shown in Figure 3.21. In addition, a fixed threshold could be applied that only allows (for example) one peak to have its normalized return be above a pre-determined threshold (0.9). Also, more detailed field data could assist in determining the reason why multiple strong peaks

appeared adjacent to one another in the waveform.

Recommendation 2: Explore alternative classification workflows. Current classification approach was developed to theoretically take advantage of both the pulse peakiness concept and classification’s ability to cluster similar shaped waveform, and can yield some accurate results in specific circumstances. However, the complexity of the waveforms proved challenging for K-means to correctly cluster them. A more sophisticated classification approach like a Convolutional Neural Network (CNN) if correctly configured and trained should be more capable of dealing with complex waveforms. While the introduction of the pulse peakiness left values did to some degrees overcome the problem of low separability between waveforms representing similar snow depths, and was able to output a specific snow depth, the two parameters used cannot always uniquely identify the correct air-snow interface peak. To further utilize the principles behind the pure pulse peakiness approach from Julita et al. (2021), a third parameter that record how far away the peaks are from the maximum return could be added as the third criterion to represent the ”choosing the first peak” element, and the Equation 3.4 would be changed to the following:

$$score = \omega_1 \times |log_{peak} - log_{target}| + \omega_2 \times |PP_{peak} - PP_{target}| - \omega_3 \times d \quad (4.1)$$

where d refers to the distance between a potential peak and the maximum return. In addition, the weights used Equation 3.4 and 4.1 should also be systematically optimized.

Recommendation 3:, Application of a wavelet interface-based approach. The wavelet approach and the GSFC-NK approach are different from the pulse peakiness approach used in this study. The pulse-peakiness approach was used because it is relatively straightfor-

ward and widely used, plus it has a demonstrated robust performance over sea ice (Julita et al., 2021). However, these alternative methods have a comparable performance and for the wavelet approach specifically, it does not need pre-defined threshold which could introduced uncertainties into the retrieval as Julita et al. (2021) pointed out. In addition for the GSFC-NK proposed by Kurtz et al. (2014), it is a physical approach instead of an empirical one, where many uncertainties in the waveforms like multiple close to maximum peaks could be explained by the physical model.

4.4 Conclusion

The overall objective of this thesis is to test the applicability of a standard interface-based snow depth retrieval algorithms for snow on sea ice to snow on land, and to explore the possibility of adapting the sea ice waveform classification approach for snow depth retrieval. In principle, the pulse peakiness approach from Julita et al. (2021) should be valid for applicability to snow on land because the snow-ground interface provides a strong dielectric contrast along with the snow-air interface; these contrasts should be somewhat similar to snow on sea ice. However, the complex ground characteristics narrowed the locations where the pulse peakiness method could yield accurate results. More specifically, the presence of tall surface vegetation combined with the hummocky terrain meant that a considerably large percentage of the waveforms were discarded, and to some degree refocused the effort into how to filter out complex waveforms that are too ambiguous for the method. With respect to the classification approach, the complex waveforms with significantly more peaks than the ones from radar altimeters seen in sea ice classification made

adapting the waveform characterizing parameters from sea ice classification considerably more challenging. Even though the final results for the classification approach were accurate for very specific situations, the inclusion of pulse peakiness and thresholds successfully transformed the method into a compromise between the two original approaches. In conclusion, for snow on land, if the ground is generally a flat surface with minimal vegetation, it is expected that the interface-based pulse peakiness algorithm will have a good performance (in this study, the MAE was found to be 0.11m for dwarf shrub, tussock and lichen combined). Development of this approach using a hybrid the pulse peakiness and k-means classification scheme demonstrated modest performance that, in its current state, requires further work to be comparable with the interface-based approach.

References

- Aldenhoff, W., Eriksson, L. E., & Heuzé, C. (2019). Comparison of sentinel-1 sar and sentinel-3 altimetry data for sea ice type discrimination. In *Igarss 2019 - 2019 ieee international geoscience and remote sensing symposium* (p. 4238-4240). doi: 10.1109/IGARSS.2019.8899041
- Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research. (2019). *Awis toolbox*. <https://github.com/awi-als-toolbox/awi-als-toolbox>. GitHub.
- Anders, K., Antonova, S., Wegener, A., Höfle, B., & Marsh, P. (2018). Airborne laser scanning (als) point clouds of trail valley creek, nwt, canada (2016) vulnerability of tundra lakes to climate change view project 4d near real-time environmental monitoring view project. Retrieved from <https://www.uni-heidelberg.de/permasar> doi: 10.1594/PANGAEA.894884
- Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005, 11). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438, 303-309. doi: 10.1038/nature04141
- Biemans, H., Siderius, C., Lutz, A. F., Nepal, S., Ahmad, B., Hassan, T., ... Immerzeel, W. W. (2019, 7). Importance of snow and glacier meltwater for agriculture on the indo-gangetic plain. *Nature Sustainability*, 2, 594-601. doi: 10.1038/s41893-019-0305-3
- Boike, J., & Grünberg, I. (2019, August). *Vegetation map of trail valley creek, northwest territories, canada*. doi: 10.1594/PANGAEA.904270
- Brooker, G. M. (2005). Understanding millimetre wave fmcw radars. In *1st international*

- conference on sensing technology* (Vol. 1).
- Chang, A., Foster, J., & Hall, D. (1987). Nimbus-7 smmr derived global snow cover parameters. *Annals of Glaciology*, 9, 39–44. doi: 10.3189/S0260305500200736
- Croce, P., Formichi, P., Landi, F., Mercogliano, P., Bucchignani, E., Dosio, A., & Di-mova, S. (2018, 1). The snow load in europe and the climate change. *Climate Risk Management*, 20, 138-154. doi: 10.1016/j.crm.2018.03.001
- Cui, Y., Xiong, C., Shi, J., Jiang, L., Peng, B., Ji, D., & Zhao, T. (2016). Estimating snow water equivalent with backscattering at x and ku bands. In *2016 ieee international geoscience and remote sensing symposium (igarss)* (p. 3454-3457). doi: 10.1109/IGARSS.2016.7729893
- Dahe, Q., Shiyin, L., & Peiji, L. (2006). Snow cover distribution, variability, and response to climate change in western china. *Journal of climate*, 19(9), 1820–1833. doi: <https://doi.org/10.1175/JCLI3694.1>
- Dettmering, D., Wynne, A., Müller, F. L., Passaro, M., & Seitz, F. (2018, 8). Lead detection in polar oceans-a comparison of different classification methods for cryosat-2 sar data. *Remote Sensing*, 10. doi: 10.3390/rs10081190
- EuroMotor virtual college. (n.d.). *Saw-tooth frequency modulation of fmcw-radar*. Retrieved from <https://www.euomotor.org/mod/resource/view.php?id=21809&forceview=1>
- European Space Agency. (n.d.). *Instrument payload*. Retrieved from <https://sentinel.esa.int/web/sentinel/missions/sentinel-1/instrument-payload>
- Farrell, S. L., Kurtz, N., Connor, L. N., Elder, B. C., Leuschen, C., Markus, T., ... Sonntag, J. G. (2012). A first assessment of icebridge snow and ice thickness data

- over arctic sea ice. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 2098-2111. doi: 10.1109/TGRS.2011.2170843
- Field, P. R., & Heymsfield, A. J. (2015). Importance of snow to global precipitation. *Geophysical Research Letters*, 42(21), 9512-9520. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GL065497> doi: 10.1002/2015GL065497
- Galín, N., Worby, A., Markus, T., Leuschen, C., & Gogineni, P. (2012, 1). Validation of airborne fmcw radar measurements of snow thickness over sea ice in antarctica. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 3-12. doi: 10.1109/TGRS.2011.2159121
- Geldsetzer, T., Langlois, A., & Yackel, J. (2009, 8). Dielectric properties of brine-wetted snow on first-year sea ice. *Cold Regions Science and Technology*, 58, 47-56. doi: 10.1016/j.coldregions.2009.03.009
- Griffiths, H. (1990). New ideas in fm radar. *Electronics & Communication Engineering Journal*, 2(5), 185-194.
- Grünberg, I., Wilcox, E. J., Zwieback, S., Marsh, P., & Boike, J. (2020, 8). Linking tundra vegetation, snow, soil temperature, and permafrost. *Biogeosciences*, 17, 4261-4279. doi: 10.5194/bg-17-4261-2020
- Hallikainen, M. (1977, January). *Dielectric properties of sea ice at microwave frequencies*. NASA STI/Recon Technical Report N.
- Hendricks, S., Ricker, R., & Juttila, A. (2019). *Icebird 2019 winter icesat-2 validation data acquisition report*.
- Juttila, A., King, J., Paden, J., Ricker, R., Hendricks, S., Polashenski, C., ... Haas, C.

- (2021). High-resolution snow depth on arctic sea ice from low-altitude airborne microwave radar data. *IEEE Transactions on Geoscience and Remote Sensing*, 60. doi: 10.1109/TGRS.2021.3063756
- King, J., Brady, M., & Newman, T. (2020, October). *kingjml/pysnowradar: Updated iee tgrs submission*. Zenodo. Retrieved from <https://doi.org/10.5281/zenodo.4071947> doi: 10.5281/zenodo.4071947
- King, J., Toose, P., Silis, A., & Derksen, C. (2020, September). *TVCSnow 2018-2019 tundra snow depth probe measurements*. Zenodo. Retrieved from <https://doi.org/10.5281/zenodo.4021401>
- Koenig, L. S., Ivanoff, A., Alexander, P. M., MacGregor, J. A., Fettweis, X., Panzer, B., ... Gogineni, P. (2016, 8). Annual greenland accumulation rates (2009-2012) from airborne snow radar. *Cryosphere*, 10, 1739-1752. doi: 10.5194/tc-10-1739-2016
- Kurtz, N. T., & Farrell, S. L. (2011). Large-scale surveys of snow depth on arctic sea ice from operation icebridge. *Geophysical Research Letters*, 38(20). Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011GL049216> doi: <https://doi.org/10.1029/2011GL049216>
- Kurtz, N. T., Farrell, S. L., Studinger, M., Galin, N., Harbeck, J. P., Lindsay, R., ... Sonntag, J. G. (2013). Sea ice thickness, freeboard, and snow depth products from operation icebridge airborne data. *Cryosphere*, 7, 1035-1056. doi: 10.5194/tc-7-1035-2013
- Kurtz, N. T., Galin, N., & Studinger, M. (2014, 7). An improved cryosat-2 sea ice freeboard retrieval algorithm through the use of waveform fitting. *Cryosphere*, 8, 1217-1237. doi: 10.5194/tc-8-1217-2014

- Kwok, R., Kurtz, N. T., Brucker, L., Ivanoff, A., Newman, T., Farrell, S. L., ... Tschudi, M. (2017, 11). Intercomparison of snow depth retrievals over arctic sea ice from radar data acquired by operation icebridge. *Cryosphere*, 11, 2571-2593. doi: 10.5194/tc-11-2571-2017
- Leinss, S., Wiesmann, A., Lemmetyinen, J., & Hajnsek, I. (2015, 8). Snow water equivalent of dry snow measured by differential interferometry. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 3773-3790. doi: 10.1109/JSTARS.2015.2432031
- Lievens, H., Demuzere, M., Marshall, H. P., Reichle, R. H., Brucker, L., Brangers, I., ... Lannoy, G. J. D. (2019, 12). Snow depth variability in the northern hemisphere mountains observed from space. *Nature Communications*, 10. doi: 10.1038/s41467-019-12566-y
- Marshall, H. P., & Koh, G. (2008, 4). Fmcw radars for snow research. *Cold Regions Science and Technology*, 52, 118-131. doi: 10.1016/j.coldregions.2007.04.008
- Marshall, H.-P., Koh, G., & Forster, R. R. (2004). Ground-based frequency-modulated continuous wave radar measurements in wet and dry snowpacks, colorado, usa: an analysis and summary of the 2002-03 nasa clpx data. *Hydrological Processes*, 18(18), 3609-3622. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.5804> doi: <https://doi.org/10.1002/hyp.5804>
- Marshall, H.-P., Koh, G., & Forster, R. R. (2005). Estimating alpine snowpack properties using fmcw radar. *Annals of Glaciology*, 40, 157-162.
- Marshall, H. P., Schneebeli, M., & Koh, G. (2007, 1). Snow stratigraphy measurements with high-frequency fmcw radar: Comparison with snow micro-penetrometer. *Cold*

- Regions Science and Technology*, 47, 108-117. doi: 10.1016/j.coldregions.2006.08.008
- Matzler, C., Aebischer, H., & Schanda, E. (1984). Microwave dielectric properties of surface snow. *IEEE Journal of Oceanic Engineering*, 9(5), 366-371. doi: 10.1109/JOE.1984.1145644
- Mostafa, A. (2019). *Fmcw radar system*. Retrieved from <https://uwaterloo.ca/centre-for-intelligent-antenna-and-radio-systems/sites/ca.centre-for-intelligent-antenna-and-radio-systems/files/uploads/files/fmcwradarsystem.pdf>
- Müller, F. L., Dettmering, D., Bosch, W., & Seitz, F. (2017, 6). Monitoring the arctic seas: How satellite altimetry can be used to detect openwater in sea-ice regions. *Remote Sensing*, 9. doi: 10.3390/rs9060551
- National Aeronautics and Space Administration. (n.d.). *Icebridge past campaigns*. Retrieved from https://www.nasa.gov/mission_pages/icebridge/news/past.html
- Oveisgharan, S., Esteban-Fernandez, D., Waliser, D., Friedl, R., Nghiem, S., & Zeng, X. (2020, 6). Evaluating the preconditions of two remote sensing swe retrieval algorithms over the us. *Remote Sensing*, 12. doi: 10.3390/rs12122021
- Palmer, M. J., Burn, C. R., & Kokelj, S. V. (2012, 8). Factors influencing permafrost temperatures across tree line in the uplands east of the mackenzie delta, 2004-2010 1,2. *Canadian Journal of Earth Sciences*, 49, 877-894. doi: 10.1139/E2012-002
- Panzer, B., Gomez-Garcia, D., Leuschen, C., Paden, J., Rodriguez-Morales, F., Patel, A., ... Gogineni, P. (2013, 4). An ultra-wideband, microwave radar for measuring snow

- thickness on sea ice and mapping near-surface internal layers in polar firn. *Journal of Glaciology*, 59, 244-254. doi: 10.3189/2013JoG12J128
- Pohl, S., Davison, B., Marsh, P., & Pietroniro, A. (2005). Modelling spatially distributed snowmelt and meltwater runoff in a small arctic catchment with a hydrology land-surface scheme (watchclass). *Atmosphere - Ocean*, 43, 193-211. doi: 10.3137/ao.430301
- Pomeroy, J. W., Gray, D. M., Shook, K. R., Toth, B., Essery, R. L. H., Pietroniro, A., & Hedstrom, N. (1998). An evaluation of snow accumulation and ablation processes for land surface modelling. *Hydrological Processes*, 12(15), 2339-2367. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291099-1085%28199812%2912%3A15%3C2339%3A%3AAID-HYP800%3E3.0.CO%3B2-L> doi: [https://doi.org/10.1002/\(SICI\)1099-1085\(199812\)12:15<2339::AID-HYP800>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1099-1085(199812)12:15<2339::AID-HYP800>3.0.CO;2-L)
- Ricker, R., Hendricks, S., Helm, V., & Gerdes, R. (2015). Classification of cryosat-2 radar echoes. In *Towards an interdisciplinary approach in earth system science* (pp. 149-158). Springer. doi: https://doi.org/10.1007/978-3-319-13865-7_17
- Rott, H., Yueh, S. H., Cline, D. W., Duguay, C., Essery, R., Haas, C., ... Thompson, A. (2010). Cold regions hydrology high-resolution observatory for snow and cold land processes. *Proceedings of the IEEE*, 98, 752-765. doi: 10.1109/JPROC.2009.2038947
- Shen, X., Zhang, J., Zhang, X., Meng, J., & Ke, C. (2017, 11). Sea ice classification using cryosat-2 altimeter data by optimal classifier-feature assembly. *IEEE Geoscience and Remote Sensing Letters*, 14, 1948-1952. doi: 10.1109/LGRS.2017.2743339
- Skolnik, M. I. (2008). *Radar handbook*. McGraw-Hill Education.
- SnowHydro. (n.d.). *Gps snow depth probe*. Retrieved from <http://www.snowhydro.com/>

[products/column2.html](#)

- Stelzer, A., Jahn, M., & Scheiblhofer, S. (2008). Precise distance measurement with cooperative fmcw radar units. In (p. 771-774). doi: 10.1109/RWS.2008.4463606
- Stove, A. G. (1992). Linear fmcw radar techniques. In *Iee proceedings f-radar and signal processing* (Vol. 139, pp. 343-350). doi: 10.1049/ip-f-2.1992.0048
- Sturm, M., Goldstein, M. A., & Parr, C. (2017). Water and life from snow: A trillion dollar science question. *Water Resources Research*, 53(5), 3534-3544. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017WR020840> doi: <https://doi.org/10.1002/2017WR020840>
- Sturm, M., Taras, B., Liston, G. E., Derksen, C., Jonas, T., & Lea, J. (2010, 12). Estimating snow water equivalent using snow depth data and climate classes. *Journal of Hydrometeorology*, 11, 1380-1394. doi: 10.1175/2010JHM1202.1
- Trail Valley Creek Arctic Research Station. (n.d.). *About us*. Retrieved from <https://www.trailvalleycreek.ca/about>
- Ulaby, F. T., Stiles, W. H., & Abdelrazik, M. (1984). Snowcover influence on backscattering from terrain. *IEEE Transactions on Geoscience and Remote Sensing*, GE-22(2), 126-133. doi: 10.1109/TGRS.1984.350604
- Ulaby, F. T., Stiles, W. H., Dellwig, L. F., & Hanson, B. C. (1977). Experiments on the radar backscatter of snow. *IEEE Transactions on Geoscience Electronics*, 15, 185-189. doi: 10.1109/TGE.1977.294490
- Verret, M., Wang, Y., Bjornson, J., & Lacelle, D. (2019). Hummocks in alpine tundra, northern british columbia, canada: Distribution, morphology and organic carbon composition. *Arctic Science*, 5, 127-147. doi: 10.1139/as-2018-0021

- Wadhams, P., Tucker III, W. B., Krabill, W. B., Swift, R. N., Comiso, J. C., & Davis, N. R. (1992). Relationship between sea ice freeboard and draft in the arctic basin, and implications for ice thickness monitoring. *Journal of Geophysical Research: Oceans*, 97(C12), 20325-20334. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/92JC02014> doi: <https://doi.org/10.1029/92JC02014>
- Wingham, D. J., Francis, C. R., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., ... Wallis, D. W. (2006). Cryosat: A mission to determine the fluctuations in earth's land and marine ice fields. *Advances in Space Research*, 37, 841-871. doi: 10.1016/j.asr.2005.07.027
- Wolff, C. (n.d.). *Frequency-modulated continuous-wave radar*. Retrieved from <https://www.radartutorial.eu/02.basics/Frequency%20Modulated%20Continuous%20Wave%20Radar.en.html>
- Yan, J.-B., Gomez-García Alvestegui, D., McDaniel, J. W., Li, Y., Gogineni, S., Rodriguez-Morales, F., ... Leuschen, C. J. (2017). Ultrawideband fmcw radar for airborne measurements of snow over sea ice and land. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2), 834-843. doi: 10.1109/TGRS.2016.2616134
- Yueh, S. H., Dinardo, S. J., Akgiray, A., West, R., Cline, D. W., & Elder, K. (2009). Airborne ku-band polarimetric radar remote sensing of terrestrial snow cover. *IEEE Transactions on Geoscience and Remote Sensing*, 47(10), 3347-3364. doi: 10.1109/TGRS.2009.2022945
- Zygmuntowska, M., Khvorostovsky, K., Helm, V., & Sandven, S. (2013). Waveform classification of airborne synthetic aperture radar altimeter over arctic sea ice. *Cryosphere*, 7, 1315-1324. doi: 10.5194/tc-7-1315-2013

