IDEA League

Master of Science in Applied Geophysics

Research Thesis

Radar terrain angle correction for glacier facies classification

UiT The Arctic University of Norway

Marion Dugué

July 29, 2024

Radar terrain angle correction for glacier facies classification

UiT The Arctic University of Norway

Master of Science Thesis

for the degree of Master of Science in Applied Geophysics

by

Marion Dugué

July 29, 2024

IDEA LEAGUE
JOINT MASTER'S IN APPLIED GEOPHYSICS

Delft University of Technology, The Netherlands
ETH Zürich, Switzerland
RWTH Aachen, Germany

Dated: *July 29, 2024*

Supervisor(s):

Anthony Doulgeris

Deyan Draganov

Committee Members:

Anthony Doulgeris

Deyan Draganov

Florian Wagner

Eric Verschuur

Abstract

Glaciers' mass balance and melting patterns can be monitored through the study of their facies. Whilst using Synthetic aperture radar (SAR) remote sensing data facilitates glaciology observations as it can be used under all weather conditions, the systemic backscatter intensity decay due to incidence angle (IA) variation makes classification even more challenging on such complex terrain. We investigate the classification accuracy of glacier facies using SAR data through a supervised learning algorithm that incorporates class-dependent local incidence angle correction. Focusing on the Holtedahlfonna and Kongsvegen glacier complexes in northeast Svalbard, we pre-process SAR data from Sentinel-1 using the SNAP toolbox. We then compare three Bayesian classifiers: one without any IA correction, one with a common IA slope correction, and the third incorporating a class-dependent IA slope correction. Our results show that per-class IA slope correction on training regions improves the models by around 30% compared to the naive one and around 10% from the common IA slope correction. When tested on the glaciers, their firn line could be mapped from 2017 to 2023 and a general retreat of 400-500 m is observed, changing to 3-4 km in some regions of Holtedalhfonna. However, when looking at regions of lower altitudes, regions with crevasses are largely misclassified. To aid crevasse classification, we finish this study by providing some insights on potential texture features, using either standard deviation or spatial Fourier transforms. In all, this work explores the extent to which class-dependent IA corrections should be included in SAR data analysis, contributing to enhanced glacier monitoring and climate research.

Acknowledgements

I want to thank my primary supervisors, Anthony Doulgeris and Deyan Draganov, for making this Master's thesis possible. Thank you, Anthony, for your guidance and time spent on the technical aspects of this thesis. These 5 months introduced me to the SAR processing chain, and I look forward to using the knowledge gained in my future scientific career. Thank you, Deyan, for finding ways to combine this project as part of the Idea League degree, ensuring this thesis was on track, and for your general advice. I am also highly grateful to Johannes Lohse for your support and for providing valuable help in understanding the algorithmic side of the project. Much of this thesis is based on Johannes' work, and he was always available to answer questions and discuss matters.
I want to extend my appreciation to the Earth Observation group at UiT. From the lunch chats to the bouldering, hikes, and board game meet-ups, you have made my short stay in Tromsø very welcoming.

This thesis is also the result of many insightful discussions with Jack Kolher and Geir Moholdt from the Norwegian Polar Institute (NPI). Your feedback on the glaciology side and your insights on the practical implications of this project on Svalbard glaciers have been invaluable. I also want to thank the glaciology group at NPI for welcoming me to their plenary meetings.

I want to express my gratitude to Stefan Baecke, CEO of Yespers, my supervisor at the part-time work that supported my studies. His belief in my abilities and his flexibility allowed me to balance professional responsibilities with academic pursuits.

Thank you to my friends and family for their encouragement and last but not least to Malte for his unwavering support.

UiT The Arctic University of Norway — Marion Dugué
July 29, 2024

Table of Contents

Abstract v

Acknowledgements vii

Nomenclature xvii

Acronyms xvii

1 Introduction 1
- 1-1 Motivation . . . 1
- 1-2 State of the art . . . 2
- 1-3 Research objectives . . . 3
- 1-4 Thesis structure . . . 4

2 Theoretical background 5
- 2-1 SAR remote sensing concepts . . . 5
  - 2-1-1 Sentinel-1 product availability . . . 5
  - 2-1-2 Principles of SAR imaging . . . 6
- 2-2 Statistics for Bayesian classifiers . . . 10
- 2-3 Glaciology . . . 11
  - 2-3-1 Anatomy of a glacier . . . 11
  - 2-3-2 Glacier zonation through SAR imagery . . . 12
  - 2-3-3 Glaciers of interest . . . 15

3 Data and method 19
- 3-1 Data selection . . . 19
  - 3-1-1 Glacier boundaries . . . 19
  - 3-1-2 Digital Elevation Model . . . 19

3-1-3 SAR products . . . 20
3-2 Classification method . . . 21
3-2-1 Data pre-processing . . . 21
3-2-2 Backscatter dependency on IA . . . 22
3-2-3 Algorithm implementation . . . 22
3-3 Texture feature . . . 23

4 Results **27**
4-1 Data selection . . . 27
4-1-1 Data availability . . . 27
4-1-2 Months . . . 27
4-1-3 Relative orbits . . . 28
4-2 Classification . . . 29
4-2-1 Backscatter dependency on IA . . . 29
4-2-2 Classification of training regions . . . 32
4-2-3 Classification of glaciers . . . 32
4-3 Texture feature . . . 37
4-3-1 Window size analysis . . . 38
4-3-2 Linearity with local IA . . . 38
4-3-3 Classification result on training data . . . 38
4-3-4 Classification results on glacier . . . 38

5 Discussion & future work **43**
Texture feature . . . 44

6 Conclusion **51**

A Additional plots - Data selection **59**
A-1 Months HV . . . 59
A-2 Local Incidence angles (projected or not) . . . 59
A-3 Glacier outlines . . . 61
A-4 SAR results example scene . . . 62

B Spatial Fourier transforms **63**

C Data management **69**

List of Figures

2-1 Synthetic Aperture Radar imaging geometry where H is the elevation of the radar antenna from the ground, θ is the look angle, and β is the width of the antenna beam. . . . 7

2-2 Geometry illustrating SAR resolution between two targets (red and blue rectangles). Based on Curlander and McDonough (1991). . . . 8

2-3 Different small-scale scattering mechanisms (dotted arrows) from an incident beam (full arrow). a) corresponds to surface scattering (for two different types of surface roughness at wavelength scale), b) to double-bounce scattering, and c) to volume scattering . . . 9

2-4 Simplified structure of a glacier's facies within the first meters of the subsurface. These facies result from the transformation of snow to ice, which is influenced by altitude (/temperature variations) and compaction levels. Diagram based on Cuffey and Paterson (2010). . . . 13

2-5 HDF and KNG glacier complexes. Each complex comprises multiple glaciers; the areas found through QGIS are detailed in the figure. Base map from Norwegian Polar Institute (2017). . . . 16

2-6 Position of the nine stakes placed along the centerline of KNG for annual mass balance monitoring. Source: Hawley et al. (2008). . . . 17

2-7 Photographs at different elevations of the surface of KRB and KNG to visualize glacier facies (Personal communication, Jack Kholer). . . . 18

3-1 Naming conventions of S1 products useful to filter the scenes per satellite, mode of acquisition, polarisation, time of acquisition, and orbit number. Source: Copernicus, 2024 . . . 20

3-2 Pre-processing graph for all the SAR IW and EW products used in this project. . 22

3-3 Training regions' locations on KNG and HDF. . . . 24

3-4 Visualisation of the standard deviation calculation. A window (here of size 1) travels through the glacier image, calculating the standard deviation for the center pixel. To get reliable results of standard deviation metrics (mean and std) for each training region, this process is done multiple times over a larger grid and then averaged. . . . 25

4-1 Mean backscatter intensity for the HH channel for IW products spanning our glacier complexes across 2021. All intensities are stable within 1 dB from January to April when the least melt and refreezing occurs. 28

4-2 Backscatter intensity spread of HH and HV bands of all training regions. 29

4-3 Plotting the backscatter intensity of the HH band versus the mean local IA for both EW and IW over the whole image. IW data does not have a large variation in mean IA as there is only one relative orbit, but we can observe that IW data for each zone fall into the corresponding EW fitted line. 30

4-4 Comparison of HH backscatter distributions of each class (left) with their 2D separability when plotted as a function of incidence angle (right). 31

4-5 Confusion matrices for training regions comparing Gaussian classifier (left), common IA slope correction (center), and per-class IA slope correction (right). SI abbreviates the Superimposed Ice zone, and GI1 & GI2 correspond to the zones Glacier Ice with texture1 and Glacier Ice with texture2, respectively. 32

4-6 Comparison of Gaussian classification without any correction (left), common IA slope correction (center), and per-class IA slope correction (right) for both EW (top) and IW (bottom) products. 33

4-7 Comparison between two SAR products close in time to show that the classification is reliable over time. For the 12-day time difference, we do not expect the glacier facies to drastically change position, and this is what we observe for firn (orange) and superimposed ice (purple). The two glacier ice textures (green and blue) show the most differences between the two product classifications. This correlates with the two texture ice distributions greatly overlapping and some visual differences observable in the backscatter intensities. 35

4-8 Firn line comparison from 2017 to 2023 for HDF (a.) and KNG (b.). A homogeneous retreat of around 570m can be observed for KNG perpendicular to its slope. For HDF, feature 1 is not classified as firn in 2023, and feature 2 & 3 are more sparse 36

4-9 Separability of the four regions firn, SI, glacier ice, and crevasse with (right) and without (left) IA considerations. The crevasse distribution significantly overlaps with the firn and SI distributions, considering the backscatter intensities. However, its IA slope is positive, which makes it more separable when correcting for per-class IA. 36

4-10 Confusion matrices over the training regions comparing training accuracy for a differently fitted algorithm. 37

4-11 Window size comparison of the std calculation for EW and IW. The objective is to find the minimum size window so that each class is the most separable. 39

4-12 Calculation of the mean, standard deviation of the HH band backscatter intensity averaged across training regions versus the local IA for years 2015 (red), 2020 (blue), and 2023 (green). 40

4-13 Mapping the labeled crevasses on optical imagery when using standard deviation as a pixel feature. