Navigating the Snowscape: Utilizing SAR Interferometric Coherence and ICESat-2 Data for Machine Learning-Based Snow Depth Estimation

MANAS KABADWAL

August 2024

SUPERVISORS:

Dr. Hossein Aghababaei (First Supervisor)

Dr. Ling Chang (Second Supervisor)

Navigating the Snowscape: Utilizing SAR Interferometric Coherence and ICESat-2 Data for Machine Learning-Based Snow Depth Estimation

MANAS KABADWAL

Enschede, The Netherlands, August 2024

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Geoinformatics

SUPERVISORS:

Dr. Hossein Aghababaei (First Supervisor)

Dr. Ling Chang (Second Supervisor)

THESIS ASSESSMENT BOARD:

Prof.Dr.ir. Claudio Persello (Chair)

Andrea Buono (External Examiner)

ABSTRACT

Accurate snow depth estimation is crucial for understanding hydrological cycles, predicting water resources, and assessing the impacts of climate change on snow-covered regions. Snow depth data provides vital insights for climate modeling, water management, and disaster risk reduction, especially in regions dependent on snowmelt for freshwater resources. Enhancing the precision of snow depth measurement techniques is, therefore, essential for both scientific research and practical applications.

This thesis delves into the utilization of ICESat-2 LIDAR data as a reference for training machine learning models to predict snow depth using SAR data from SAOCOM 1B and Sentinel-1. A meticulously crafted methodology was employed, encompassing advanced preprocessing, coregistration techniques, absolute coherence calculation, and sophisticated machine learning-based prediction models. The study's comprehensive approach revealed that SAOCOM 1B's L-band SAR data significantly surpasses Sentinel-1's C-band SAR data in terms of snow depth prediction accuracy, primarily due to its superior penetration depth and more dependable coherence measurements.

The Random Forest model was identified as the most effective machine learning algorithm, achieving an exceptional R-squared value of 0.8212 with the gt1r beam from ICESat-2, thereby underscoring the potential use of LIDAR-SAR data for snow-monitoring model generation. The feature importance analysis underscored the pivotal role of multi-polarization data (HH, HV, VH, VV) in augmenting the prediction accuracy, emphasizing the necessity of utilizing diverse polarization channels to capture the comprehensive range of snowpack characteristics.

Despite the comparatively lower performance of Sentinel-1 data, this study highlighted potential avenues for enhancement in C-band SAR technology and data processing techniques. The findings elucidate the critical importance of selecting appropriate SAR datasets and leveraging robust machine learning models for accurate snow depth estimation. This research provides invaluable insights into the comparative strengths and limitations of different SAR datasets and different beams of ICESat-2, significantly contributing to the advancement of snow monitoring practices and a deeper understanding of snowpack dynamics within the broader context of climate change.

Keywords: ICESat-2, SAOCOM 1B, Sentinel-1, Snow Depth Estimation, SAR data, LIDAR, Machine Learning, Random Forest, Absolute Interferometric Coherence, Multi-polarization.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my supervisors, Dr. Hossein Aghababaei and Dr. Ling Chang. I am especially thankful to Dr. Hossein Aghababaei, my primary supervisor, for his unwavering support, guidance, and encouragement throughout the course of my research. His insightful feedback and constructive criticism have been invaluable in shaping this thesis. I am truly grateful for his expertise and the time he dedicated to helping me achieve my academic goals.

I would also like to extend my heartfelt thanks to Dr. Ling Chang for her valuable contributions and support as my co-supervisor. Her expertise and thoughtful advice have greatly enriched my research.

I am deeply grateful to my family for their endless love, support, and patience. To my parents, thank you for believing in me and for your constant encouragement. Your sacrifices and unwavering faith have been the driving force behind my achievements. To my brother, Divyanshu, thank you for always being there for me and for your words of encouragement during challenging times.

Additionally, I am thankful to my friends and colleagues who have provided me with moral support and companionship throughout this journey. Your understanding and camaraderie have made this experience enjoyable and memorable.

Finally, I would like to acknowledge the support from the University of Twente, particularly the Faculty of Geoinformation Science and Earth Observation for providing the resources and facilities necessary for conducting this research. The academic environment and the access to state-of-the-art resources at the University have been instrumental in the completion of this thesis. I am grateful for the opportunities and support provided by the university.

Thank you all for your invaluable contributions and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background

Measuring snow depth is essential for comprehending global hydrological cycles, climate patterns, and ecological stability. Accurate snow depth data is vital for managing water resources, predicting avalanche risks, and studying climate change impacts (Enderlin et al., 2022). Traditional methods for measuring snow depth, such as ground-based observations and airborne LiDAR, provide high accuracy but are limited in spatial coverage and temporal frequency. Satellite-based remote sensing technologies, particularly Synthetic Aperture Radar (SAR) and Light Detection and Ranging (LiDAR), offer a promising solution for extensive and frequent snow depth monitoring.

Optical remote sensing utilizes satellite sensors to provide high-resolution imagery for snow cover mapping by relying on visible and near-infrared wavelengths to distinguish snow from other surface features. However, optical remote sensing is limited by cloud cover and the inability to penetrate the snowpack, which restricts its use in cloudy conditions and only provides surface snow information (Tsang et al., 2022).

Airborne and spaceborne LiDAR systems utilize laser pulses to calculate the distance from the sensor to the Earth's surface, enabling highly accurate elevation data collection. Airborne LiDAR can provide highly accurate snow depth data over specific areas, but it is expensive and logistically challenging to cover large regions frequently (Lu et al., 2022). Spaceborne LiDAR, like ICESat-2, offers global coverage but with limitations in spatial and temporal resolution due to its narrow swath width and sparse ground tracks (Kwok et al., 2020.). Equipped with the Advanced Topographic Laser Altimeter System (ATLAS), the ICESat-2 satellite offers highly accurate surface elevation measurements and is widely employed in cryospheric studies. The ATL03 product from ICESat-2 delivers photon heights, while the ATL08 product provides terrain and canopy heights, facilitating the extraction of snow depth by comparing surface and ground elevations (Zhao et al., 2022). However, ICESat-2's spatial coverage is limited due to its narrow swath width and sparse ground tracks.

SAR sensors, such as those on SAOCOM 1B (L-band) and Sentinel-1 (C-band), emit microwaves and measure the backscattered signals, allowing for snow depth estimation under various weather conditions and independent of daylight. L-band SAR, with its longer wavelength, penetrates deeper into the snowpack, providing valuable information on snow volume and structure (Benedikter et al., 2022). In contrast, C-band SAR offers higher resolution surface scattering information, useful for capturing surface features and shallow snow depths (Patil et al., 2020). However, SAR data requires complex processing and interpretation. Terrestrial methods, such as ground-based measurements with snow stakes, ultrasonic sensors, and manual probes, provide direct and accurate snow depth data. However, these methods are labor-intensive and limited in spatial coverage, making them impractical for large-scale monitoring (Bernard, 2017).

The limitations of each method highlight the need for an approach that leverages the strengths of multiple remote sensing technologies to achieve comprehensive and accurate snow depth estimation. This thesis focuses on using ICESat-2 LiDAR data as a reference for training machine learning models to predict snow depth using SAR data from SAOCOM 1B and Sentinel-1. By employing ICESat-2 LiDAR data, the gaps in spatial coverage and temporal frequency inherent in each individual dataset can be addressed.

The combined use of SAR and LiDAR data offers a robust framework for snow depth estimation across diverse terrains and weather conditions. This method enhances the precision of snow depth measurements and offers a scalable solution for monitoring extensive areas. The motivation behind this research is to enhance our understanding of snow dynamics, support water resource management, and contribute to climate change studies.

1.2. Literature Review

Recent research has extensively investigated the use of Synthetic Aperture Radar (SAR) and Light Detection and Ranging (LiDAR) data for snow depth estimation due to their complementary characteristics. SAR's ability to penetrate snow and provide detailed information on surface roughness and structure, combined with the high-precision elevation data from LiDAR, enhances the accuracy and reliability of snow depth measurements. However, for my research, it is important to clarify that I will not integrate SAR and LiDAR directly. Instead, I will use LiDAR data as a reference to train machine learning models, which will then predict snow depth using SAR data.

Leinss et al., (2014) utilized polarimetric SAR data from the TerraSAR-X satellite to determine snow height by analyzing the phase differences between the co-polarized channels. Their study revealed a strong correlation between the Co-Polar Phase Difference (CPD) of VV and HH polarizations and snow depth, suggesting a viable approach for snow depth retrieval across various snow conditions. However, their approach faced significant limitations related to the phase unwrapping process and the influence of surface roughness on phase measurements. These limitations can lead to inaccuracies in snow depth estimation, especially in areas with complex terrain. This highlights the need for methodologies that can account for these limitations and improve the robustness of snow depth retrievals.

Further emphasizing the complementary nature of SAR and LiDAR, Kwok et al. (2020) investigated the use of ICESat-2 and CryoSat-2 data for estimating snow depth and sea ice thickness in the Arctic. Their research illustrated the synergy between LiDAR and radar altimetry, showing that combining these datasets can enhance the spatial and temporal resolution of snow depth measurements. However, the study also identified challenges related to the differing spatial resolutions and temporal coverages of the datasets, which can introduce errors in snow depth estimation and limit the method's applicability in regions with rapidly changing snow conditions. These findings emphasize the need for developing techniques that effectively integrate diverse datasets to improve the accuracy of snow depth estimation.

Awasthi et al., (2021) applied pursuit-monostatic TanDEM-X datasets for snow depth retrieval in the North-Western Himalayan region. Using polarimetric SAR interferometry (PolInSAR), the study retrieved snow depth by exploiting the phase information of SAR signals. This technique proved effective in mountainous regions, where traditional methods are challenging. However, the study faced challenges related to the decorrelation of SAR signals in heavily forested areas and steep terrains. Additionally, the accuracy of snow depth estimates was affected by the temporal gap between SAR acquisitions. These limitations suggest the need for continuous SAR acquisitions and improved decorrelation techniques to enhance the reliability of snow depth estimates in complex terrains.

Tebaldini et al., (2016) took a different approach by using L-band airborne SAR tomography (TomoSAR) to image the internal structure of an alpine glacier. The study provided detailed 3-D maps of glacier subsurface structures, crucial for understanding glacier dynamics and mass balance. Although promising, this method requires multiple SAR acquisitions from different angles, which can be logistically challenging and costly. The high computational demand of TomoSAR also limits its practical application for large-scale

snow depth estimation. Consequently, creating cost-effective and computationally efficient methods for large-scale snow depth estimation continues to be a crucial research focus.

Building on the need for deeper penetration capabilities, Benedikter et al., (2022) highlighted the advantages of L-band SAR from the ALOS-2 satellite for snow depth estimation. They emphasized its deeper penetration capability compared to C-band SAR from the Sentinel-1 satellite. Their findings indicated that L-band SAR provides more accurate snow volume and structure information, essential for regions with thick snow cover. However, the study faced limitations related to the influence of underlying terrain and vegetation on SAR signal penetration, which can introduce errors in snow depth estimates. This underscores the necessity of developing techniques to mitigate these influences and improve snow depth estimation accuracy.

(Majumdar et al., 2019) further explored the application of polarimetric SAR data from the Radarsat-2 satellite for snow depth estimation, underscoring the need for multi-frequency and multi-polarization approaches to capture the complex interactions between radar signals and snow. Their research demonstrated that combining data from different SAR frequencies (e.g., C-band and L-band) with LiDAR improves the robustness of snow depth estimates, particularly in heterogeneous snow conditions. However, the study highlighted challenges related to the calibration and validation of SAR data, which can impact the accuracy of snow depth retrievals. This calls for standardized calibration and validation protocols to enhance the reliability of snow depth estimates.

Tsang et al., (2022) used polarimetric SAR data from the Sentinel-1 satellite to estimate snow depth and SWE in alpine regions, applying machine learning algorithms to integrate SAR data with ground-based measurements. This study demonstrated the potential of SAR data to capture the spatial variability of snow cover in complex terrains. However, the accuracy of the machine learning models was constrained by the limited availability of ground truth data for training and validation, underscoring the necessity of extensive field measurements. This highlights the importance of integrating machine learning with comprehensive ground truth data to enhance the accuracy of snow depth estimation.

Recent studies have also explored the potential of using coherence information derived from SAR interferometry for snow depth estimation. For instance, (Kumar & Venkataraman, 2011) utilized coherence data from SAR to map snow depth variations in alpine regions. This approach leverages the phase stability of SAR signals to detect changes in snow cover, offering a high-resolution method for monitoring snow dynamics. However, coherence-based methods can be affected by temporal decorrelation and changes in surface roughness, which can introduce errors in snow depth estimates. This highlights the need for robust coherence-based methodologies that can account for these factors.

Machine learning models have shown significant potential in remote sensing applications, including snow depth estimation. These models are capable of managing large datasets and capturing complex, non-linear relationships between input features (such as SAR coherence values) and snow depth. Random Forest (RF) and Gradient Boosting (GB) models are ensemble learning techniques that combine multiple decision trees to improve predictive performance. (Immerzeel et al., 2014) demonstrated the effectiveness of Random Forest in integrating multi-source remote sensing data for environmental monitoring. These models are robust to overfitting and can handle high-dimensional data, making them suitable for snow depth estimation using SAR and LiDAR data. However, the effectiveness of these models is highly dependent on the quality and quantity of the training data, and they can be vulnerable to noise and outliers in the input data. Support Vector Regression (SVR) is another robust machine learning technique used for regression tasks.

SVR is effective in high-dimensional spaces and performs well with limited data points, making it suitable for remote sensing applications where data availability may be limited. Studies have shown that SVR can achieve high accuracy in predicting snow depth by capturing the complex interactions between radar signals and snow properties (Li et al., 2021). However, SVR can be computationally demanding and often requires meticulous tuning of hyperparameters to achieve optimal performance.

Recent advancements in deep learning have also been explored for snow depth estimation. Convolutional neural networks (CNNs) are being implemented to SAR data for snow depth prediction, demonstrating the potential of deep learning in capturing spatial patterns and improving estimation accuracy. However, the computational complexity and requirement for large training datasets pose challenges for the widespread application of deep learning methods in snow depth estimation. Additionally, deep learning models can be prone to overfitting, particularly when trained on limited datasets.

A study by (Frey et al., 2015) utilized a combination of SAR and LiDAR data to improve snow depth estimation in forested areas. Their approach involved using LiDAR to map the forest canopy and SAR to detect the snow surface beneath the canopy. This integration provided a more accurate assessment of snow depth in complex environments, highlighting the potential of combining different remote sensing technologies for comprehensive snow monitoring. However, the study faced challenges related to the calibration of SAR and LiDAR data and the influence of forest canopy on SAR signal penetration.

1.3. Research Gap & Problem Statement

Despite significant advancements in utilizing SAR and LiDAR data for snow depth estimation, several critical gaps remain unaddressed. Current literature lacks comprehensive research on the optimal methodologies for feeding SAR coherence information into machine learning models for snow depth prediction. Existing studies often fail to systematically evaluate the impact of different SAR data preprocessing techniques and machine learning model configurations on the accuracy of snow depth estimates. Moreover, while various machine learning models have been employed, there is insufficient exploration of hybrid models and ensemble learning techniques that could leverage the strengths of different algorithms to enhance predictive performance.

Another significant gap is the limited comparative analysis of SAR datasets from different satellites, such as SAOCOM 1B and Sentinel-1, in conjunction with ICESat-2 LiDAR data. There is a need for detailed evaluations of how different SAR polarizations and frequencies influence snow depth predictions. Additionally, existing studies often focus on specific geographic areas, limiting the generalizability of their findings. There is a pressing need for research that validates snow depth estimation methodologies across diverse snow-covered regions to ensure robustness and applicability in different climatic and topographic conditions.

Furthermore, ICESat-2, despite providing high-precision elevation data, suffers from limited spatial coverage due to its narrow swath width and sparse ground tracks. This research seeks to explore the incorporation of SAR data to fill these information gaps, leveraging the extensive spatial coverage of SAR to derive continuous snow depth estimates across larger areas. Accurately estimating snow depth across diverse terrains and varying snow conditions remains a significant challenge in remote sensing. Traditional ground-based methods, although accurate, are limited by their spatial coverage and accessibility. Remote sensing technologies like LiDAR and SAR offer the potential for extensive and continuous monitoring but have their own limitations. LiDAR provides high-resolution elevation data but is constrained by its limited spatial coverage, whereas SAR offers extensive spatial coverage but faces challenges related to coherence

loss and noise. This research aims to tackle these challenges by developing a robust methodology that employs ICESat-2 LiDAR data as a reference for training machine learning models with SAR data from SAOCOM 1B and Sentinel-1 to estimate snow depth. By harnessing the unique strengths of both LiDAR and SAR, this study aims to improve the accuracy of snow depth predictions, assess the performance of various SAR datasets and polarizations, and identify the most reliable approach for snow depth estimation. The findings of this research will enhance the understanding and monitoring of snow dynamics, which is essential for managing water resources, predicting avalanche risks, and studying the impacts of climate change.

1.4. Objectives

1. Explore the Use of SAR Datasets to Fill Spatial Coverage Gaps across ICESat-2 Data beams
ICESat-2, despite providing high-precision elevation data, suffers from limited spatial coverage due to its narrow swath width and sparse ground tracks. This research aims to leverage the extensive spatial coverage of SAR data from SAOCOM 1B and Sentinel-1 to derive continuous snow depth estimates across larger areas, thereby filling the information gaps in ICESat-2 data.

2. Evaluate the Effectiveness of Different Machine Learning Models for Predicting Snow Depth
This objective aims to apply and compare the performance of standard machine learning models, including Random Forest, Gradient Boosting, and Support Vector Regression, in predicting snow depth from SAR absolute coherence data. The focus is on identifying the most accurate model and concluding the reliability of machine learning for this task.

3. Compare the Snow Depth Prediction Performance of SAOCOM 1B and Sentinel-1 Datasets
This objective seeks to assess and compare the accuracy of snow depth predictions derived from SAOCOM 1B's L-band SAR data and Sentinel-1's C-band SAR data. The comparison aims to highlight the advantages and limitations of each dataset in snow depth estimation, providing insights into their relative effectiveness.

1.5. Research Questions

This study aims to address the following research questions:

1. How can the extensive spatial coverage of SAR data from SAOCOM 1B and Sentinel-1 be utilized to fill the spatial coverage gaps in ICESat-2 data for snow depth estimation?

2. Which machine learning model, Random Forest, Gradient Boosting, or Support Vector Regression provides the most accurate predictions of snow depth from SAR absolute coherence data?

3. What are the comparative strengths and limitations of SAOCOM 1B's L-band SAR data and Sentinel-1's C-band SAR data in predicting snow depth?

1.6. Significance of the Study

This study significantly advances snow depth estimation techniques by utilizing ICESat-2 LiDAR data as a reference to train machine learning models with SAR data from SAOCOM 1B and Sentinel-1. The innovative use of different machine learning models enhances the accuracy and reliability of snow depth measurements across various terrains and snow conditions. Additionally, the study offers a comprehensive evaluation of interferometric coherence derived from L-band and C-band SAR data, highlighting their strengths and limitations for snow depth estimation. Accurate snow depth data is vital for understanding

the Earth's hydrological cycle and its impact on climate change. This research enhances climate monitoring by providing improved snow depth estimates, which are essential for assessing the effects of climate variations on snow and ice-covered regions. The methodologies and findings of this study support water resource management, disaster risk reduction, and policy and decision-making efforts related to climate change adaptation, contributing to a deeper understanding of snowpack dynamics and fostering further research and innovation in remote sensing and cryospheric science.

1.7. Thesis Structure

This thesis is structured into six chapters, starting with the introduction. Chapter 2 discusses the study area, while Chapter 3 describes the datasets utilized in this research. Chapter 4 explains the proposed methodology, including preprocessing and coregistration of SAR datasets, extracting snow depth from ICESat-2 data, and applying machine learning models for snow depth prediction. Chapter 5 presents the results, evaluating the performance of the machine learning models and comparing snow depth predictions from SAOCOM 1B and Sentinel-1 datasets. This chapter also discusses the strengths and limitations of the proposed approach. Chapter 6 concludes the thesis by summarizing key findings, discussing the implications, and offering recommendations for future research.

2. STUDY AREA

The study area for this research is in the Sermersooq province in Southern Greenland, characterized by its dynamic and ever-changing snowpack (Figure 1). Greenland is a critical area for studying snow depth and snowpack dynamics due to its significant contribution to global sea-level rise and unique climatic conditions. The Sermersooq region experiences substantial snowfall and snowpack variations due to its polar climate (Fischer et al., 2019). The region's challenging snow conditions make it an excellent location to test the capabilities of advanced snow depth estimation methods. The study area's latitudinal extent ranges from -47.1846° to -47.2453°, and its longitudinal extent ranges from 61.5916° to 62.4779°.

Sermersooq features diverse terrain, including glaciers, ice caps, fjords, and mountainous areas. These various topographic features present challenges in snow depth estimation, and the effectiveness of the research in such complex terrain can be a valuable demonstration of its capabilities. Greenland is highly susceptible to the effects of climate change, making it an important area for climate research (Fischer et al., 2019). The study in Sermersooq is relevant for understanding how changes in snowpack and snow depth contribute to broader climate change research, including the impact on sea-level rise. The region aligns with many scientific priorities and initiatives focused on polar analysis, ice sheet dynamics, and the changing Arctic environment (Jouvet et al., 2019).

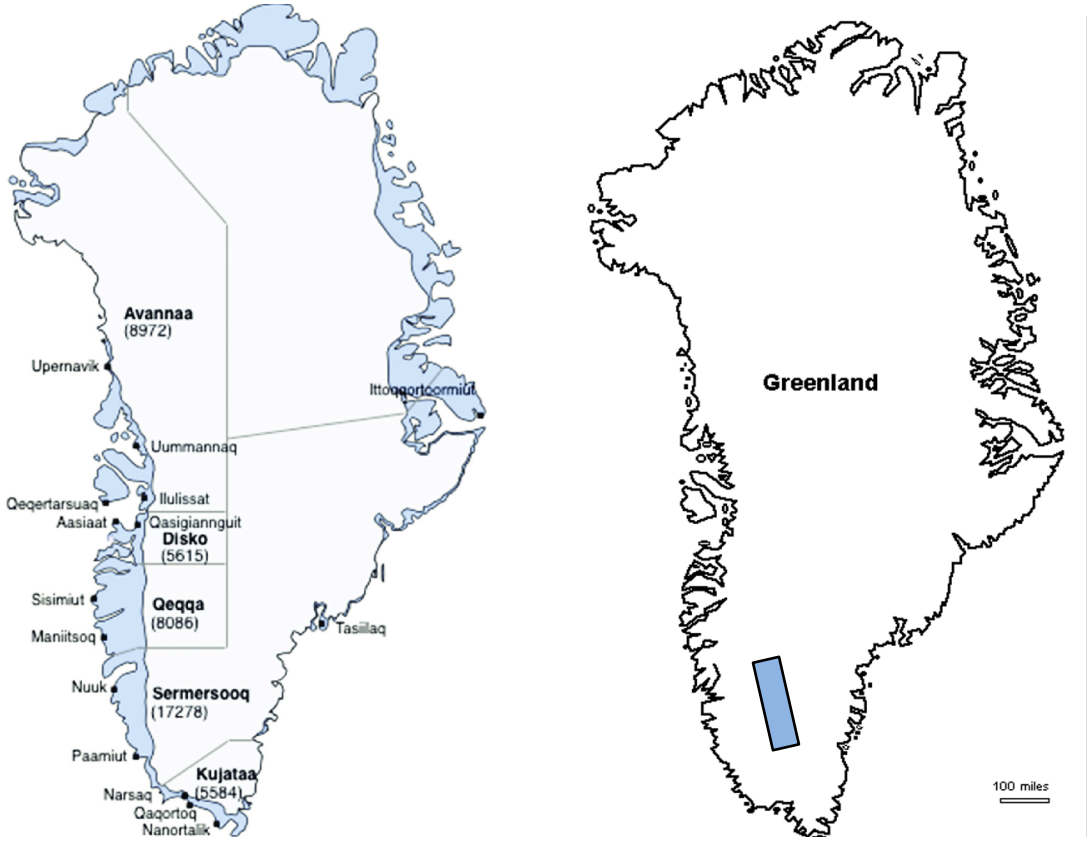


Figure 1: The blue polygon in the right part of the figure indicates the study area located in the Sermersooq province of Southern Greenland

3. DATASETS USED

In this chapter, we elaborate on the specific datasets utilized in this research, discussing the rationale behind selecting SAOCOM 1B, ICESat-2, and Sentinel-1 for snow depth estimation. Each dataset brings unique strengths that enhance the thorough analysis and accurate prediction of snow depth.

3.1. SAR Data

SAR data serve as the primary input for training the machine learning model, which is then used to predict and fill gaps in the ICESat-2 LiDAR data. This study utilizes two SAR datasets: SAOCOM 1B and Sentinel-1. The simultaneous use of these datasets offers a comprehensive approach to snow depth estimation.

3.1.1. SAOCOM 1B

The SAOCOM 1B satellite, operated by the Argentine Space Agency (CONAE), functions in the L-band frequency, which is particularly advantageous for snow depth estimation due to its deep penetration capabilities. The specific data used in this study were acquired in Stripmap mode with Quad Polarization (HH, HV, VH, VV) on two dates: September 13, 2023, and September 29, 2023. This configuration enables detailed analysis of snow properties by examining the backscatter response in different polarizations. The dataset features a ground range resolution of 10 meters and an azimuth resolution of 6 meters, with a swath width ranging from 20 to 40 kilometers.

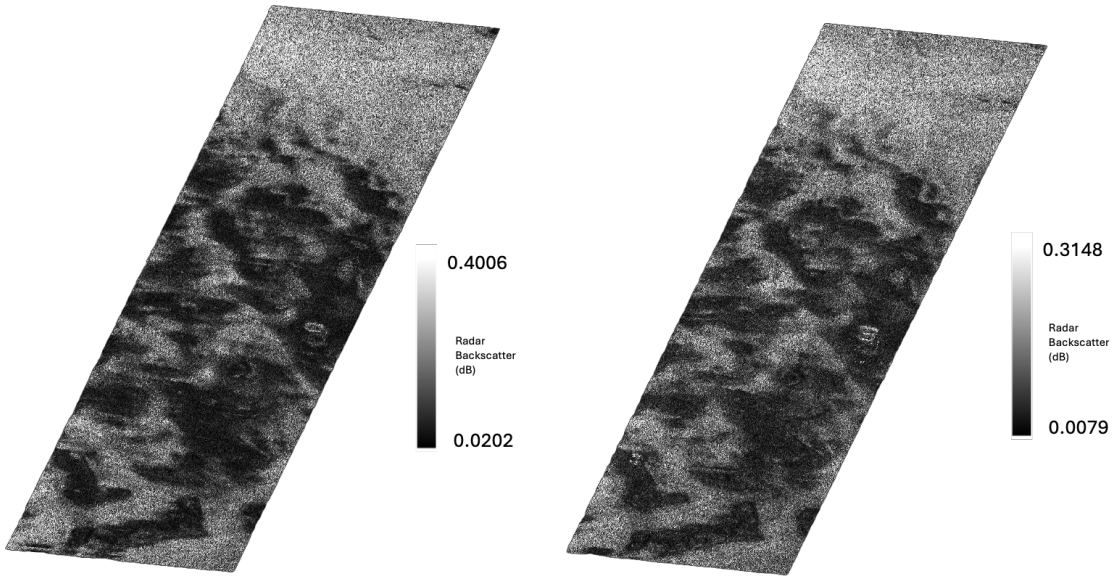


Figure 2: SAOCOM 1B images representing intensity of HV polarisation from master (on left) and slave (on right) datasets

The L-band radar waves can penetrate through dense vegetation and snow, providing valuable information about the subsurface characteristics (Cazcarra-Bes et al., 2020). This makes SAOCOM 1B ideal for monitoring snowpack in complex terrains where traditional methods may fall short. Additionally, the radar signals are highly sensitive to changes in snow characteristics, such as density and wetness, making it an effective tool for estimating snow water equivalent (SWE) (Rott et al., 2021).

Table 1: Parameters Associated with SAOCOM 1B

Parameter	Description
Platform	SAOCOM 1B
Instrument	L-Band SAR, 1.275 GHz
Sensor Mode	Stripmap
Processing Levels	L1A - SLC, L1B - DI, L1C - GEC, L1D - GTC
Resolution	10m (ground range) x 6m (azimuth)
Swath Width	20 - 40 km
Polarization	Quad Polarization (HH, HV, VH, VV)

The capability of L-band SAR to penetrate deeper into the snowpack allows for capturing more comprehensive data about the snow's subsurface layers. This is particularly useful for regions with thick snow cover, where surface observations alone may not provide an accurate representation of snow depth. (Rekioua et al., 2017). Furthermore, the multi-polarization channels (VV, VH, HH, HV) enhance the ability to analyze the anisotropy of snow, which is crucial for accurate snow depth estimation.

3.1.2. Sentinel-1

The Sentinel-1 mission, a component of the European Space Agency's (ESA) Copernicus program, includes two satellites, Sentinel-1A and Sentinel-1B, operating in the C-band frequency. Although the C-band does not penetrate as deeply as the L-band, it offers high spatial resolution and frequent revisit times, which are advantageous for temporal analysis of snow cover (Garg et al., 2022). The data used in this study were acquired from Sentinel-1A in dual polarization (HH, HV) mode on two dates: September 18, 2023, and September 30, 2023. These dates were chosen to ensure a consistent temporal framework as same as SAOCOM. Sentinel-1 provides detailed imagery essential for capturing fine-scale snow cover dynamics in open and less vegetated areas (Dahhani et al., 2022). With frequent revisit times, typically every 6-12 days, Sentinel-1 allows for continuous monitoring of snow cover changes over time. The use of Sentinel-1 data for snow depth prediction through machine learning integration with ICESat-2 is primarily exploratory and comparative in this study.

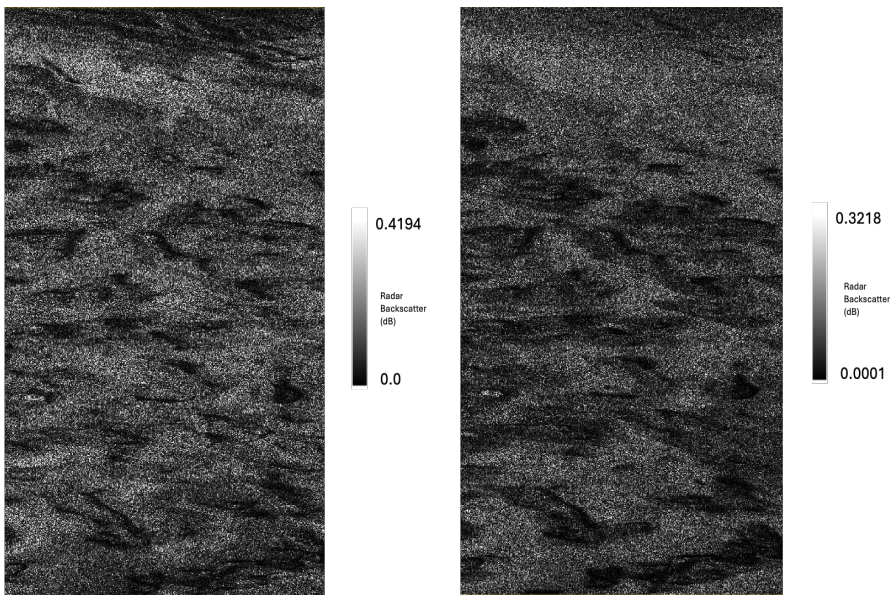


Figure 3: Sentinel-1 images representing intensity of HV polarisation from master (on left) and slave (on right) datasets

Despite the shallower penetration of C-band SAR, its higher spatial resolution is advantageous for detecting surface characteristics and monitoring snow cover changes with high temporal resolution (Garg et al., 2022). Sentinel-1's ability to capture fine-scale dynamics of snow cover in less vegetated areas makes it particularly useful for monitoring seasonal variations and short-term changes in snowpack. This feature is critical for understanding snowmelt patterns, predicting water availability, and managing water resources in regions where snow is a primary source of freshwater (Askne et al., 2017).

Table 2: Parameters Associated with Sentinel-1

Parameter	Description
Platform	Sentinel-1A
Instrument	C-Band SAR, 5.405 GHz
Sensor Mode	Interferometric Wide Swath (IW)
Processing Levels	Level-1 (Single Look Complex - SLC)
Resolution	5m (range) x 20m (azimuth)
Swath Width	250 km
Polarization	Dual Polarization (HH, HV)

Table 3 provides a summary of the acquisition dates for the datasets used in this research. These dates were strategically chosen to maximize data relevance and coherence across the different satellite missions and ensure temporal alignment between datasets for accurate comparative analysis.

Table 3: Acquisition Dates for the Datasets used

Dataset	Acquisition Date 1	Acquisition Date 2
SAOCOM 1B	2023-09-13	2023-09-29
Sentinel-1A	2023-09-18	2023-09-30
ICESat-2 (ATL03 & ATL08)	2023-09-13	

3.2. Space-based Lidar data (ICESat-2)

ICESat-2 data will be considered as reference data when building the machine learning model. ICESat-2, launched by NASA, uses the Advanced Topographic Laser Altimeter System (ATLAS) to provide precise surface elevation measurements through LIDAR technology (Narine et al., 2020). This dataset serves as the ground truth for snow depth estimation in this study, offering highly accurate elevation data that can be directly related to snow depth. ICESat-2 provides elevation measurements with centimeter-level accuracy, making it a reliable source for validating snow depth estimates (Lu et al., 2022). The ICESat-2 mission produces several key data products relevant to this study:

ATL03 (Photon Heights): This product contains geolocated photon data, which includes the precise height measurements of individual photons reflected off the Earth's surface (Zhao et al., 2022). These photon heights are critical for determining the surface elevation, particularly during snow-covered periods. The ATL03 data is used to capture the snow surface height, which is essential for calculating the snow depth when compared to ground heights.

ATL08 (Terrain and Canopy Heights): This product provides data on terrain and canopy height, offering a detailed representation of the ground surface beneath the snow (Enderlin et al., 2022). The ATL08 data is instrumental in identifying the ground elevation, serving as a reference to determine the snow depth by subtracting this value from the snow surface height captured in the ATL03 product.

The snow depth is not directly available as a product from ICESat-2 data; it requires pre-processing to extract snow depth. This pre-processing involves comparing the surface elevation measurements taken from ATL03 during snow-covered periods with the ground heights from ATL08. This methodology ensures that the extracted snow depth values are accurate and can be effectively used for validating and training machine learning models. The acquisition dates for both ATL03 and ATL08 is September 13, 2023.

Table 4: Parameters Associated with ICESat-2

Parameter	Description
Platform	ICESat-2
Instrument	Advanced Topographic Laser Altimeter System (ATLAS)
Data Products	ATL03 (Photon Heights), ATL08 (Terrain and Canopy Heights)
Elevation Accuracy	Centimeter-level
Spatial Coverage	Extensive, but with narrow swath width and sparse ground tracks
Vertical Resolution	~0.2 meters
Along-Track Resolution	0.7 meters
Across-Track Resolution	17 meters
Revisit Time	91 days
Beam Configuration	6 beams (3 strong and 3 weak), spaced 3.3 km apart
Application	Surface elevation measurements, snow depth estimation

ICESat-2 operates with six beams (see Figure 4), arranged in three pairs, with each pair consisting of a weak beam and a strong beam. These beams are spaced approximately 3.3 km apart, creating gaps in the data coverage. The gaps across the beams necessitate the integration of additional data sources to fill these spatial gaps, making SAR data a valuable complement.

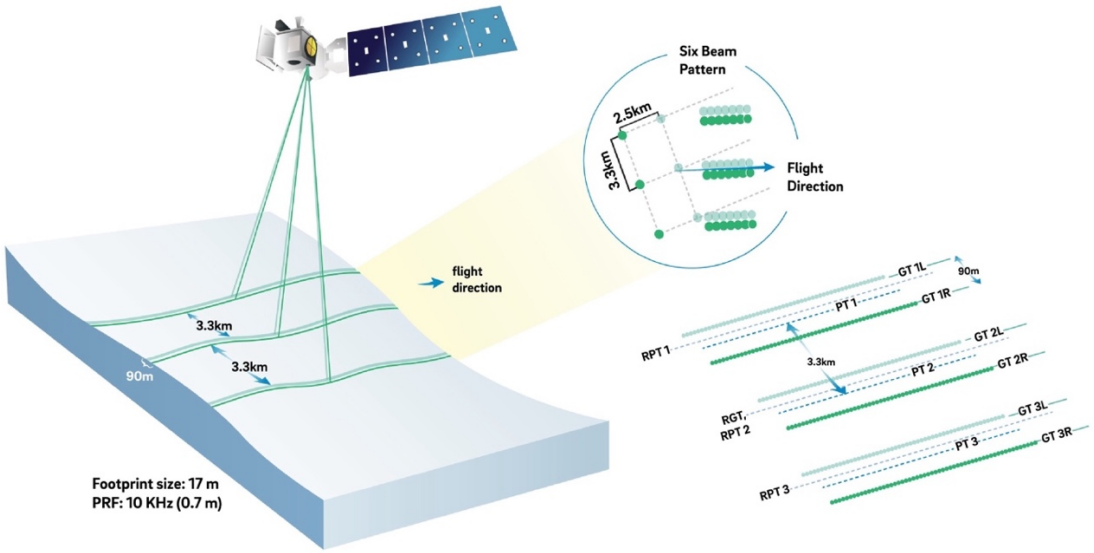


Figure 4: This graphic shows the six-beam pattern from the ATLAS instrument and how it measures ice thickness (Smith et al., 2019)

4. METHODOLOGY

This chapter provides a comprehensive overview of the methodology employed in this research to estimate snow depth using SAR data from SAOCOM 1B and Sentinel-1, with ICESat-2 LIDAR data serving as a reference for training machine learning models. The primary goal is to leverage the extensive spatial coverage and high-resolution data from SAR alongside the precise elevation measurements from ICESat-2 to develop robust and accurate snow depth prediction models. The following flowchart outlines the key steps involved in the methodology, with each step discussed in detail below.

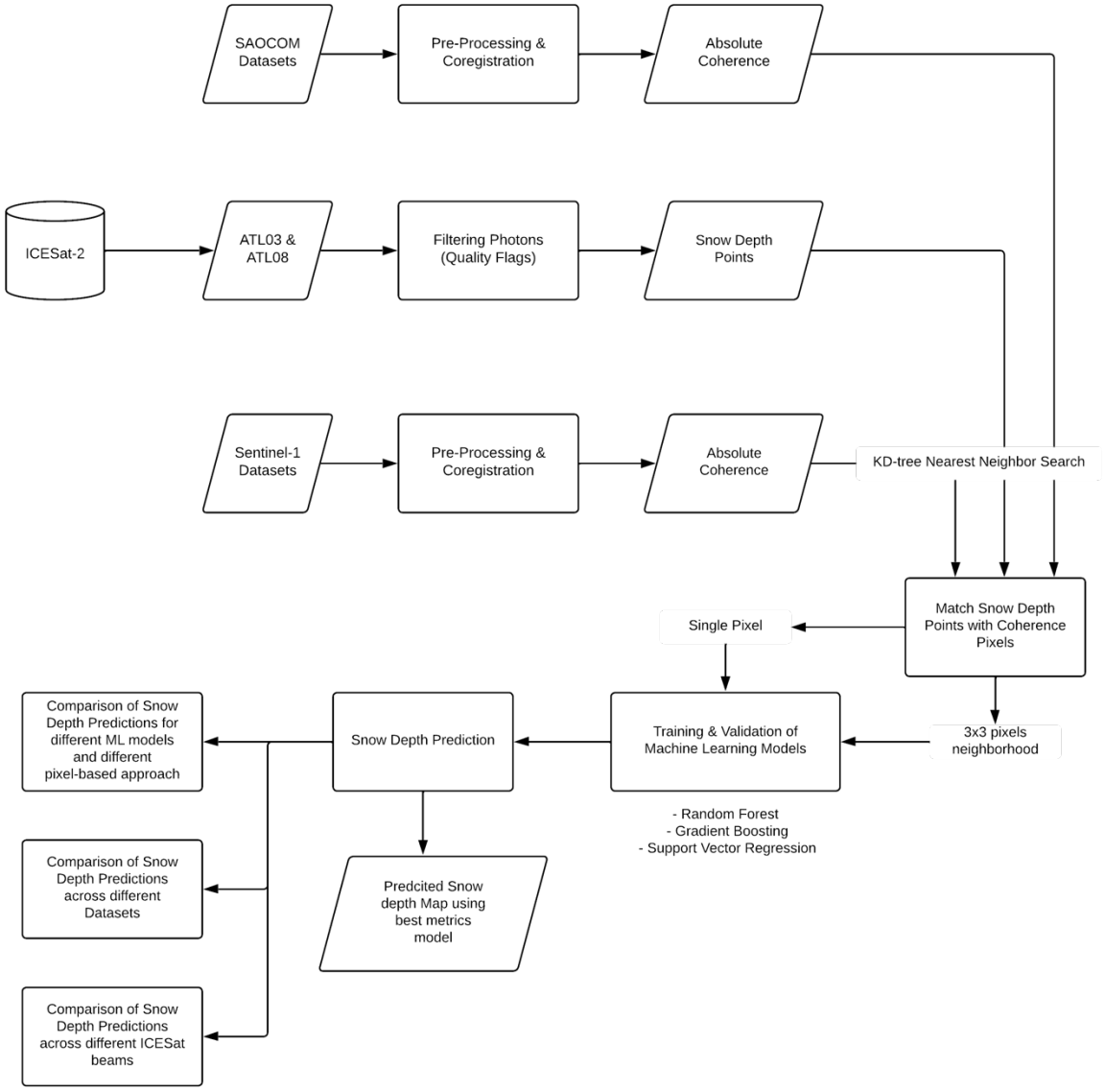


Figure 5: Methodology Flowchart

4.1. Pre-processing of ICESat-2 Data to extract snow depth

This section outlines the comprehensive methodology involved in pre-processing ICESat-2 data, specifically the ATL03 and ATL08 products, to derive accurate snow depth measurements. These measurements serve as a critical reference for training machine learning models with SAR data.