Appendix A

Full Optimization Results for the Pulse Peakiness Approach

Table A.1: Pulse peakiness parameters optimization results for trees

Tree (100-0)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.6	0.5	25	0.12	0.65	11 / 84
Tree (80-20)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.7	0.7	15	0.22	0.47	8 / 17
Run #2	0.6	0.5	25	0.14	N/A	2 / 17
Run #3	0.6	0.5	25	0.05	N/A	2 / 17
Run #4	0.5	0.5	25	0.30	N/A	2 / 17
Run #5	0.5	0.5	25	0.28	0.20	4 / 17
Tree (50-50)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.3	0.7	25	0.28	-0.11	11 / 42
Run #2	0.6	0.8	15	0.28	0.16	27 / 42
Run #3	0.8	0.7	10	0.30	0.21	13 / 42
Run #4	0.7	0.7	15	0.22	0.35	20 / 42
Run #5	0.2	0.8	15	0.38	0.07	36 / 42

*indicates significant P-value at $\alpha < .01$

Table A.2: Pulse peakiness parameters optimization results for tall shrubs

Tall Shrub (100-0)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.3	0.8	20	0.27	0.20*	226 / 441
Tall Shrub (80-20)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.3	0.8	20	0.2443	0.28	47 / 89
Run #2	0.1	0.8	20	0.3515	0.01	66 / 89
Run #3	0.3	0.8	20	0.2494	0.27	46 / 89
Run #4	0.2	0.8	20	0.2728	0.14	62 / 89
Run #5	0.1	0.8	20	0.3578	-0.36*	64 / 89
Tall Shrub (50-50)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.3	0.8	20	0.2761	0.13	114 / 221
Run #2	0.3	0.8	20	0.2727	0.2	110 / 221
Run #3	0.1	0.9	15	0.4293	-0.07	211 / 221
Run #4	0.2	0.8	20	0.3314	0.05	119 / 221
Run #5	0.1	0.7	20	0.3298	-0.06	154 / 221

*indicates significant P-value at $\alpha < .01$

Table A.3: Pulse peakiness parameters optimization results for riparian shrubs

Riparian Shrub (100-0)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.2	0.7	15	0.29	0.35*	209 / 321
Riparian Shrub (80-20)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.2	0.6	15	0.301	0.27	37 / 65
Run #2	0.2	0.8	15	0.3287	0.32	48 / 65
Run #3	0.7	0.4	20	0.2008	0.6	7 / 65
Run #4	0.6	0.4	20	0.2581	-0.01	10 / 65
Run #5	0.2	0.7	15	0.2654	0.54*	40 / 65
Riparian Shrub (50-50)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.2	0.7	15	0.2958	0.35*	107 / 161
Run #2	0.2	0.4	15	0.2927	0.28	71 / 161
Run #3	0.2	0.5	15	0.3037	0.35*	86 / 161
Run #4	0.1	0.8	20	0.2978	0.22	113 / 161
Run #5	0.2	0.7	15	0.3153	0.35*	110 / 161

*indicates significant P-value at $\alpha < .01$

Table A.4: Pulse peakiness parameters optimization results for dwarf shrubs

Dwarf Shrub (100-0)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.4	0.5	25	0.13	0.29*	135 / 500
Dwarf Shrub (80-20)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.4	0.5	25	0.1308	-0.07	28 / 100
Run #2	0.4	0.5	25	0.1176	0.3	26 / 100
Run #3	0.4	0.5	25	0.0921	0.41	25 / 100
Run #4	0.3	0.5	25	0.1876	0.24	24 / 100
Run #5	0.4	0.5	25	0.1336	0.45	24 / 100
Dwarf Shrub (50-50)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.4	0.5	25	0.1351	0.23	65 / 250
Run #2	0.4	0.5	25	0.129	0.37*	73 / 250
Run #3	0.5	0.5	25	0.1219	0.27	57 / 250
Run #4	0.3	0.5	25	0.1388	0.36*	75 / 250
Run #5	0.4	0.5	25	0.1226	0.36*	64 / 250

*indicates significant P-value at $\alpha < .01$

Table A.5: Pulse peakiness parameters optimization results for tussocks

Tussock (100-0)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.5	0.5	30	0.09	0.55*	141 / 1224
Tussock (80-20)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.3	0.6	25	0.14	0.37*	75 / 245
Run #2	0.3	0.6	25	0.17	0.05	75 / 245
Run #3	0.3	0.6	25	0.15	0.30*	75 / 245
Run #4	0.3	0.6	25	0.15	-0.02	61 / 245
Run #5	0.3	0.6	25	0.13	0.37*	70 / 245
Tussock (50-50)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.4	0.8	25	0.14	0.34*	190 / 612
Run #2	0.3	0.6	25	0.13	0.33*	200 / 612
Run #3	0.4	0.7	25	0.14	0.43*	184 / 612
Run #4	0.4	0.6	25	0.13	0.39*	163 / 612
Run #5	0.3	0.6	25	0.13	0.36*	205 / 612

*indicates significant P-value at $\alpha < .01$

Table A.6: Pulse peakiness parameters optimization results for Lichens

Lichen (100-0)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.3	0.5	30	0.10	0.46*	74 / 558
Lichen (80-20)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.3	0.5	30	0.10	0.65	13 / 112
Run #2	0.3	0.5	30	0.13	0.41	14 / 112
Run #3	0.3	0.5	25	0.22	0.09	31 / 112
Run #4	0.3	0.5	30	0.08	0.65*	17 / 112
Run #5	0.3	0.5	30	0.07	0.73*	14 / 112
Lichen (50-50)						
	Log	Lin	PP_L	MAE(m)	Pearson	Valid / Total
Run #1	0.3	0.7	25	0.17	0.14	118 / 279
Run #2	0.5	0.5	25	0.20	0.30*	71 / 279
Run #3	0.3	0.5	25	0.15	0.19	79 / 279
Run #4	0.3	0.5	25	0.15	0.31*	80 / 279
Run #5	0.3	0.7	20	0.23	0.22*	171 / 279

*indicates significant P-value at $\alpha < .01$

Appendix B

Full Results for the K-means Approach

Table B.1: K-means approach results for trees

	Tree		
	MAE(m)	Pearson	Valid / Total
Run#1	0.24	0.35	31 / 42
Run#2	0.32	0.07	29 / 42
Run#3	0.30	0.06	28 / 42
Run#4	0.21	0.43	29 / 42
Run#5	0.24	0.28	29 / 42

*indicates significant P-value at $\alpha < .01$

Table B.2: K-means approach results for tall shrubs

Tall Shrub			
	MAE(m)	Pearson	Valid / Total
Run#1	0.23	0.24*	151 / 221
Run#2	0.24	0.32*	150 / 221
Run#3	0.28	0.13	149 / 221
Run#4	0.27	0.12	144 / 221
Run#5	0.25	0.25*	140 / 221

*indicates significant P-value at $\alpha < .01$

Table B.3: K-means approach results for riparian shrubs

Riparian Shrub			
	MAE(m)	Pearson	Valid / Total
Run#1	0.23	0.45*	63 / 161
Run#2	0.24	0.21	68 / 161
Run#3	0.25	0.26	65 / 161
Run#4	0.23	0.28	65 / 161
Run#5	0.24	0.32*	63 / 161

*indicates significant P-value at $\alpha < .01$

Table B.4: K-means approach results for dwarf shrubs

Dwarf Shrub			
	MAE(m)	Pearson	Valid / Total
Run#1	0.19	0.11	158 / 250
Run#2	0.19	0.13	174 / 250
Run#3	0.19	0.10	186 / 250
Run#4	0.16	0.14	182 / 250
Run#5	0.19	0.23*	167 / 250

*indicates significant P-value at $\alpha < .01$

Table B.5: K-means approach results for tussocks

Tussock			
	MAE(m)	Pearson	Valid / Total
Run#1	0.16	0.26*	316 / 612
Run#2	0.16	0.17*	319 / 612
Run#3	0.16	0.22*	313 / 612
Run#4	0.15	0.26*	320 / 612
Run#5	0.17	0.22*	312 / 612

*indicates significant P-value at $\alpha < .01$

Table B.6: K-means approach results for lichens

Lichen			
	MAE(m)	Pearson	Valid / Total
Run#1	0.2	0.17	191 / 279
Run#2	0.2	0.18*	180 / 279
Run#3	0.2	0.18*	187 / 279
Run#4	0.18	0.23*	185 / 279
Run#5	0.21	0.14	190 / 279

*indicates significant P-value at $\alpha < .01$