4.1.1. Data Acquisition and Initial Filtering

The ICESat-2 satellite, equipped with the Advanced Topographic Laser Altimeter System (ATLAS), provides high-precision elevation measurements crucial for snow depth estimation. This study utilizes two specific ICESat-2 products: ATL03, which provides photon heights, and ATL08, which delivers terrain and canopy heights.

The first step involves extracting ground height data from the ATL08 product, including latitude (lat_ground), longitude (lon_ground), and terrain height (h_ground). Similarly, photon data is extracted from the ATL03 product, encompassing photon latitude (lat_ph), longitude (lon_ph), and height (h_ph). Only photons with heights greater than 0 are retained, ensuring that only measurements above the ground surface are considered. Additionally, the ICESat-2 ATL03 dataset includes signal confidence and quality flags, which are used to identify and filter out low-confidence signals. By retaining only high-confidence photons, the quality and reliability of the snow depth measurements are significantly improved.

An overlay of ATL03 photon heights and ATL08 terrain heights is created to visualize alignment and consistency between these datasets. Figure 6 below shows this overlay, where photon data from ATL03 (indicated in various colors representing different height ranges) is superimposed on terrain heights from ATL08. This visualization aids in understanding the spatial distribution and accuracy of the photon measurements in relation to the terrain data, confirming the datasets validity for snow depth estimation.

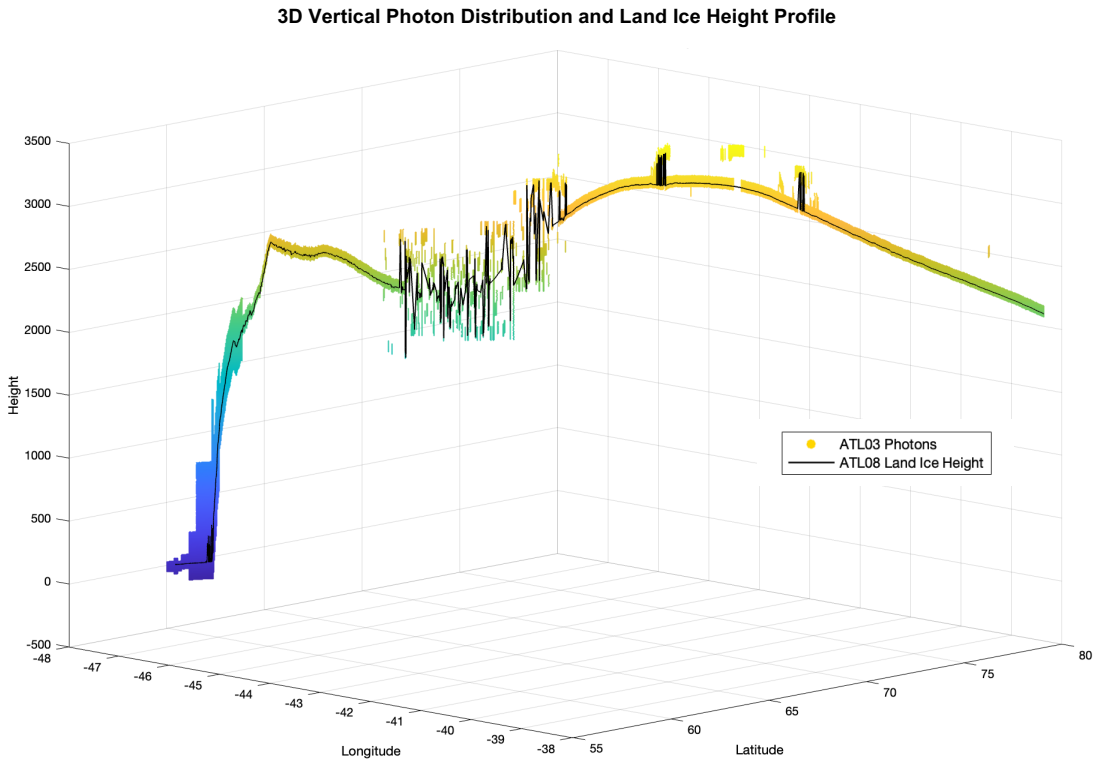


Figure 6: ATL03 Photons and ATL08 surface height overlay (for gt1l beam)

4.1.2. Signal Confidence and Quality Flags

The ICESat-2 ATL03 dataset includes signal confidence and quality flags for photons, which identify and filter out low-confidence signals. Figure 7 shows the distribution of photon heights with different signal confidence levels: low (red), medium (blue), and high (green). Retaining only high-confidence photons enhances the reliability of the snow depth measurements. This figure demonstrates the importance of using quality flags to ensure that only the most reliable photon returns are used in the analysis.

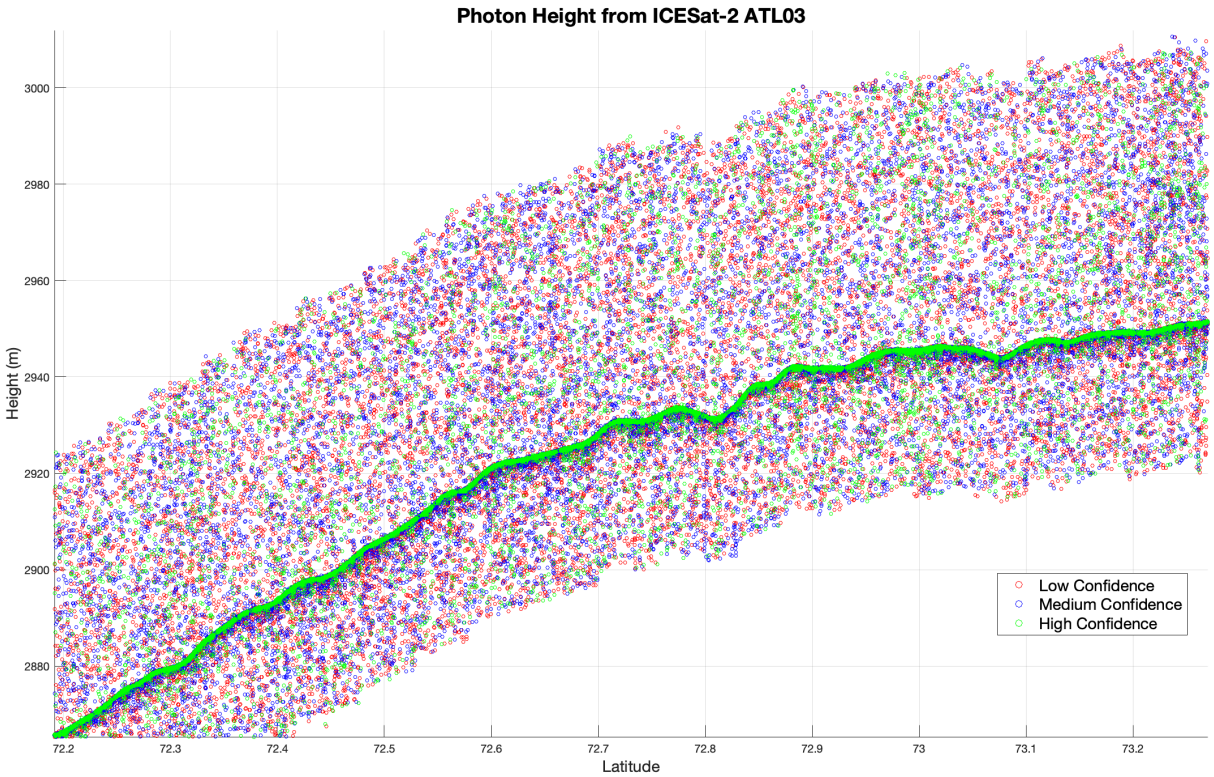


Figure 7: Photon profile for gt1l beam: Signal Confidence (zoomed view)

4.1.3. Smoothing and Interpolation

A smoothing technique using a moving average filter is applied to the ground height data from ATL08. A moving average filter with a window size of 10 reduces noise and provides a more accurate representation of the terrain surface. This smoothed data is crucial for accurate interpolation (Kwok et al., 2019). Smoothing eliminates high-frequency noise that could interfere with the interpolation process, ensuring the ground height data represents the true terrain surface. An interpolant is created using the smoothed ground height data to estimate the ground surface height at each photon location. A scattered interpolation method constructs a surface passing through known ground height points, estimating the ground surface height at the photon locations. This process is critical for determining the reference ground height directly beneath each photon, necessary for calculating snow depth. Scatter plots verify the accuracy of the interpolated ground surface data, allowing for visual inspection of the interpolated values and identification of any anomalies or inconsistencies. Verification helps detect and correct potential errors in the interpolation process, ensuring the reliability of the subsequent snow depth calculations.

4.1.4. Calculation of Snow Depths

Snow depths are calculated by subtracting the interpolated ground surface heights from the photon heights, providing an initial estimation of the snow depth at each photon location. This measures the vertical distance between the snow surface (indicated by the photon heights) and the underlying ground surface (provided by the interpolated ground heights). Additional filtering removes any physically implausible snow depth values. Snow depths below 0 meters are excluded, indicating errors in the photon or ground height data. Focusing on valid photon heights and utilizing quality flags ensures that the calculated snow depths are within a plausible range, thereby improving the reliability of the snow depth estimates.

Figure 8 illustrates the variability in photon heights within a single pixel across latitude, longitude, and time. This detailed analysis of photon distribution is crucial for correlating with SAR coherence values. Understanding the precise spatial and temporal distribution of photon heights allows researchers to better match these LiDAR data points with SAR data, enhancing the robustness and accuracy of the snow depth models.

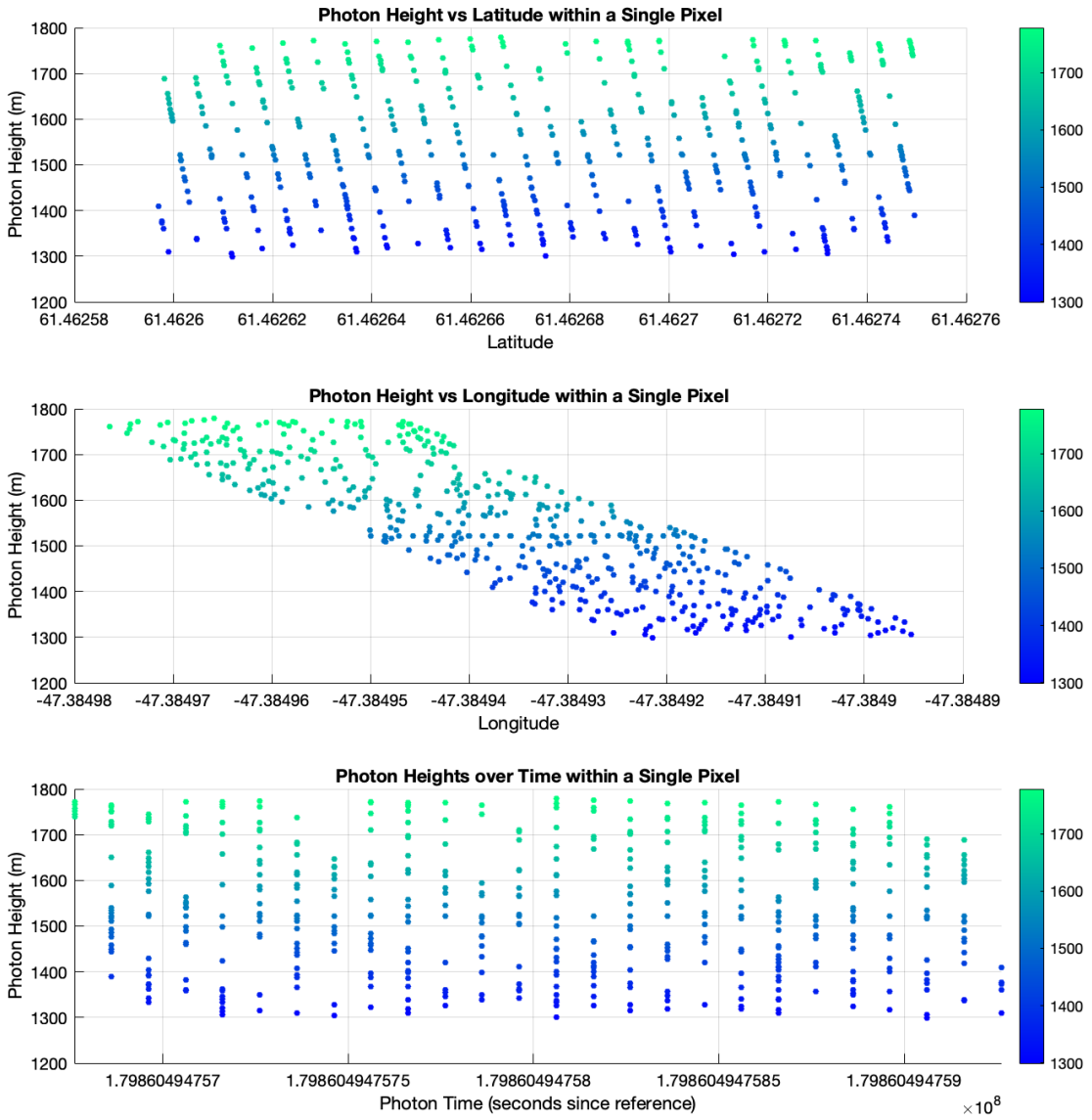


Figure 8: Photon distribution within a single pixel from ICESat-2, illustrating variability in photon heights across latitude, longitude, and time.

Following these detailed pre-processing steps transforms the ICESat-2 data into a reliable reference dataset for snow depth estimation. The rigorous application of filtering based on photon heights, and quality flags, combined with smoothing and interpolation techniques, ensures the accuracy and reliability of the derived snow depth measurements. This pre-processed data is then ready to be used in conjunction with SAR data to train machine learning models and enhance the accuracy of snow depth predictions. This comprehensive approach leverages the high-precision elevation data from ICESat-2 while addressing the spatial coverage limitations, providing a robust framework for snow depth estimation.

4.2. Preprocessing and Coregistration of SAOCOM Datasets

The preprocessing of SAR data is crucial to ensure the accuracy and reliability of subsequent analyses, particularly for snow depth estimation. This section details the steps taken to preprocess SAOCOM 1B and Sentinel-1 SAR data, focusing on radiometric and geometric corrections, and coregistration.

4.2.1. Radiometric and Geometric Correction

Radiometric and geometric corrections are essential preprocessing steps for SAR data. Radiometric correction addresses sensor-specific distortions and standardizes backscatter values across different acquisitions. This ensures that the SAR data is consistent and comparable, which is vital for accurate analysis. In SNAP, the "Radiometric Calibration" module is used for this purpose. This module adjusts the SAR data based on the sensor characteristics and imaging geometry, converting raw SAR data into calibrated backscatter values that are reliable for further analysis.

Geometric correction aligns the SAR data to a common geographic coordinate system, ensuring accurate spatial positioning on the Earth's surface. This is particularly important when integrating SAR data with other spatial datasets such as optical imagery and digital elevation models (DEMs). In SNAP, the "Range-Doppler Terrain Correction" module, which uses a Global Earth Topography and Sea Surface Elevation (GETASSE30) DEM, is employed to correct distortions caused by the imaging geometry and Earth's curvature. These corrections ensure that the SAR data is geometrically consistent and accurate, laying a solid foundation for subsequent coherence analysis and snow depth estimation.

4.2.2. Coregistration of Multitemporal SAR Datasets

Coregistration involves aligning multiple SAR images of the same region captured at different times. This step is critical for generating coherence maps, which are used to analyze surface changes over time, including snow dynamics (Pardini et al., 2019). Accurate coregistration ensures that corresponding pixels in the SAR images represent the same ground location, facilitating reliable temporal analysis.

For SAOCOM 1B data, **DEM-assisted coregistration** is employed using the SNAP software. This method leverages a digital elevation model (DEM) to accurately align the SAR images, correcting for any geometric distortions. The key parameters for this coregistration include the use of a high-resolution DEM (GETASSE30), nearest-neighbor resampling to preserve original data values, and setting the RMS threshold to 1.0 pixel for high-precision alignment. Additionally, the polynomial order for warping is set to 1, indicating a linear transformation, which is sufficient for most terrain types. The result of this meticulous coregistration process is the generation of six complex bands for each polarization channel (HH, HV, VH, VV) from SAOCOM 1B, corresponding to the real and imaginary parts, as well as intensity, leading to a total of 24 bands.

For Sentinel-1 data, a different approach is employed due to its unique data acquisition method. Initially, Sentinel-1 data is processed using the **Sentinel-1 TOPS Split** and **Deburst** modules in SNAP. These steps handle the burst mode acquisition of Sentinel-1, ensuring the data is split into consistent segments and debursted to remove redundant information. Subsequently, the **Sentinel-1 TOPS Coregistration** module aligns the Sentinel-1 datasets accurately. Each polarization channel (HH, HV) is coregistered separately, resulting in six complex bands for each polarization channel (real and imaginary parts, as well as intensity), leading to a total of 12 bands (6 bands per polarization * 2 polarizations). This precise coregistration is crucial for subsequent coherence analysis and accurate snow depth estimation.

4.3. Generating Absolute Coherence from SAR Datasets

Complex coherence is a critical measure that quantifies the similarity between two SAR images acquired at different times(Brolly et al., 2016). The process of generating absolute coherence from SAR datasets, specifically SAOCOM 1B and Sentinel-1, is based on the Random Volume over Ground (RVoG) model, which describes the volume scattering effects in interferometric SAR (InSAR) coherence (Olesk et al., 2016). The RVoG model assumes that a homogeneous layer of volume scatterers, such as a forest canopy, is located above a reflective ground layer (Olesk et al., 2016).

The coherence γ is defined as the normalized complex cross-correlation between two SAR images, separated by a baseline (Olesk et al., 2016), and is given by:

$$\gamma = \frac{\langle s_1 s_2^* \rangle}{\sqrt{\langle s_1 s_1^* \rangle \langle s_2 s_2^* \rangle}} \quad \text{Eq. (1)}$$

In general, the measured coherence γ can be expressed as a product of several factors:

$$\gamma = \gamma_{\text{System}} \cdot \gamma_{\text{SNR}} \cdot \gamma_{\text{Temp}} \cdot \gamma_{\text{Vol}} \quad \text{Eq. (2)}$$

where γ_{System} accounts for system-related decorrelation, γ_{SNR} for signal-to-noise ratio effects, γ_{Temp} for temporal decorrelation, and γ_{Vol} for volume scattering-induced decorrelation. In our study, the focus is on volumetric coherence (γ_{Vol}), which can be linked to forest height using the RVoG model. The volumetric coherence is represented by:

$$\gamma_{\text{Vol}} = \frac{\sigma(e^{h\sigma} e^{i\kappa_z h} - 1)}{(\sigma + i\kappa_z)(e^{h\sigma} - 1)} \quad \text{Eq. (3)}$$

where **σ** is the volume extinction coefficient, *h* is the height of the volume layer, and κ_z is the vertical wavenumber. According to the RVoG model, the total coherence can be written as:

$$\gamma_{\text{RVoG}} = e^{i\phi} \left(\gamma_{\text{Vol}} + \frac{\mu(\omega)}{1 + \mu(\omega)} \right) \quad \text{Eq. (4)}$$

where **μ(ω)** represents the ground reflection term and **ϕ** is the ground phase.

According to Equation 3, snow depth (parameter (h)) is embedded within the volumetric coherence. The presence of temporal decorrelation in the measured coherence complicates the direct retrieval of volume coherence for snow depth estimation. However, when SAR images are acquired simultaneously, such as with Tandem-X, which captures two images of the same area at the same time, the measured coherence is not affected by temporal decorrelation. By excluding noise and SNR decorrelation, the coherence measured by Tandem-X can be accurately modeled as in Equation 4, facilitating the recovery of volume coherence and subsequently the height (Olesk et al., 2016).

Some studies have used simplified versions of the RVoG model, as seen in this paper (Olesk et al., 2016) which introduces Equation 11 and correlates the absolute value of coherence with the parameter h or snow depth. The study demonstrated a strong correlation between the absolute value of volume coherence and snow depth. To further illustrate, the left part of the image below from the same paper shows the relationship between the absolute value of coherence and height, supporting how the absolute value of coherence is related to snow depth.

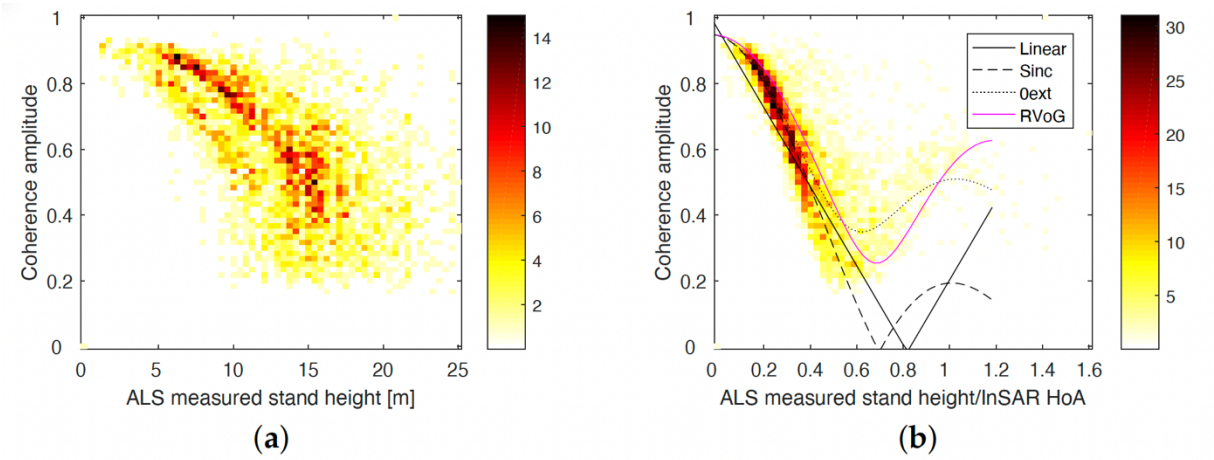


Figure 9: Coherence against Airborne LiDAR Scanned (ALS) forest height The colors represent how many coherence-stand height pairs fall into the value range. On the left (a), the coherence magnitude is compared to forest stand height, and on the right (b), the coherence magnitude is compered to ALS forest height divided by HoA (9) values (Olesk et al., 2016).

However, coherence measurements from SAOCOM 1B and Sentinel-1 are influenced by various disturbing factors since the SAR images are obtained at different times, introducing temporal decorrelation. This makes it challenging to recover volumetric coherence (γ_{Vol}) directly. Therefore, instead of directly recovering (γ_{Vol}), the measured coherence is provided to a machine learning model that identifies relevant informative features and extracts snow depth. By feeding the measured coherence values into the machine learning model, the complexities of isolating volumetric coherence are bypassed. The model leverages the correlation between coherence values and snow depth to predict snow depth accurately, even when the measured coherence is influenced by various disturbing factors.

4.3.1. Absolute Coherence Calculation (SAOCOM 1B)

The absolute coherence maps for the SAOCOM 1B dataset across four polarization channels (HH, HV, VH, VV) are presented in Figure 10. These maps visually represent coherence values, indicating the correlation level between SAR images from different acquisition dates. High coherence values suggest minimal surface changes, while low coherence values indicate significant changes or decorrelation.

The coherence maps for HH and HV polarizations exhibit similar patterns, with high coherence areas indicating stable surface conditions and low coherence areas suggesting snow accumulation or melting. The VH and VV polarization maps show slight variations in spatial distribution, reflecting different sensitivities to surface features. These coherence values are crucial for understanding snow-covered region dynamics, providing insights into accumulation and melting areas. The absolute coherence values generated for SAOCOM 1B data are integral to predicting snow depth and are used for subsequent analysis and correlation with ICESat-2 snow depth measurements.

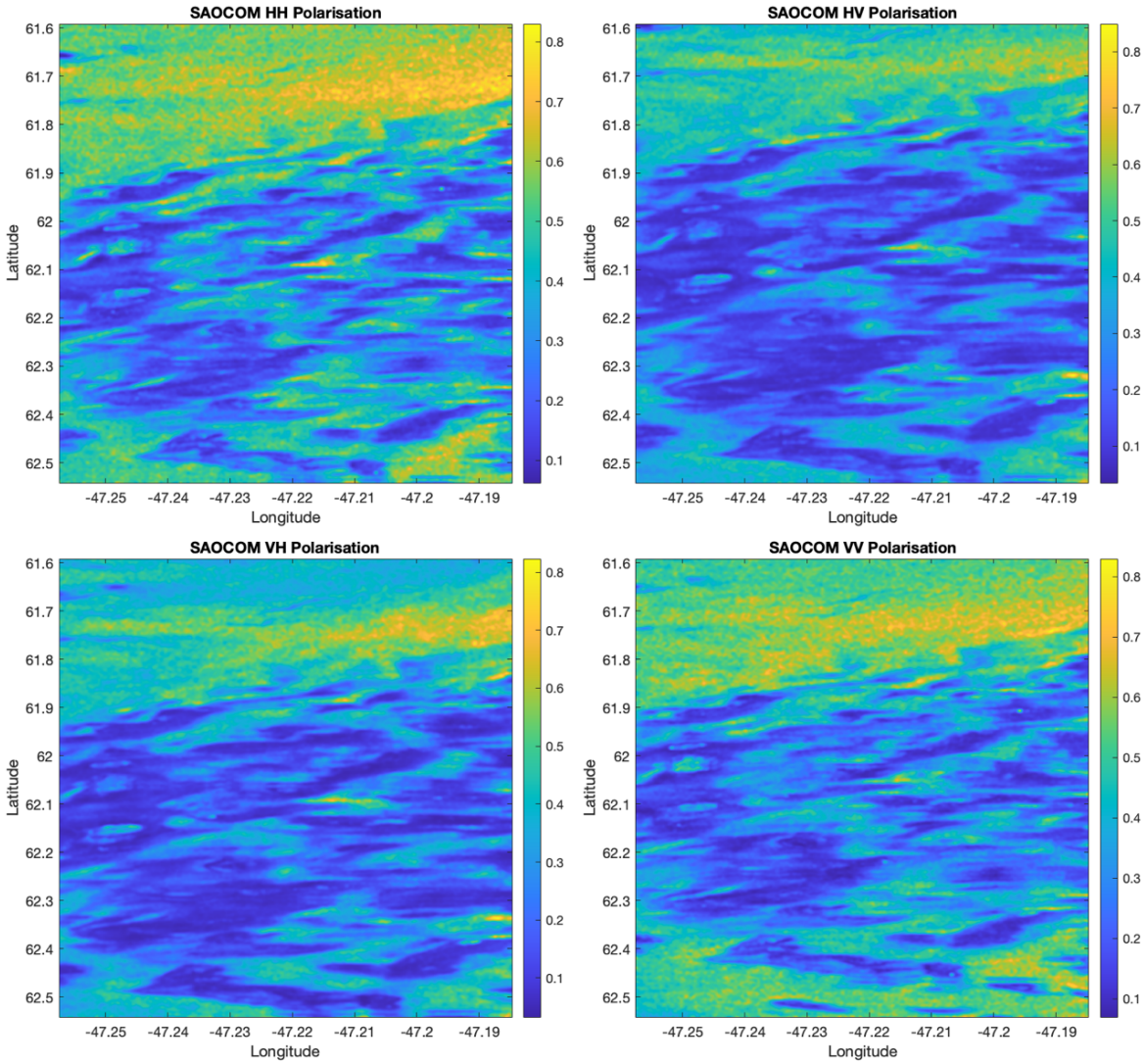


Figure 10: Coherence maps across four SAOCOM 1B polarization channels

4.3.2. Absolute Coherence Calculation (Sentinel-1)

The absolute coherence maps for the Sentinel-1 dataset across HH and HV polarization channels, presented in Figure 11 and 12 respectively, similarly indicate the correlation level between SAR images from different acquisition dates. High coherence values denote stable surface properties, while low coherence values point to significant changes or decorrelation. The HH and HV polarization maps of Sentinel-1 show variability in surface stability across the study area, with high coherence values indicating minimal surface changes and low values suggesting changes due to snow dynamics. These coherence values are essential for snow depth prediction, providing critical data for further analysis and integration with ICESat-2 snow depth measurements.

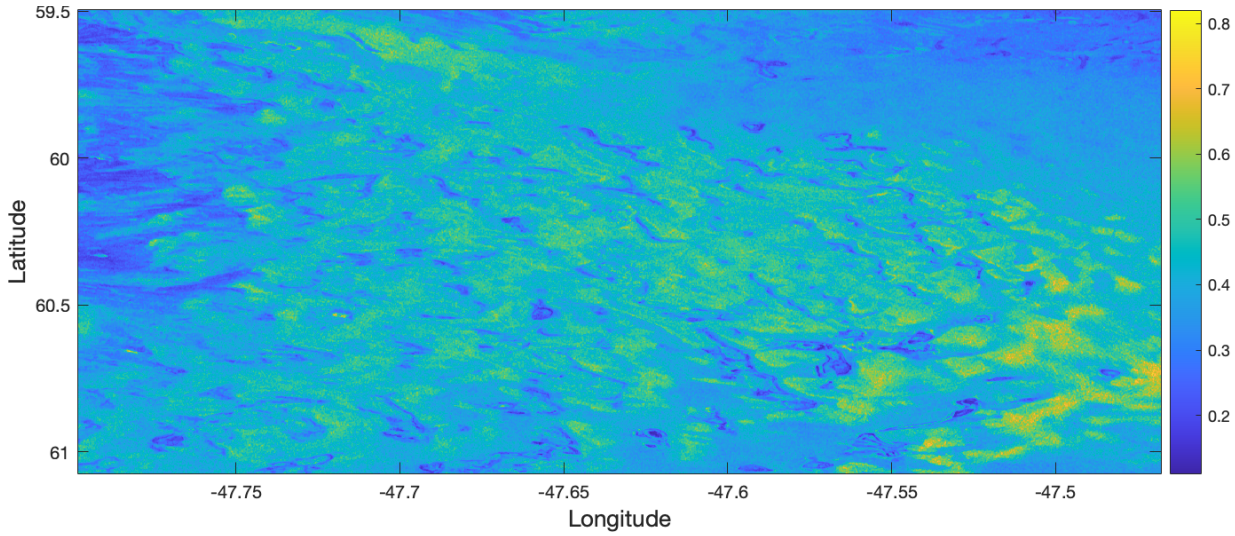


Figure 11: Absolute Coherence Map for Sentinel-1 Polarization HH

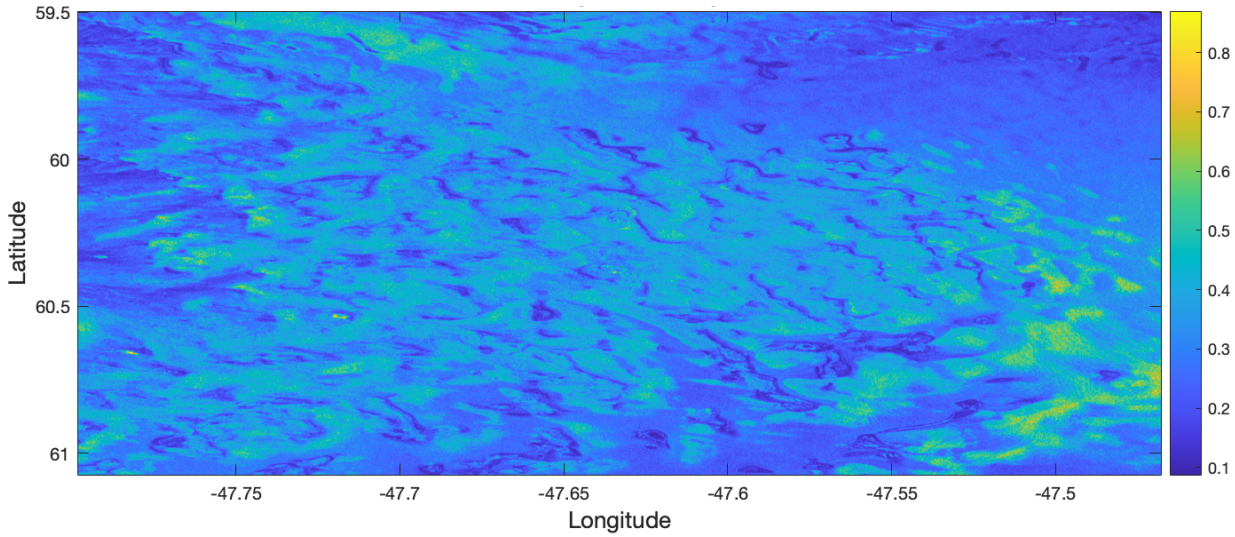


Figure 12: Absolute Coherence Map for Sentinel-1 Polarization HV

The clipped version of the Sentinel-1 coherence maps, shown in Figure 13, aligns with the extent of the SAOCOM 1B data. This clipping ensures a consistent spatial extent for comparing dataset-specific predictions. By examining these coherence maps, regions of interest can be identified for detailed analysis and correlation with snow depth measurements from ICESat-2.

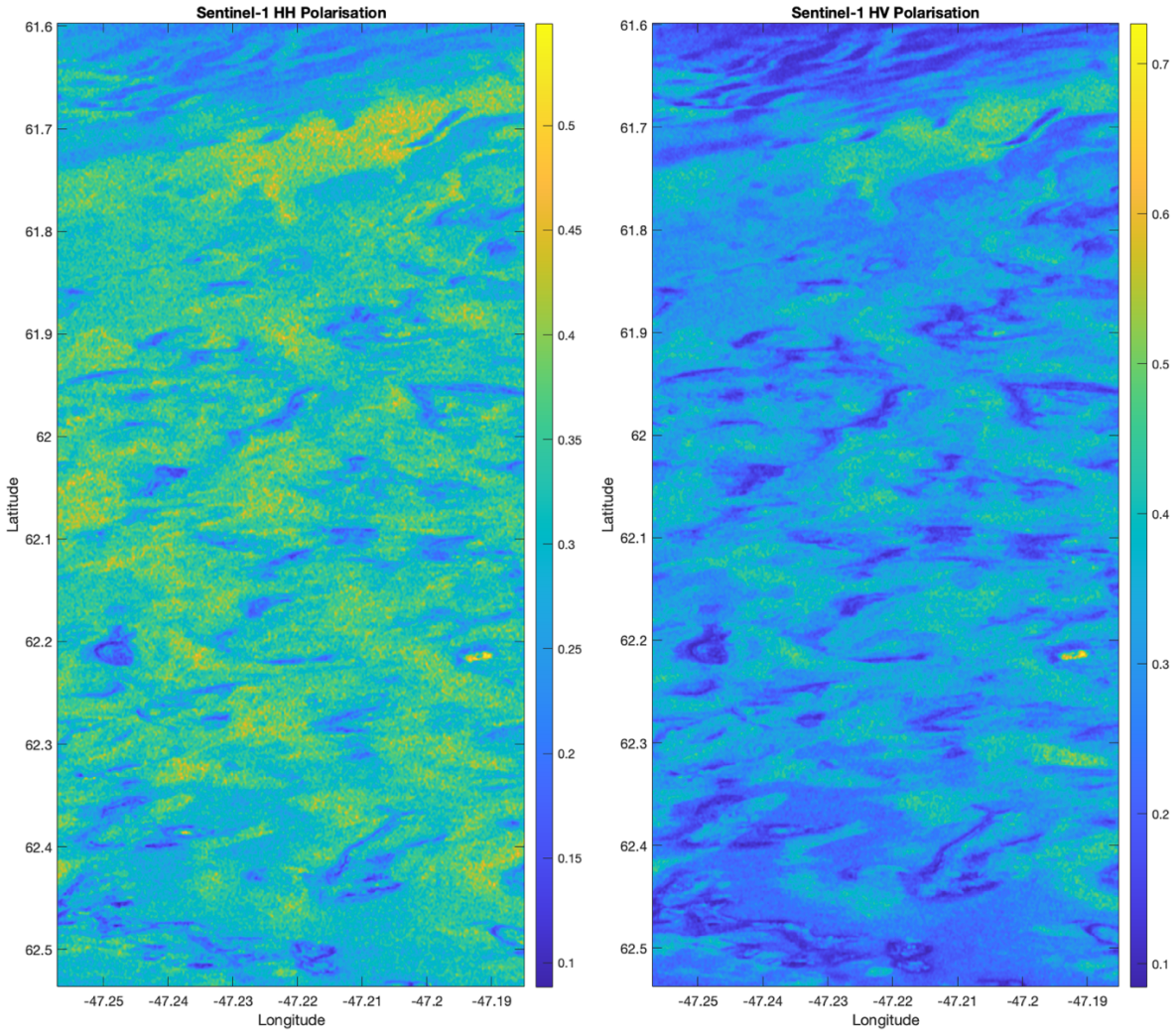


Figure 13: Clipped Absolute Coherence Maps for Sentinel-1 Polarization Channels (HH, HV) with the same extent as SAOCOM 1B

In conclusion, the absolute coherence results for both SAOCOM 1B and Sentinel-1 datasets demonstrate spatial variability in surface stability and provide essential data for predicting snow depth. These results lay the groundwork for further analysis and incorporation with ICESat-2 LiDAR data, ultimately enhancing the accuracy and reliability of snow depth predictions.

4.4. Matching ICESat-2 Snow Depths with SAR Interferometric Coherence Values

Matching ICESat-2 snow depths with SAR coherence values is a critical step in leveraging high-resolution LiDAR data to train machine learning models for snow depth prediction.

Grid Generation and Data Overlay

The process begins with generating a grid for the SAR coherence data based on the latitude and longitude ranges of the ICESat-2 photon data. Both ICESat-2 data points and SAR image pixels have associated latitude and longitude coordinates, allowing for direct overlay. This alignment ensures that the spatial domains of both datasets are properly matched, facilitating subsequent analysis. By overlaying the ICESat-2 photon data points onto the SAR coherence image, corresponding pixels in the SAR image can be identified and matched with ICESat-2 data points.

KD-Tree Construction and Nearest Neighbor Search

To efficiently match the snow depths with corresponding coherence values, a KD-tree data structure is employed. The KD-tree organizes points in a k-dimensional space (latitude and longitude in this case), significantly reducing the computational complexity of nearest neighbor searches. The KD-tree is constructed using the locations of the ICESat-2 photon data points. Each node in the KD-tree represents a photon data point, and the tree is built by recursively splitting the data along the latitude and longitude dimensions. Once constructed, the KD-tree allows efficient querying to find the nearest neighbor of any given point in the coherence data grid. For each point in the coherence grid, the KD-tree is queried to identify the nearest ICESat-2 photon data point, matching each grid point with the corresponding snow depth value from the ICESat-2 dataset. After identifying the nearest neighbors, snow depth values are assigned to the coherence grid. To account for spatial variability and ensure robust matching, a 3x3 window is applied around each matched pixel. This approach helps to smooth out potential mismatches and incorporates spatial context into the analysis. The 3x3 window method assigns the snow depth value not only to the central pixel but also to its immediate neighbors in a 3x3 grid around the matched point, ensuring that the central pixel and its surrounding pixels (a total of nine pixels) receive the snow depth value, thereby enhancing the data's continuity.

Rationale for Using a 3x3 Window

The rationale for employing a 3x3 window includes several factors. First, ICESat-2 measures height using only six beams, resulting in data gaps between beams. Expanding the matched pixels into a 3x3 window helps fill these gaps, ensuring more continuous spatial coverage. Second, the 3x3 window helps mitigate issues of spatial misalignment by averaging out small positional errors between the ICESat-2 and SAR datasets. Third, incorporating spatial context through a 3x3 window reduces the impact of noise and measurement errors, capturing the local spatial variability of snow depth and leading to more reliable predictions. Finally, the 3x3 window approach enhances the spatial continuity and reliability of the matched data, making it more suitable for subsequent predictive modeling. Ensuring that each pixel and its immediate neighbors have consistent snow depth values allows the model to learn from a more coherent dataset.

Both the single pixel approach and the 3x3 pixel neighborhood method will be performed and analysed using machine learning models with the SAR datasets, SAOCOM 1B and Sentinel-1. This comprehensive approach ensures accurate snow depth predictions and prepares the dataset for effective machine learning model training.

4.5. Predictive Modeling Using Machine Learning

In this study, predictive modeling was employed to estimate snow depth using coherence values derived from SAR datasets (SAOCOM 1B and Sentinel-1) with ICESat-2 LIDAR data serving as the reference. The coherence values from the SAR datasets were used as features, and the snow depths from ICESat-2 were used as labels. To ensure the robustness of the prediction models, a comprehensive training, testing, and validation process was implemented.

4.5.1. Training, Testing, and Validation

To thoroughly assess the models, a 10-fold cross-validation method was utilized. This technique divides the dataset into 10 subsets, or folds. Each fold is used as a validation set once, while the remaining nine folds are employed for training. This cycle is repeated 10 times, allowing each fold to serve as the validation set once. This approach ensures that every data point is used for both training and validation, providing a comprehensive evaluation of the model's performance (Phinzi et al., 2021). Unlike a simple train-test split that separates the dataset into a single training and testing set, 10-fold cross-validation offers several benefits. It optimizes data usage, mitigates the risk of overfitting, and yields a stable and reliable estimate of the model's performance by averaging the results across multiple iterations. To evaluate the predictive capabilities of the models during each iteration, performance metrics such as R-squared (R^2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) were computed.

The datasets used for the machine learning models included four different configurations based on the number of matched points (pixels):

1. 1x1 Window (single pixel) for SAOCOM 1B
2. 1x1 Window (single pixel) for Sentinel-1
3. 3x3 Window for SAOCOM 1B
4. 3x3 Window for Sentinel-1

These configurations allowed for a thorough evaluation of the model's performance, providing insights into how the inclusion of neighborhood information (3x3 window) affects the prediction accuracy. The normalization of features and the use of k-fold cross-validation ensured that the models were trained on a balanced and representative dataset, reducing the likelihood of overfitting and enhancing the generalizability of the models. The results from these various configurations are essential for comprehending the influence of spatial context on snow depth prediction and for optimizing the models accordingly.

4.5.2. Random Forest

Random Forest utilizes the principle of ensemble learning by constructing multiple decision trees and combining their predictions to achieve a more accurate and stable output. This method excels in handling large datasets with numerous variables without overfitting (Reinan Assis Conceição et al., 2021). To optimize the model, hyperparameters such as the number of trees and the minimum number of samples per leaf are fine-tuned using a grid search approach. Various configurations—50, 100, 200, and 300 trees with leaf sizes of 1, 5, 10, and 20—are assessed using k-fold cross-validation, evaluating their performance based on R-squared values. The configuration with the highest average R-squared is selected, ensuring the model captures the maximum variance in the snow depth data.

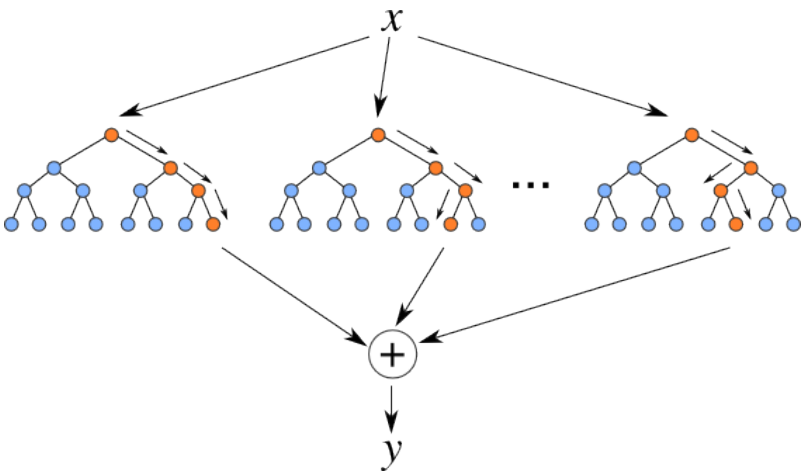


Figure 14: Random Forest Illustration

The Random Forest prediction is given by:

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N} T_i(x) \qquad \text{Eq. (5)}$$

Where $\hat{\mathbf{y}}$ is the predicted snow depth, **N** is the number of trees in the forest, and $\mathbf{T_{i(x)}}$ is the prediction from the $\mathbf{i^{th}}$ tree. Hyperparameters such as the number of trees $\mathbf{(n_{trees})}$ and the minimum leaf size **(min_leaf_size)** were tuned using cross-validation to find the best model.

4.5.3. Gradient Boosting

Gradient Boosting is another ensemble technique but differs from Random Forest by building trees sequentially, with each tree attempting to correct the errors of its predecessors (Musyimi et al., 2022). Gradient Boosting is known for its high predictive accuracy, particularly in regression tasks. By iteratively improving the model, it effectively minimizes the prediction error. Similar to Random Forests, Gradient Boosting is adept at identifying intricate, non-linear relationships between input features and the target variable. This capability is essential for accurately modeling snow depth variations based on coherence values. The tuning process for Gradient Boosting mirrors that of Random Forest, examining the same range of trees and leaf sizes. Gradient Boosting provides various parameters, such as the number of boosting stages, learning rate, and maximum tree depth, which can be tuned to optimize model performance (Tamiminia et al., 2022). This flexibility allows for fine-tuning the model to achieve the best possible predictions.

The Gradient Boosting prediction is given by:

$$\hat{y} = \sum_{i=1}^{N} \alpha_i T_i(x) \qquad \text{Eq. (6)}$$

Where $\hat{\mathbf{y}}$ is the prediction, **N** is the number of trees, $\boldsymbol{\alpha_i}$ is the weight of the $\mathbf{i_{th}}$ tree, and $\mathbf{T_{i(x)}}$ is the prediction of the $\mathbf{i_{th}}$ tree.

4.5.4. Support Vector Regression

Support Vector Regression (SVR) extends the capabilities of Support Vector Machines to regression problems, focusing on fitting data within a certain threshold of error, which is defined by the epsilon parameter (Mehravar et al., 2023). SVR is particularly adept at handling outliers and can model complex, non-linear relationships with high precision. For SVR, the critical hyperparameters—Box Constraint and Epsilon—are optimized through a rigorous search. A range of values for Box Constraint (0.1, 1, 10) and Epsilon (0.1, 0.5, 1) are tested in a nested loop, with each combination evaluated under k-fold cross-validation settings to assess their effectiveness using the R-squared metric. The combination providing the highest R-squared is deemed the best, reflecting the model's capacity to predict snow depths with maximum accuracy.

SVR constructs a function that has at most **ϵ** (epsilon) deviation from the actual observed targets y for all the training data, and at the same time, is as flat as possible:

$$\min_{w,b} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^)$$

Subject to:

$$y_i - \langle w, x_i\rangle - b \leq \epsilon + \xi_i$$

$$\langle w, x_i\rangle + b - y_i \leq \epsilon + \xi_i$$

$$\xi_i, \xi_i^* \geq 0$$

Eq. (7)

Where **w** is the weight vector, **b** is the bias, $\mathbf{x_i}$ are the training samples, $\mathbf{y_i}$ are the target values, $\boldsymbol{\xi_i}$, $\boldsymbol{\xi_i^*}$ are the slack variables, **C** is the penalty parameter, and **ϵ** (epsilon) is the margin of tolerance.

4.5.5. Hyperparameter Tuning

Hyperparameter tuning is essential for optimizing machine learning models to ensure effective performance on SAR datasets for predicting snow depths. Each model: Random Forest, Support Vector Regression (SVR), and Gradient Boosting, requires careful adjustment of specific hyperparameters to enhance their accuracy and generalization capabilities.

For the Random Forest model, crucial hyperparameters include the number of trees and the minimum leaf size. The number of trees impacts the model's robustness and its ability to generalize, as having more trees reduces variance through aggregation of multiple decision trees. The minimum leaf size determines the granularity of the splits; smaller leaf sizes enable the model to capture finer details in the data. Optimizing these parameters requires a balance between having enough trees to enhance accuracy without overfitting and ensuring the leaf size is small enough to detect subtle patterns without becoming overly sensitive to noise (Reinan Assis Conceição et al., 2021).

The performance of the Gradient Boosting model is significantly influenced by the number of trees, the learning rate, and the maximum depth of the trees. The number of trees indicates the number of boosting iterations, with more trees typically improving performance but increasing the risk of overfitting. The learning rate governs the contribution of each tree to the overall model, where lower learning rates necessitate more trees for better accuracy. The maximum depth of the trees determines the complexity of

each individual tree, with deeper trees capturing more intricate patterns but also raising the risk of overfitting (Tamiminia et al., 2022). Fine-tuning these parameters involves balancing these factors to achieve the optimal model performance.

For the SVR model, key hyperparameters include the choice of kernel function, the box constraint (C), and the epsilon parameter. The kernel function determines how the input data is transformed, with the radial basis function (rbf) kernel often preferred for its ability to handle non-linear relationships. The box constraint parameter manages the trade-off between a low training error and a low testing error, effectively balancing the model's complexity (Mehravar et al., 2023). The epsilon parameter defines a margin of tolerance where no penalty is applied to errors, aiding in managing the model's sensitivity to noise. Optimizing these hyperparameters involves selecting the appropriate kernel and adjusting C and epsilon to balance the model's fit and generalization ability (Mehravar et al., 2023). By systematically adjusting parameters for Random Forest, Gradient Boosting, and SVR models, it is possible to achieve high accuracy in snow depth estimation.

4.6. Model Evaluation

Model evaluations are performed using standard metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R-squared (R^2). These metrics provide insights into the accuracy, error magnitude, and explanatory power of each model. Cross-validation, which involves splitting the dataset into k subsets where each subset is used once as a test set while the others are used for training, ensures that the models are robust and generalize well to new data. This systematic approach to model selection and evaluation guarantees that the most effective predictive model is used, enhancing the reliability and applicability of snow depth estimations from satellite-derived SAR data.

4.6.1. Root Mean Square Error (RMSE)

RMSE is a widely used measure that quantifies the average magnitude of the prediction error, which is the differences between the values predicted by a model and the values observed from the environment that is being modelled (Zhang et al., 2018). It is calculated by taking the square root of the average of the square of all the error (Tamiminia et al., 2022). The formula is as follows:

The model performance was evaluated using the R-squared (R^2) score, calculated as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad \text{Eq. (8)}$$

Where $\mathbf{y_i}$ = actual values, $\hat{\mathbf{y}}_\mathbf{i}$ = predicted values, and n = total number of observations.

The lower the RMSE, the better a model is able to fit the observed data. A lower RMSE value indicates that the model's predictions are closely aligned with the actual data, reflecting higher accuracy and precision in the prediction of snow depths (Zhao et al., 2022).

4.6.2. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) is a critical metric in remote sensing for evaluating model performance. It calculates the average magnitude of prediction errors, without considering their direction. Specifically, MAE is derived by averaging the absolute differences between the predicted values and the actual observations, ensuring that all errors are equally weighted in the assessment of the model's predictive accuracy (Feng et al., 2023). The formula for MAE is given by:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad \text{Eq. (9)}$$

Where $\mathbf{y_i}$ and $\hat{\mathbf{y}}_\mathbf{i}$ are the actual and predicted values respectively, n is the number of observations.

MAE is predominantly valuable because it provides a direct indication of average error magnitudes and is less sensitive to outliers than RMSE. In the context of snow depth prediction, MAE helps assess the model's practical utility by quantifying the average prediction error in the same units as the measurement (Qiao et al., 2023).

4.6.3. R-squared (R^2)

R^2, or the coefficient of determination, is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model (Yang et al., 2021). The formula to calculate R^2 is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}i)^2}{\sum i = 1^n(y_i - \bar{y})^2} \qquad \text{Eq. (10)}$$

Where $\bar{\mathbf{y}}$ is the mean of the observed data $\mathbf{y_i}$, $\mathbf{y_i}$ and $\hat{\mathbf{y}}_\mathbf{i}$ are the actual and predicted values respectively, and n is the number of observations.

An R^2 of 1 indicates that the regression predictions perfectly fit the data. In snow depth prediction, a higher R^2 value would indicate that the model explains a large portion of the variance in observed snow depths, which is critical for understanding how well the model performs in terms of explaining the outcomes in relation to the variance observed in the actual data (Ahmadi et al., 2020). Together, these metrics provide a comprehensive framework for evaluating the performance of predictive models, guiding the selection of the most suitable model for accurately predicting snow depths. Each metric addresses different aspects of model accuracy and fit, ensuring a balanced and nuanced approach to model evaluation.

4.7. Model Evaluation Across ICESat-2 Beams

Following the initial performance assessment of the machine learning models, the methodology progresses to apply the best-performing model across all six beams of the ICESat-2 satellite, see Figure 15. This phase is designed to examine how model predictions fluctuate across different ICESat-2 beams as each beam has a different topography, and to identify which beam provides the most precise snow depth estimates.

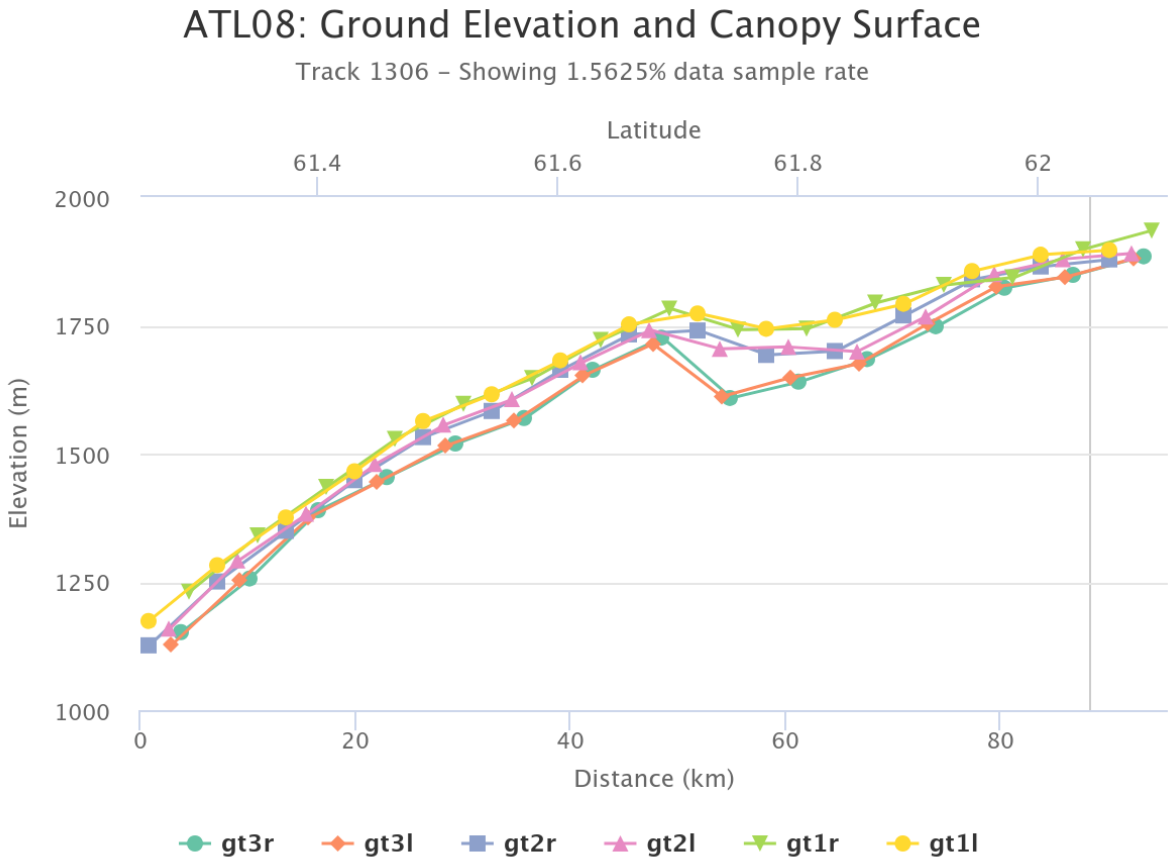


Figure 15: Six different beams from ICESat-2 (ATL08)

In addition to evaluating individual beams, a combined beam analysis was conducted by integrating data from multiple beams (gt1l, gt2l, gt3l). This combined dataset was used to train a Random Forest model, aiming to capture a broader range of spatial and environmental conditions. The combined analysis allowed for a comprehensive assessment of the model's ability to generalize across different regions within the study area. Performance metrics for the combined dataset were calculated, and scatter plots were generated to compare the results with those from individual beams. This approach helped to understand the overall predictive capability of the model across larger spatial extents and diverse conditions, highlighting the potential benefits or loss of increased training data and diversity against the challenges of managing more complex and varied datasets.

5. RESULTS AND DISCUSSION

This chapter presents the results of the snow depth estimation using the ICESat-2 and SAR datasets. The findings from the machine learning models are discussed in detail, highlighting the performance metrics and the comparative analysis between SAOCOM 1B and Sentinel-1 datasets. The results are further analysed to understand the strengths and limitations of each dataset and the overall effectiveness of the predictive models.

5.1. Snow Depth Measurements from ICESat-2 Data

The snow depth estimations derived from ICESat-2 photon data provide a detailed and high-resolution understanding of snow cover variability across the study area, as illustrated in figure 16.

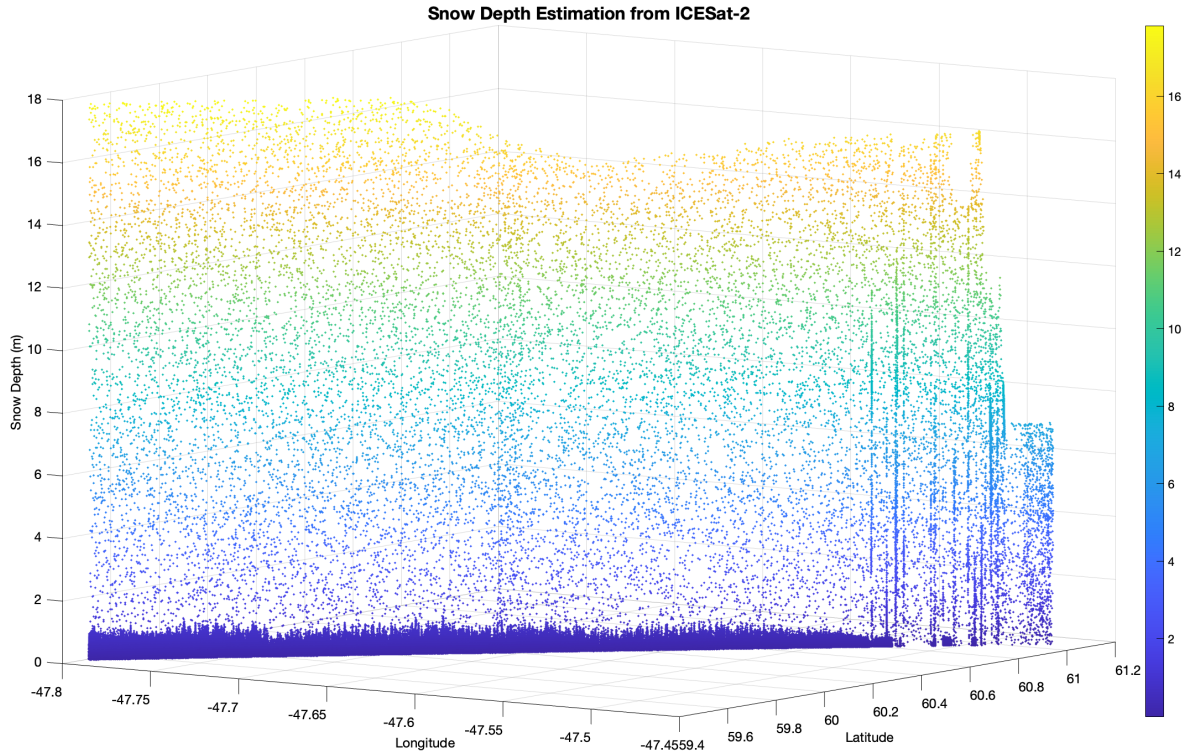


Figure 16: Snow depth measurements from ICESat-2

Snow depth values range from 0 meters (dark blue) to approximately 18 meters (yellow), indicating substantial spatial variability. This variation reflects different snow accumulation patterns influenced by factors such as topography, vegetation, and local climatic conditions.

Regions with high densities of lower snow depths (0-2 meters) are widespread, indicating that many areas within the study region have relatively shallow snow cover. These shallow snow depths could be due to lower elevations, open terrains, or areas experiencing higher temperatures and greater wind exposure, which lead to reduced snow accumulation. Additionally, the ATLAS instrument on ICESat-2, known for its precise surface elevation measurements, tends to capture backscatter signals more effectively from shallower snow depths. The instrument's sensitivity to snow surfaces results in a higher frequency of detected shallow snow depths, as deeper snow can attenuate the signal.

Conversely, regions with lower densities but significant snow depths (exceeding 10 meters) indicate substantial snow accumulation. These areas are likely located at higher elevations or within heavily vegetated regions, where snow accumulates and persists due to sheltering effects and lower temperatures. The extensive spatial coverage and uniform distribution of data points across latitude and longitude demonstrate the comprehensive reach achieved by ICESat-2. This ensures that the snow depth estimations are representative of the entire study area. The consistent patterns in snow depth measurements highlight the reliability and accuracy of the data collected by ICESat-2.

These high-resolution snow depth measurements are vital for training and validating machine learning models designed to predict snow depth from SAR coherence values. The detailed spatial variability captured in the ICESat-2 data enables robust model training, thereby improving the accuracy of snow depth predictions in remote sensing and GIS applications.

5.2. Model Performance Using Single Pixel and 3x3 Pixel Neighborhood Approaches

5.2.1. Results from SAOCOM

The performance of snow depth prediction models using single pixel, and 3x3 pixel neighborhood approaches were evaluated using SAOCOM 1B data. The table below summarizes the metrics for each model type and pixel approach:

Table 5: Performance Metrics for Machine Learning Models in Snow Depth Prediction for SAOCOM 1B

Pixel Base	Model	RMSE	MAE	R-squared
Single Pixel	Random Forest	1.0334	0.3299	0.6238
	Gradient Boosting	0.9596	0.4359	0.6756
	SVR	1.7032	0.3957	-0.0220
3x3 Pixel	Random Forest	0.8674	0.2613	0.7063
	Gradient Boosting	1.2284	0.4509	0.4109
	SVR	1.6130	0.3611	-0.0157

Random Forest Model

The Random Forest model with single pixel data achieved an RMSE of 1.0334, an MAE of 0.3299, and an R-squared value of 0.6238. The scatter plot (Figure 17 (a)) shows a moderate correlation between actual and predicted snow depths, but there is noticeable variability, especially at higher snow depths. Incorporating a 3x3 pixel neighborhood significantly improved the model's performance. The RMSE decreased to 0.8674, the MAE to 0.2613, and the R-squared value increased to 0.7063. The scatter plot (b) shows a tighter clustering of points around the 1:1 line, indicating better prediction accuracy. This improvement can be attributed to the additional spatial context provided by neighboring pixels, which helps the model capture local variability and structure of the snowpack more effectively.

Gradient Boosting Model

The Gradient Boosting model performed well with single pixel data, achieving an RMSE of 0.9596, an MAE of 0.4359, and the highest R-squared value (0.6756) among the single pixel models. The scatter plot (c) reveals a relatively good correlation but with some spread, particularly at higher snow depths. The performance declined with the 3x3 pixel approach, resulting in an RMSE of 1.2284, an MAE of 0.4509, and an R-squared of 0.4109. The scatter plot (d) shows a wider spread of predicted values, indicating reduced accuracy. This suggests that the Gradient Boosting model may be more sensitive to the specifics of individual pixel data rather than aggregated neighborhood data.

Support Vector Regression (SVR) Model

The SVR model exhibited poor performance with single pixel approach, with an RMSE of 1.7032, an MAE of 0.3957, and a negative R-squared value (-0.0220). The scatter plot (e) shows a large dispersion of points, particularly at lower snow depths, highlighting the model's inadequacy. The 3x3 pixel approach did not significantly improve the SVR model's performance, with an RMSE of 1.6130, an MAE of 0.3611, and an R-squared of -0.0157. The scatter plot (f) still shows a high level of dispersion, indicating that the SVR model struggles to capture snow depth variability.

The figure below presents scatter plots of actual vs. predicted snow depths for each model under both single pixel and 3x3 pixel neighborhood approaches for SAOCOM data.

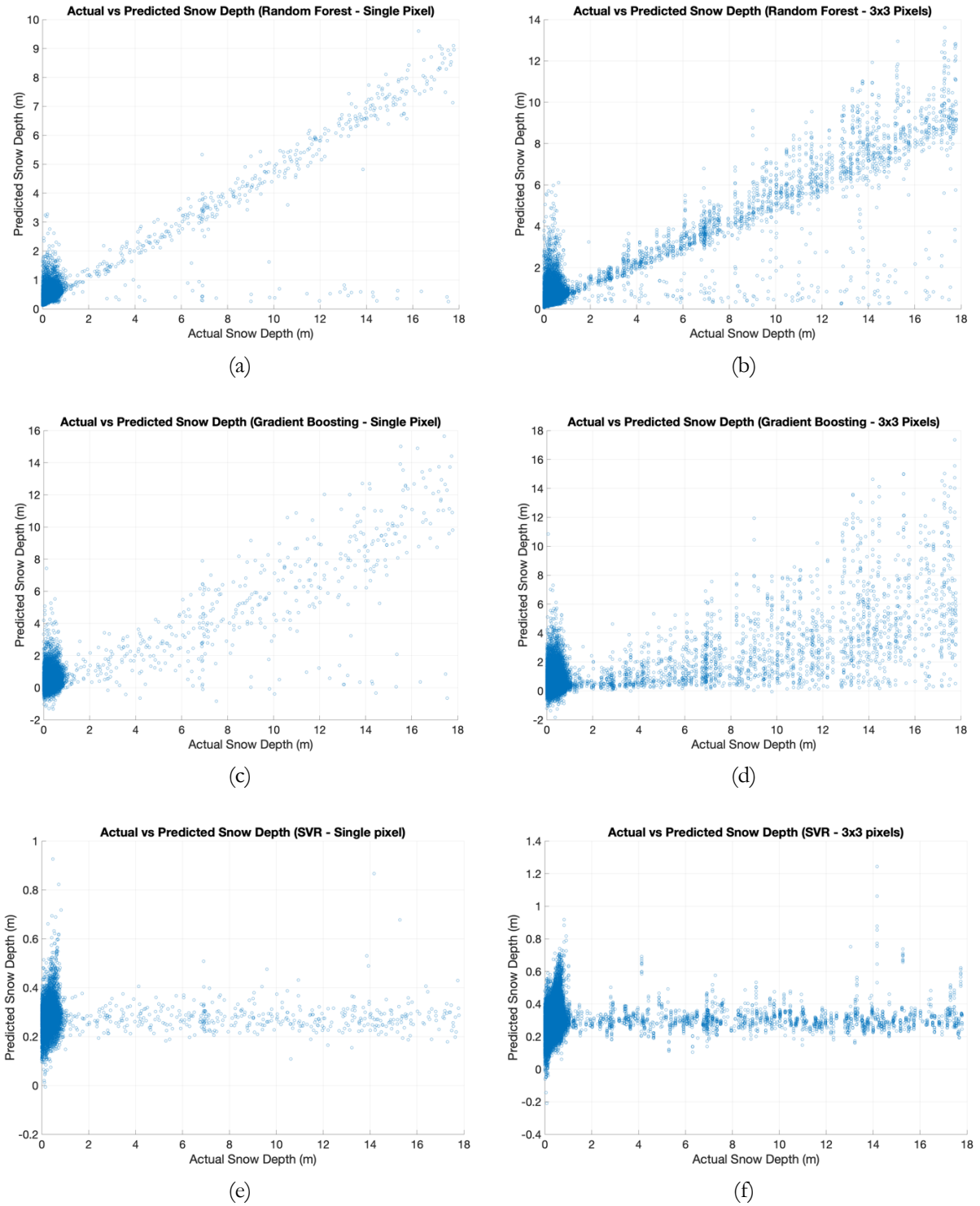


Figure 17: Scatter plots for snow depth predictions for single pixel and 3x3 pixel neighborhood methods for all three models using SAOCOM

The results underscore the importance of spatial context in snow depth prediction models. The 3x3 pixel neighborhood approach generally enhances the model's ability to capture local spatial variability and reduces prediction errors. This is most evident in the Random Forest model, where incorporating neighborhood information significantly improved performance metrics and prediction accuracy.

5.2.2. Results from Sentinel-1

The performance of snow depth prediction models using single pixel, and 3x3 pixel neighborhood approaches was evaluated using Sentinel-1 satellite data as well. The table below summarizes the metrics for each model type under both pixel-based approaches:

The figure below presents scatter plots of actual vs. predicted snow depths for each model under both single pixel and 3x3 pixel neighborhood approaches for Sentinel-1 data.

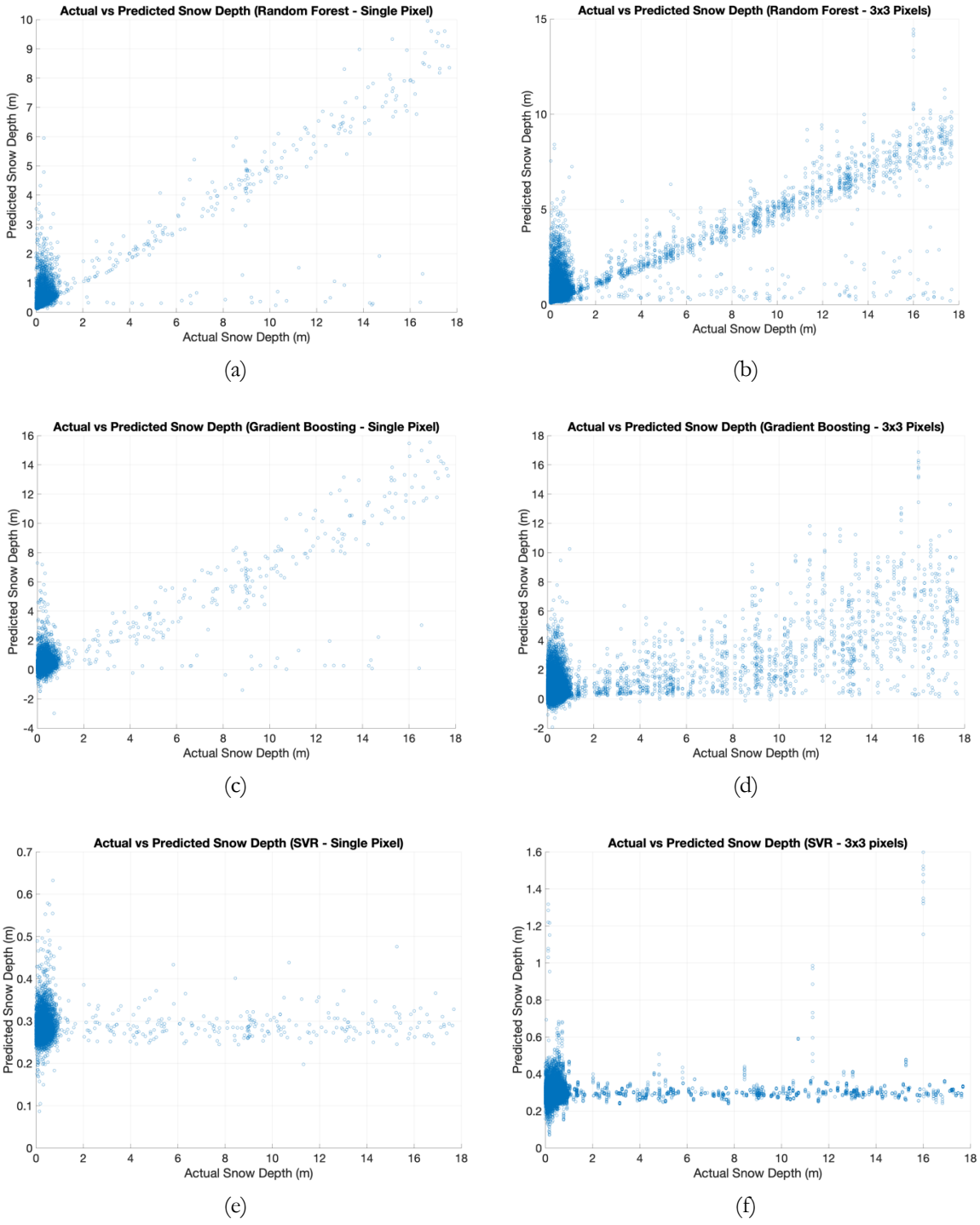


Figure 18: Scatter plots for snow depth predictions for single pixel and 3x3 pixel neighborhood methods for all three models using Sentinel-1

Table 6: Performance Metrics for Machine Learning Models in Snow Depth Prediction for Sentinel-1

Pixel Base	Model	RMSE	MAE	R-squared
Single Pixel	Random Forest	1.1357	0.3843	0.6054
	Gradient Boosting	0.9490	0.4186	0.7245
	SVR	1.8337	0.4478	-0.0286
3x3 Pixel	Random Forest	1.0573	0.3448	0.6314
	Gradient Boosting	1.3387	0.4989	0.4091
	SVR	1.7599	0.4149	-0.0213

Random Forest Model

The Random Forest model with single pixel data achieved an RMSE of 1.1357, an MAE of 0.3843, and an R-squared value of 0.6054. The scatter plot (Figure 18 (a)) for Sentinel-1 data shows a moderate correlation between actual and predicted snow depths, with some variability at higher snow depths. Incorporating a 3x3 pixel neighborhood improved the model's performance, with an RMSE of 1.0573, an MAE of 0.3448, and an R-squared value of 0.6314. The scatter plot (b) shows a tighter clustering of points around the 1:1 line, indicating better prediction accuracy. The additional spatial context from neighboring pixels likely enhances the model's ability to capture local variability in the snowpack.

Gradient Boosting Model

The Gradient Boosting model performed well with single pixel data, achieving an RMSE of 0.9490, an MAE of 0.4186, and the highest R-squared value (0.7245) among the single pixel models. The scatter plot (c) reveals a relatively good correlation but with some spread, especially at higher snow depths. The performance declined with the 3x3 pixel approach, resulting in an RMSE of 1.3387, an MAE of 0.4989, and an R-squared of 0.4091. The scatter plot (d) shows a wider spread of predicted values, indicating reduced accuracy. This suggests that the Gradient Boosting model may be better suited for single pixel data from Sentinel-1.

Support Vector Regression (SVR) Model

The SVR model exhibited poor performance with single pixel data, with an RMSE of 1.8337, an MAE of 0.4478, and a negative R-squared value (-0.0286). The scatter plot (e) shows a large dispersion of points, particularly at lower snow depths, highlighting the model's inadequacy. The 3x3 pixel approach did not significantly improve the SVR model's performance, with an RMSE of 1.7599, an MAE of 0.4149, and an R-squared of -0.0213. The scatter plot (f) still shows a high level of dispersion, indicating that the SVR model struggles to capture snow depth variability.

The results underscore the importance of spatial context in snow depth prediction models. The 3x3 pixel neighborhood approach generally enhances the model's ability to capture local spatial variability and reduces prediction errors. This is most evident in the Random Forest model, where incorporating neighborhood information significantly improved performance metrics and prediction accuracy. The consistent patterns observed in the scatter plots for Sentinel-1 data reinforce the quantitative metrics, providing visual validation of the models' performance. The Random Forest model, particularly with the 3x3 pixel neighborhood approach, demonstrates the most reliable performance, making it a preferred choice for snow depth prediction using Sentinel-1 data.

5.3. Comparative Analysis of SAOCOM and Sentinel-1 Results

The comparative analysis of model performance using SAOCOM and Sentinel-1 data highlights the variability in snow depth prediction accuracy based on different SAR data sources. The Random Forest model consistently outperformed other models in both datasets, particularly with the 3x3 pixel neighborhood approach.

Random Forest Model

The Random Forest model with the 3x3 pixel neighborhood approach achieves an RMSE of 0.8674, an MAE of 0.2613, and an R-squared value of 0.7063. These metrics indicate strong predictive accuracy and a high correlation between predicted and actual snow depths. The improvement over the single pixel approach demonstrates the importance of incorporating spatial context. For Sentinel-1 data, the Random Forest model with the 3x3 pixel neighborhood approach achieves an RMSE of 1.0573, an MAE of 0.3448, and an R-squared value of 0.6314. Although the performance is slightly lower than with SAOCOM data, it still shows significant improvement over the single pixel approach, highlighting the model's robustness.

Gradient Boosting Model

The Gradient Boosting model shows moderate performance with the 3x3 pixel neighborhood approach, achieving an RMSE of 1.2284, an MAE of 0.4509, and an R-squared value of 0.4109. The decline in performance compared to the single pixel approach suggests sensitivity to neighborhood data. Similarly, for Sentinel-1 data, the Gradient Boosting model with the 3x3 pixel approach achieves an RMSE of 1.3387, an MAE of 0.4989, and an R-squared value of 0.4091. The model struggles with higher snow depths, indicating limitations in capturing variability.

Support Vector Regression (SVR) Model

The SVR model performs poorly with both single pixel and 3x3 pixel neighborhood approaches, with an RMSE of 1.6130, an MAE of 0.3611, and an R-squared value of -0.0157. The negative R-squared value indicates poor predictive performance. The SVR model shows similar poor performance with Sentinel-1 data, achieving an RMSE of 1.7599, an MAE of 0.4149, and an R-squared value of -0.0213. This reinforces the model's inadequacy in capturing snow depth variability.

Overall, the Random Forest model consistently outperforms other models in both datasets, particularly with the 3x3 pixel neighborhood approach. This highlights the model's robustness in handling spatial variability and its effectiveness in snow depth prediction. The Gradient Boosting model, while performing well with single pixel data, shows reduced accuracy with 3x3 pixel approach, indicating potential sensitivity to spatial context. The SVR model's poor performance across all approaches suggests it may not be suitable for this application.

Several factors contribute to the differences in model performance between SAOCOM and Sentinel-1 data:

Frequency Band Differences: SAOCOM 1B operates in the L-band, which has a longer wavelength compared to the C-band of Sentinel-1. This allows the L-band to penetrate deeper into the snowpack, capturing more detailed subsurface information which is crucial for accurate snow depth estimation.

Signal-to-Noise Ratio: The L-band SAR data from SAOCOM 1B tends to have a higher signal-to-noise ratio when interacting with snow, leading to more reliable coherence measurements. In contrast, C-band signals are more susceptible to scattering and noise, which can degrade the accuracy of the coherence values used for snow depth prediction.

Coherence Levels: The absolute coherence values obtained from SAOCOM 1B data generally exhibited higher correlation with snow depth, as seen in the scatter plots, compared to Sentinel-1. Higher coherence levels indicate better preservation of phase information, which is crucial for the integrity of snow depth predictions.

Temporal Resolution: The temporal resolution and revisit frequency of Sentinel-1 might not align as well with the dynamics of snow cover changes, leading to less accurate temporal coherence measurements compared to SAOCOM 1B.

Beam Configuration and Ground Tracks: The differences in beam configuration and ground tracks of ICESat-2 can influence the quality of the data collected. Beams that capture more consistent snowpack characteristics tend to perform better in predictive models. Variability in snow depth data across different beams could also affect the performance metrics, particularly for the less effective Sentinel-1 dataset.

5.4. Model Comparison for Different ICESat-2 Beams

The performance of the Random Forest model in predicting snow depth was also evaluated using different ICESat-2 beams for both SAOCOM 1B and Sentinel-1 datasets. The beams analysed include gt1l, gt2l, gt3l, gt1r, gt2r, and gt3r.

5.4.1. Performance Analysis for SAOCOM 1B Data

The evaluation metrics for different ICESat-2 beams using the Random Forest model with SAOCOM 1B data are presented in Table 7.

Table 7: Performance metrics for all ICESat-2 Beams (SAOCOM 1B)

Beam	RMSE	MAE	R-squared
Random Forest Model Parameters: NumTrees = 300, MinLeafSize = 1			
gt1r	0.5817	0.2246	0.8212
gt2r	0.7150	0.2955	0.7513
gt3r	0.7283	0.3109	0.7719
gt1l	0.3456	0.1058	0.7657
gt2l	0.4335	0.1404	0.7407
gt3l	0.4387	0.1439	0.7746

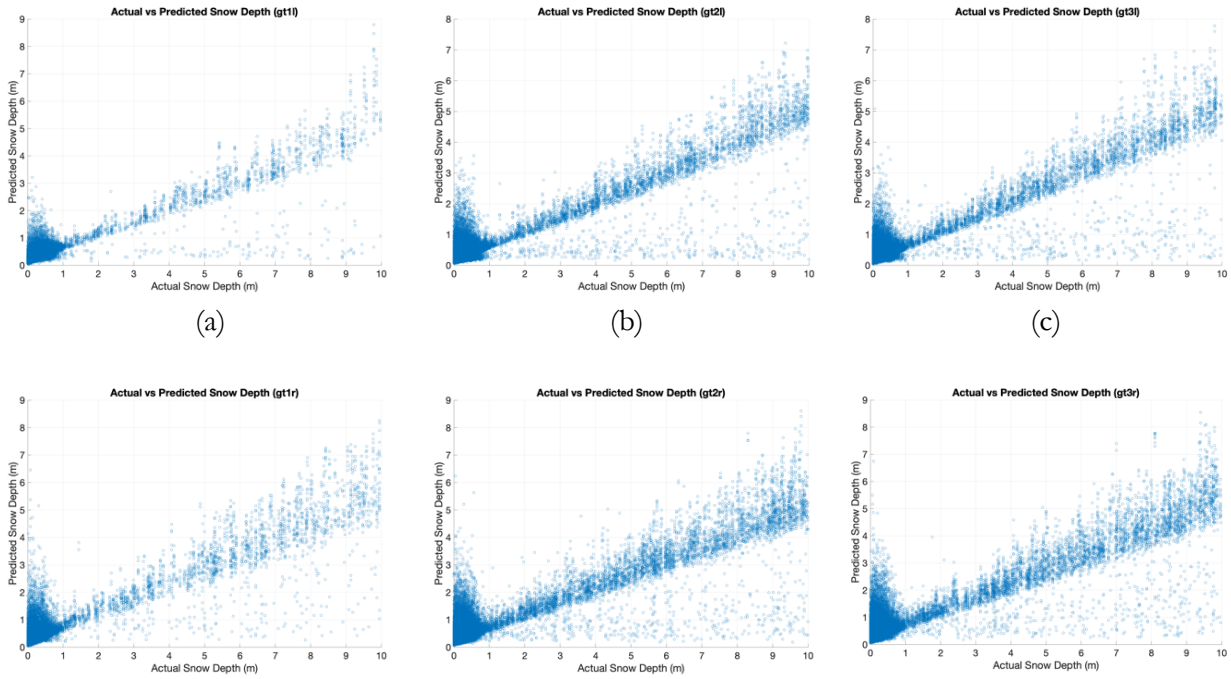


Figure 19: Predictions for different ICESat-2 Beams: SAOCOM 1B

The Random Forest model exhibited strong performance across all ICESat-2 beams when using SAOCOM data. Notably, beam gt1r showed the highest predictive accuracy with an RMSE of 0.5817, an MAE of 0.2246, and an R-squared value of 0.8212. This indicates a robust correlation between actual and predicted snow depths, with the scatter plot (Figure 19 (a)) showing a tight clustering of points around the 1:1 line.

For beam gt2r, the model's performance slightly declined, achieving an RMSE of 0.7150, an MAE of 0.2955, and an R-squared value of 0.7513. The scatter plot indicates a wider spread of points, suggesting variability in the predictions. Similarly, beam gt3r produced an RMSE of 0.7283, an MAE of 0.3109, and an R-squared value of 0.7719, with the scatter plot showing a comparable performance to gt2r.

Beam gt1l demonstrated the best overall performance among all beams, with an RMSE of 0.3456, an MAE of 0.1058, and an R-squared value of 0.7657. The scatter plot shows a very tight clustering of points, indicating excellent predictive accuracy. Beams gt2l and gt3l also performed well, with RMSEs of 0.4335 and 0.4387, MAEs of 0.1404 and 0.1439, and R-squared values of 0.7407 and 0.7746, respectively. The scatter plots for these beams illustrate consistent predictive performance.

The results indicate that the SAOCOM 1B data consistently yielded higher R-squared values and lower RMSE and MAE across all beams, particularly for the beams on the left side (gt1l, gt2l, gt3l). The best performance was observed for beam gt1r, with an R-squared value of 0.8212 and an RMSE of 0.5817. This demonstrates the effectiveness of the L-band data from SAOCOM 1B in capturing snow depth variations.

5.4.2. Performance Analysis for Sentinel-1 Data

The performance of the Random Forest model was assessed across different ICESat-2 beams using Sentinel-1 data. The table below summarizes the RMSE, MAE, and R-squared values for each beam:

Table 8: Performance metrics for all ICESat-2 Beams (Sentinel-1)

Beam	RMSE	MAE	R-squared
Random Forest Model Performance: NumTrees = 300, MinLeafSize = 1			
gt1r	0.8126	0.3462	0.6311
gt2r	0.9556	0.4318	0.6064
gt3r	0.9879	0.4540	0.6054
gt1l	0.4136	0.1552	0.6225
gt2l	0.5216	0.1799	0.6167
gt3l	0.5436	0.1877	0.6192

When using Sentinel-1 data, the Random Forest model showed lower performance metrics compared to SAOCOM data. For beam gt1r, the model achieved an RMSE of 0.8126, an MAE of 0.3462, and an R-squared value of 0.6311. The scatter plot reveals a moderate correlation between actual and predicted snow depths, with more variability than observed with SAOCOM data.

Beam gt2r exhibited further reduced performance, with an RMSE of 0.9556, an MAE of 0.4318, and an R-squared value of 0.6064. The scatter plot shows a wider spread of points, indicating greater prediction errors. Beam gt3r had similar metrics, with an RMSE of 0.9879, an MAE of 0.4540, and an R-squared value of 0.6054, as shown in the scatter plot. The model's performance for beam gt1l was the best among Sentinel-1 data beams, with an RMSE of 0.4136, an MAE of 0.1552, and an R-squared value of 0.6225. The scatter plot shows a tighter clustering of points, although still less accurate than SAOCOM data. Beams gt2l and gt3l had RMSEs of 0.5216 and 0.5436, MAEs of 0.1799 and 0.1877, and R-squared values of 0.6167 and 0.6192, respectively, as illustrated in scatter plots.

The scatter plots below illustrate the actual vs. predicted snow depths for each beam:

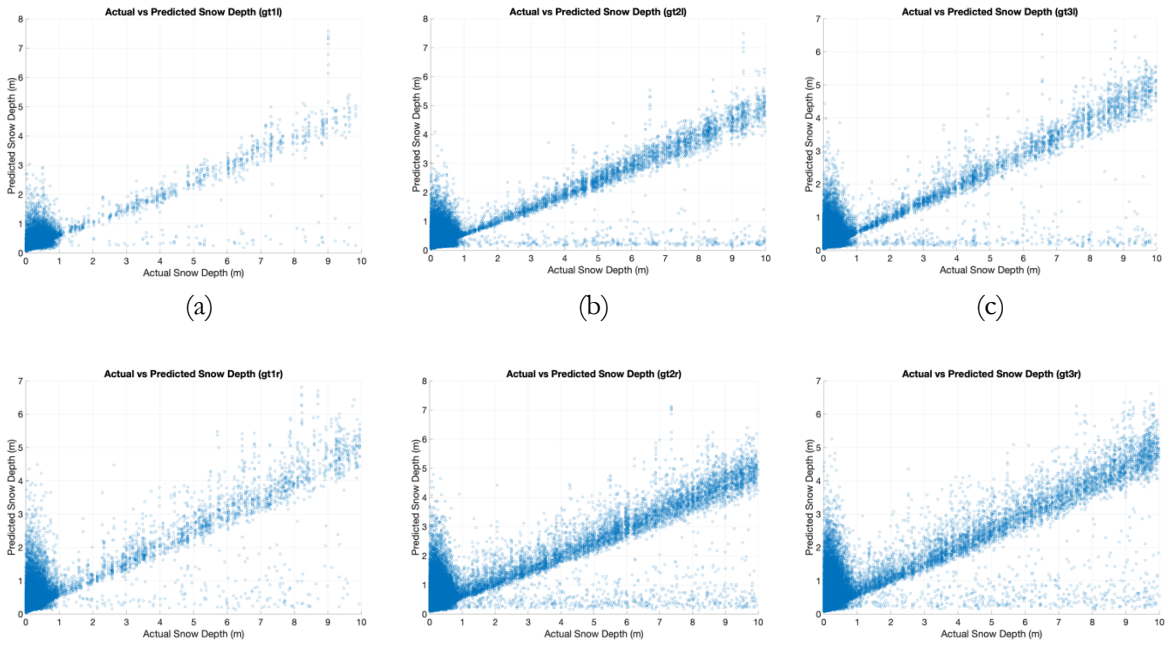


Figure 20: Predictions for different ICESat-2 Beams: Sentinel-1

Analyzing snow depth predictions for each ICESat-2 beam is essential due to the spatial separation of beams, which are approximately 3.3 km apart, and their coverage of different areas. This spatial separation results in variations in terrain, vegetation cover, and snow conditions, significantly influencing prediction accuracy. Evaluating model performance across different beams provides insights into how regional variations affect predictive accuracy.

For example, beams gt1r and gt1l demonstrated higher accuracy, likely due to traversing areas with more consistent snow cover or better data quality. Conversely, beams gt2r and gt3r showed lower accuracy, reflecting more heterogeneous or noisier conditions. Identifying specific beams with more reliable data helps optimize data collection and processing efforts, prioritizing high-accuracy beams in future studies.

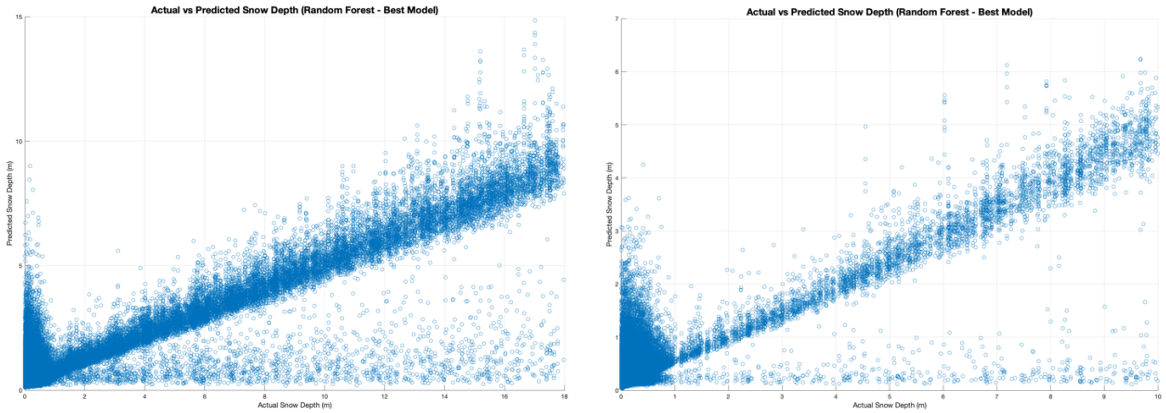
5.5. Combined Beam Analysis for SAOCOM and Sentinel-1

The analysis of snow depth predictions using combined beam data from ICESat-2 for both SAOCOM and Sentinel-1 datasets revealed several critical insights.

Table 9: Random Forest Model Performance Across Combined Beams

SAR Dataset	RMSE	MAE	R-squared
SAOCOM-1B (gt1l, gt2l, gt3l)	1.1949	0.4636	0.6645
Sentinel-1 (gt1l, gt2l, gt3l)	0.6070	0.2151	0.6188

For the combined SAOCOM-1B data, the Random Forest model achieved an RMSE of 1.1949, an MAE of 0.4636, and an R-squared value of 0.6645. In contrast, for the combined Sentinel-1 data, the model demonstrated better performance with an RMSE of 0.6070, an MAE of 0.2151, and an R-squared value of 0.6188. These results indicate a noticeable decline in model accuracy and predictive capability when beams are combined compared to individual beam analyses.



(a) (b)

Figure 21: Combined Beam Prediction for SAOCOM and Sentinel-1 Data

Several factors contribute to the decreased performance metrics in the combined beam analysis. One primary factor is the increased data variability. Each ICESat-2 beam covers distinct regions approximately 3.3 km apart, leading to a wide range of snowpack conditions, terrain features, and vegetation cover. This added heterogeneity complicates the model's task of identifying consistent patterns and relationships, resulting in reduced accuracy. This complexity is particularly evident in the SAOCOM dataset, where the

scatter plot shows a broader spread of points and greater variability compared to Sentinel-1 data. Furthermore, the risk of overfitting is heightened with the increased number of training points from combined beams. Overfitting occurs when the model learns specific details and noise within the training data, diminishing its ability to generalize to new, unseen data. The scatter plots for combined SAOCOM data indicate higher variability, suggesting the model is capturing noise rather than true underlying patterns.

Another crucial factor influencing model performance is the difference in signal-to-noise ratios and coherence levels between the datasets. SAOCOM's L-band data penetrates deeper into the snowpack and generally exhibits higher coherence values, which are beneficial for capturing detailed subsurface information. However, combining beams with varying coherence levels and noise characteristics introduces inconsistencies that challenge the model's ability to maintain high accuracy. Conversely, Sentinel-1's C-band data, while more susceptible to scattering and noise, provides more uniform characteristics across beams, potentially explaining its relatively better performance in the combined analysis. The increase in training data volume from combining beams can also introduce more noise and less relevant features, complicating the model's training process and highlighting the need for advanced preprocessing and regularization techniques to improve robustness and accuracy.

5.6. Feature Importance & Final Predicted Snow Depth Using the Best Model

The map generated from SAOCOM 1B data (Figure 23) exhibits detailed snow depth distribution across the study area. Among all beams, gt1l demonstrated the best overall performance, achieving an RMSE of 0.3456, an MAE of 0.1058, and an R-squared value of 0.7657. This highlights the effectiveness of using specific beams for snow depth prediction. The feature importance analysis from the Random Forest model underscores the significance of using multiple polarization channels. For SAOCOM 1B, all polarizations (HH, HV, VH, VV) contributed substantially to snow depth prediction, with HV polarization showing slightly higher importance, as depicted in Figure 22. This indicates that integrating various polarization data can enhance the model's predictive accuracy, as different polarizations provide complementary information about the snowpack's structure and properties.

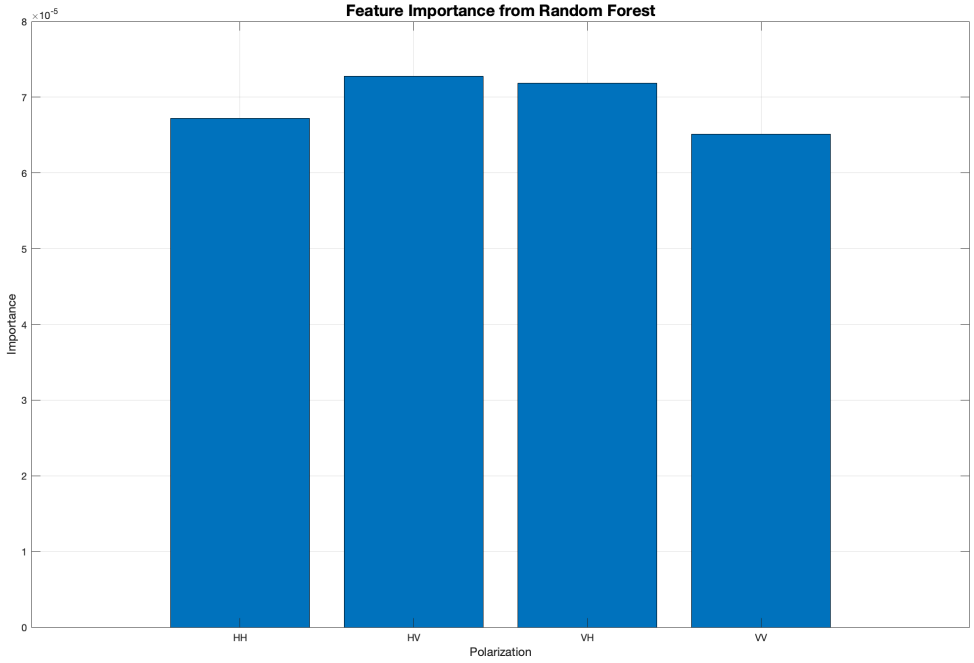
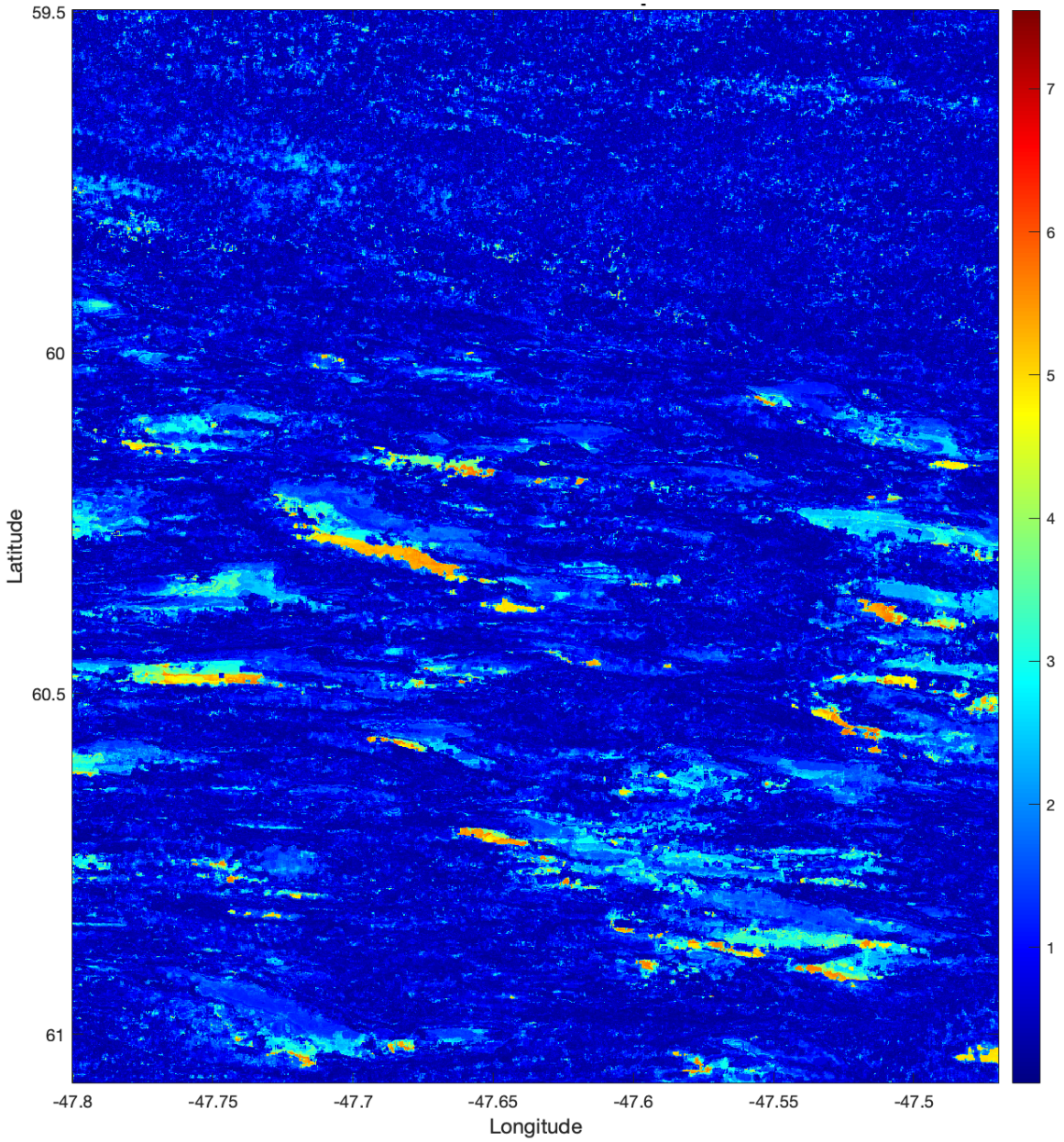


Figure 22: Feature Importance

The reduced penetration depth and higher noise levels of C-band SAR data make it challenging to obtain accurate coherence measurements, resulting in less reliable snow depth predictions. The feature importance analysis further highlights the benefit of using multi-polarization data, which enhances the model's ability to discern different snowpack characteristics. The results clearly demonstrate the superiority of the SAOCOM 1B dataset over Sentinel-1 for snow depth estimation. The higher resolution and better penetration capabilities of L-band SAR data in SAOCOM 1B result in more reliable and accurate coherence measurements, crucial for predicting snow depth. The Random Forest model effectively captures the variability in the snowpack, providing high-fidelity predictions. In contrast, the Sentinel-1 dataset's lower performance can be linked to the inherent limitations of C-band SAR data.

Figure 23: Predicted snow depth map from best model



Overall, this study underscores the importance of selecting appropriate SAR datasets and leveraging robust machine learning models with ICESat-2 LiDAR measurements for accurate snow depth estimation. The findings provide valuable insights for future research and practical applications in snow monitoring and climate studies, emphasizing the use of SAR data in snow-covered regions.

6. CONCLUSION

6.1. Conclusion

This research demonstrates the effectiveness of utilizing ICESat-2 LIDAR data in conjunction with Interferometric Coherence values from SAOCOM 1B and Sentinel-1 to enhance snow depth estimation through various machine learning techniques. The comprehensive methodological approach adopted in this study, including preprocessing, coregistration, coherence analysis, and machine learning-based prediction models, offers a robust framework for utilizing satellite data in snow monitoring.

The findings reveal that SAOCOM 1B's L-band SAR data significantly outperforms Sentinel-1's C-band SAR data in predicting snow depth. The deeper penetration of L-band SAR facilitates more accurate and reliable coherence measurements, which are essential for precise snow depth predictions. The Random Forest model was identified as the best-performing machine learning model, achieving a high R-squared value of 0.8212 with the gt1r beam from ICESat-2. This highlights the potential of using LIDAR and SAR data together for improved snow monitoring.

A significant portion of this research focused on analyzing the performance of different ICESat-2 beams. Each beam, covering distinct regions approximately 3.3 km apart, provided unique insights into snowpack characteristics. The gt1l beam demonstrated the best overall performance among all beams, with an RMSE of 0.3456, an MAE of 0.1058, and an R-squared value of 0.7657. This detailed beam-level analysis highlighted the importance of considering regional differences and beam-specific characteristics in improving model accuracy. The combined beam analysis for both SAOCOM and Sentinel-1 datasets showed a decrease in performance metrics, likely due to increased data variability and potential overfitting. The larger dataset introduced by combining beams covered a wider range of snowpack conditions, adding complexity for the model to handle. Nevertheless, Sentinel-1 data demonstrated relatively better performance in the combined analysis, indicating that uniform data characteristics play a critical role in maintaining model accuracy.

The feature importance analysis emphasized the significance of multi-polarization data in enhancing prediction accuracy. All four polarization channels (HH, HV, VH, VV) contributed significantly to the model's performance, highlighting the necessity of using diverse polarization data to capture the full spectrum of snowpack properties. Despite the lower performance of Sentinel-1 data, the study identified areas for improvement, particularly in overcoming the challenges posed by C-band SAR's limited penetration and higher noise levels.

How can the extensive spatial coverage of SAR data from SAOCOM 1B and Sentinel-1 be utilized to fill the spatial coverage gaps in ICESat-2 data for snow depth estimation?

The extensive spatial coverage of SAR data, particularly from SAOCOM 1B, can be effectively utilized to fill the spatial gaps in ICESat-2 data by providing continuous and comprehensive snow depth estimates. Despite the lack of strong correlation between coherence values and snow depth, robust machine learning models can predict snow depth over larger areas, addressing the limitations of ICESat-2's narrow swath width and sparse ground tracks.

Which machine learning model, Random Forest, Gradient Boosting, or Support Vector Regression provides the most accurate predictions of snow depth from SAR absolute coherence data?

The Random Forest model provided the most accurate predictions of snow depth from SAR absolute interferometric coherence data. It consistently outperformed Gradient Boosting and Support Vector Regression models, as evidenced by lower RMSE, lower MAE, and higher R-squared values. The model's ability to handle complex interactions and its robustness in capturing snow depth variations contributed to its superior performance.

What are the comparative strengths and limitations of SAOCOM 1B's L-band SAR data and Sentinel-1's C-band SAR data in predicting snow depth?

SAOCOM 1B's L-band SAR data demonstrated significant strengths in predicting snow depth due to its greater penetration depth and higher signal-to-noise ratio, which provided more reliable coherence measurements. In contrast, Sentinel-1's C-band SAR data, while offering higher spatial resolution, exhibited lower predictive accuracy due to its lesser penetration depth and susceptibility to scattering and noise. This highlights the importance of selecting the appropriate SAR frequency band for specific snow depth estimation applications.

6.2. Limitations

Despite the promising results, this study has several limitations. The spatial and temporal coverage of ICESat-2 data, limited by its narrow swath width and sparse ground tracks, can affect the generalizability of the results. Although SAR data helps mitigate these gaps, the temporal alignment between datasets can pose challenges. The calculation of absolute coherence from SAR data is influenced by various factors, including environmental conditions and surface properties, which can introduce uncertainties in snow depth estimation. Additionally, the performance of machine learning models is greatly influenced by the quality and quantity of the training data. In this study, ICESat-2 data served as a high-accuracy benchmark, offering reliable reference points. However, its relatively sparse coverage could limit the model's applicability in regions with varying snow characteristics.

6.3. Future Work and Recommendations

Exploring additional machine learning models and ensemble learning techniques could enhance the robustness of snow depth predictions. Expanding the study to other snow-covered regions with diverse climatic and topographic conditions is essential to validate and refine the proposed methodology. Integrating additional remote sensing data sources, such as optical and thermal imagery, could provide complementary information to further improve snow depth estimation. Finally, developing real-time snow depth monitoring systems using SAR and LIDAR data would offer significant benefits for water resource management, avalanche risk prediction, and climate change studies. By addressing these areas, future research can build on the findings of this study and contribute to more accurate and reliable snow depth estimation methods.

7. DATA MANGEMENT & ETHICAL CONSIDERATIONS

7.1. Data Management

Effective data management was crucial for this study, involving extensive datasets from ICESat-2, SAOCOM 1B, and Sentinel-1. The data were systematically organized and stored to ensure easy access and efficient analysis throughout the research. To address potential data quality issues, comprehensive preprocessing steps such as filtering, radiometric calibration, and geometric correction were implemented. These steps were essential to maintain the accuracy and reliability of the data. Additionally, redundant data storage and regular backups on secure external storage devices were carried out to prevent data loss, ensuring the continuity and integrity of the research process.

7.2. Ethical Considerations

This research was conducted with a strong commitment to ethical guidelines and data usage agreements. The SAOCOM-1B dataset was obtained with explicit consent from the Comisión Nacional de Actividades Espaciales (CONAE) and used exclusively for the purposes of this study, as facilitated by the Faculty of Geo-information Science and Earth Observation (ITC) at the University of Twente. Publicly accessible datasets, such as ICESat-2 from NASA and Sentinel-1 from the European Space Agency (ESA), were also utilized responsibly, adhering strictly to research objectives. The research findings were presented with integrity, ensuring accuracy and transparency.

7.3. Use of AI

To enhance the quality of this thesis, AI tools were utilized for some tasks. OpenAI's ChatGPT was used to help understand complex concepts, which were then extensively studied and validated using reputable academic sources. Grammarly was employed to proofread the text, ensuring grammatical correctness. Despite the support from these AI tools, the author carefully reviewed and revised all content, ensuring the final document's accuracy and scholarly integrity. This approach balanced the efficiency of AI with rigorous academic standards, producing a polished and reliable thesis.

LIST OF REFERENCES

Ahmadi, K., Kalantar, B., Saeidi, V., Harandi, E. K. G., Janizadeh, S., & Ueda, N. (n.d.). *remote sensing Comparison of Machine Learning Methods for Mapping the Stand Characteristics of Temperate Forests Using Multi-Spectral Sentinel-2 Data. 12*, 3019. https://doi.org/10.3390/rs12183019

Askne, J. I. H., Soja, M. J., & Ulander, L. M. H. (2017). Biomass estimation in a boreal forest from TanDEM-X data, lidar DTM, and the interferometric water cloud model. *Remote Sensing of Environment*, *196*, 265–278. https://doi.org/10.1016/j.rse.2017.05.010

Awasthi, S., Kumar, S., Thakur, P. K., Jain, K., Kumar, A., & Snehmani. (2021). Snow depth retrieval in North-Western Himalayan region using pursuit-monostatic TanDEM-X datasets applying polarimetric synthetic aperture radar interferometry based inversion Modelling. *International Journal of Remote Sensing*, *42*(8), 2872–2897. https://doi.org/10.1080/01431161.2020.1862439

Benedikter, A., Rodriguez-Cassola, M., Betancourt-Payan, F., Krieger, G., & Moreira, A. (2022). Autofocus-Based Estimation of Penetration Depth and Permittivity of Ice Volumes and Snow Using Single SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, *60*. https://doi.org/10.1109/TGRS.2021.3135026

Bernard, E. (2017). INVESTIGATING SNOWPACK VOLUMES AND ICING DYNAMICS IN THE MORAINE OF AN ARCTIC CATCHMENT USING UAV PHOTOGRAMMETRY. *The Photogrammetric Record*, *32*(160), 497–512. https://doi.org/10.1111/phor.12217

Brolly, M., Simard, M., Tang, H., Dubayah, R. O., Fisk, J. P., & Tang, H. (2016). A Lidar-Radar Framework to Assess the Impact of Vertical Forest Structure on Interferometric Coherence. *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING*, *9*(12). https://doi.org/10.1109/JSTARS.2016.2527360

Cazcarra-Bes, V., Pardini, M., Tello, M., & Papathanassiou, K. P. (2020). Comparison of Tomographic SAR Reflectivity Reconstruction Algorithms for Forest Applications at L-band. *IEEE Transactions on Geoscience and Remote Sensing*, *58*(1), 147–164. https://doi.org/10.1109/TGRS.2019.2934347

Dahhani, S., Raji, M., Hakdaoui, M., & Lhissou, R. (2022). *Land Cover Mapping Using Sentinel-1 Time-Series Data and Machine-Learning Classifiers in Agricultural Sub-Saharan Landscape.* https://doi.org/10.3390/rs15010065

Enderlin, E. M., Elkin, C. M., Gendreau, M., Marshall, H. P., O'Neel, S., McNeil, C., Florentine, C., & Sass, L. (2022). Uncertainty of ICESat-2 ATL06- and ATL08-derived snow depths for glacierized and vegetated mountain regions. *Remote Sensing of Environment*, *283*. https://doi.org/10.1016/j.rse.2022.113307

Feng, T., Hao, X., Wang, J., Luo, S., Huang, G., Li, H., & Zhao, Q. (2023). Applicability of alpine snow depth estimation based on multitemporal UAV-LiDAR data: A case study in the Maxian Mountains, Northwest China. *Journal of Hydrology*, *617*. https://doi.org/10.1016/j.jhydrol.2022.129006

Fischer, G., Jager, M., Papathanassiou, K. P., & Hajnsek, I. (2019). Modeling the Vertical Backscattering Distribution in the Percolation Zone of the Greenland Ice Sheet with SAR Tomography. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(11), 4389–4405. https://doi.org/10.1109/JSTARS.2019.2951026

Frey, O., Werner, C. L., & Wiesmann, A. (2015). *Tomographic Profiling of the Structure of a Snow Pack at X-/Ku-Band using SnowScat in SAR Mode.*

Garg, V., Thakur, P. K., Rajak, D. R., Aggarwal, S. P., & Kumar, P. (2022). Spatio-temporal changes in radar zones and ELA estimation of glaciers in NyÅlesund using Sentinel-1 SAR. *Polar Science*, *31*. https://doi.org/10.1016/j.polar.2021.100786

Immerzeel, W. W., Kraaijenbrink, P. D. A., Shea, J. M., Shrestha, A. B., Pellicciotti, F., Bierkens, M. F. P., & de Jong, S. M. (2014). High-resolution monitoring of Himalayan glacier dynamics using unmanned aerial vehicles. *Remote Sensing of Environment*, *150*, 93–103. https://doi.org/10.1016/j.rse.2014.04.025

Jouvet, G., Weidmann, Y., van Dongen, E., Lüthi, M. P., Vieli, A., & Ryan, J. C. (2019). High-Endurance UAV for Monitoring Calving Glaciers: Application to the Inglefield Bredning and Eqip Sermia, Greenland. *Frontiers in Earth Science*, *7*. https://doi.org/10.3389/feart.2019.00206

Kumar, V., & Venkataraman, G. (2011). SAR interferometric coherence analysis for snow cover mapping in the western Himalayan region. *International Journal of Digital Earth*, *4*(1), 78–90. https://doi.org/10.1080/17538940903521591

Kwok, R., Kacimi, S., Webster, M. A., Kurtz, N. T., & Petty, A. A. (n.d.). *Arctic Snow Depth and Sea Ice Thickness From ICESat-2 and CryoSat-2 Freeboards: A First Examination*. https://doi.org/10.1029/2019JC016008

Leinss, S., Parrella, G., & Hajnsek, I. (2014). *Snow Height Determination by Polarimetric Phase Differences in X-Band SAR Data*. https://doi.org/10.1109/JSTARS.2014.2323199

Li, L., Chen, H., & Guan, L. (2021). *remote sensing Retrieval of Snow Depth on Arctic Sea Ice from the FY3B/MWRI*. https://doi.org/10.3390/rs13081457

Lu, X., Hu, Y., Zeng, X., Stamnes, S. A., Neuman, T. A., Kurtz, N. T., Yang, Y., Zhai, P.-W., Gao, M., Sun, W., Xu, K., Liu, Z., Omar, A. H., Baize, R. R., Rogers, L. J., Mitchell, B. O., Stamnes, K., Huang, Y., Chen, N., … P-w, Z. (2022). *Deriving Snow Depth From ICESat-2 Lidar Multiple Scattering Measurements: Uncertainty Analyses*. *3*, 1. https://doi.org/10.3389/frsen.2022.891481

Majumdar, S., Thakur Ling Chang Advisor, P. K., & Kumar, S. (2019). *SNOW DEPTH AND SWE ESTIMATION USING SPACEBORNE POLARIMETRIC AND INTERFEROMETRIC SYNTHETIC APERTURE RADAR]*.

Mehravar, S., Razavi-Termeh, S. V., Moghimi, A., Ranjgar, B., Foroughnia, F., & Amani, M. (2023). Flood susceptibility mapping using multi-temporal SAR imagery and novel integration of nature-inspired algorithms into support vector regression. *Journal of Hydrology*, *617*. https://doi.org/10.1016/j.jhydrol.2023.129100

Musyimi, P. K., Sahbeni, G., Timár, G., Weidinger, T., & Székely, B. (2022). Actual Evapotranspiration Estimation Using Sentinel-1 SAR and Sentinel-3 SLSTR Data Combined with a Gradient Boosting Machine Model in Busia County, Western Kenya. *Atmosphere*, *13*(11). https://doi.org/10.3390/atmos13111927

Narine, L. L., Popescu, S. C., & Malambo, L. (2020). Using ICESat-2 to estimate and map forest aboveground biomass: A first example. *Remote Sensing*, *12*(11). https://doi.org/10.3390/rs12111824

Olesk, A., Praks, J., Antropov, O., Zalite, K., Arumäe, T., & Voormansik, K. (2016). Interferometric SAR coherence models for Characterization of hemiboreal forests using TanDEM-X dssata. *Remote Sensing*, *8*(9). https://doi.org/10.3390/rs8090700

Pardini, M., Armston, J., Qi, W., Lee, S. K., Tello, M., Cazcarra Bes, V., Choi, C., Papathanassiou, K. P., Dubayah, R. O., & Fatoyinbo, L. E. (2019). Early Lessons on Combining Lidar and Multi-baseline SAR Measurements for Forest Structure Characterization. In *Surveys in Geophysics* (Vol. 40, Issue 4, pp. 803–837). Springer Netherlands. https://doi.org/10.1007/s10712-019-09553-9

Patil, A., Singh, G., & Rüdiger, C. (n.d.). *remote sensing Retrieval of Snow Depth and Snow Water Equivalent Using Dual Polarization SAR Data*. https://doi.org/10.3390/rs12071183

Phinzi, K., Abriha, D., & Szabó, S. (2021). *remote sensing Classification Efficacy Using K-Fold Cross-Validation and Bootstrapping Resampling Techniques on the Example of Mapping Complex Gully Systems*. https://doi.org/10.3390/rs13152980

Qiao, H., Zhang, P., Li, Z., Huang, L., Zhao, C., Gao, S., Liu, C., Wu, Z., Liang, S., Zhou, J., Sun, W., & Wang, L. (2023). Snow profile reconstruction from tomographic UAV SAR. *International Journal of Applied Earth Observation and Geoinformation*, *118*. https://doi.org/10.1016/j.jag.2023.103291

Reinan Assis Conceição, M., Felipe Ferreira de Mendonça, L., Alessandre Domingos Lentini, C., Telles da Cunha Lima, A., Marques Lopes, J., Nogueira de Vasconcelos, R., Biazati Gouveia, M., José Porsani, M., & Salehi, B. (2021). *remote sensing SAR Oil Spill Detection System through Random Forest Classifiers*. https://doi.org/10.3390/rs13112044

Rekioua, B., Davy, M., Ferro-Famil, L., & Tebaldini, S. (2017). Snowpack permittivity profile retrieval from tomographic SAR data. *Comptes Rendus Physique*, *18*(1), 57–65. https://doi.org/10.1016/j.crhy.2015.12.016

Rott, H., Scheiblauer, S., Wuite, J., Krieger, L., Floricioiu, D., Rizzoli, P., Libert, L., & Nagler, T. (2021). Penetration of interferometric radar signals in Antarctic snow. *Cryosphere*, *15*(9), 4399–4419. https://doi.org/10.5194/tc-15-4399-2021

Smith, B., Fricker, H. A., Holschuh, N., Gardner, A. S., Adusumilli, S., Brunt, K. M., Csatho, B., Harbeck, K., Huth, A., Neumann, T., Nilsson, J., & Siegfried, M. R. (2019). Land ice height-retrieval algorithm for NASA's ICESat-2 photon-counting laser altimeter. *Remote Sensing of Environment*, *233*. https://doi.org/10.1016/j.rse.2019.111352

Tamiminia, H., Salehi, B., Mahdianpari, M., Beier, C. M., & Johnson, L. (2022). Evaluating pixel-based and object-based approaches for forest above-ground biomass estimation using a combination of optical, sar, and an extreme gradient boosting model. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *5*(3), 485–492. https://doi.org/10.5194/isprs-Annals-V-3-2022-485-2022

Tebaldini, S., Nagler, T., Rott, H., & Heilig, A. (2016). Imaging the Internal Structure of an Alpine Glacier via L-Band Airborne SAR Tomography. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(12), 7197–7209. https://doi.org/10.1109/TGRS.2016.2597361

Tsang, L., Durand, M., Derksen, C., Barros, A. P., Kang, D. H., Lievens, H., Marshall, H. P., Zhu, J., Johnson, J., King, J., Lemmetyinen, J., Sandells, M., Rutter, N., Siqueira, P., Nolin, A., Osmanoglu, B., Vuyovich, C., Kim, E., Taylor, D., … Xu, X. (2022). Review article: Global monitoring of snow water equivalent using high-frequency radar remote sensing. In *Cryosphere* (Vol. 16, Issue 9, pp. 3531–3573). Copernicus Publications. https://doi.org/10.5194/tc-16-3531-2022

Yang, J. W., Jiang, L. M., Lemmetyinen, J., Pan, J. M., Luojus, K., & Takala, M. (2021). Improving snow depth estimation by coupling HUT-optimized effective snow grain size parameters with the random forest approach. *Remote Sensing of Environment*, *264*. https://doi.org/10.1016/j.rse.2021.112630

Zhang, H., Wang, C., Zhu, J., Fu, H., Xie, Q., & Shen, P. (2018). *Forest Above-Ground Biomass Estimation Using Single-Baseline Polarization Coherence Tomography with P-Band PolInSAR Data.* https://doi.org/10.3390/f9040163

Zhao, Y., Wu, B., Shu, S., Yang, L., Wu, J., & Yu, B. (2022). Evaluation of ICESat-2 ATL03/08 Surface Heights in Urban Environments Using Airborne LiDAR Point Cloud Data. *IEEE Geoscience and Remote Sensing Letters*, *19*. https://doi.org/10.1109/LGRS.2021.3127540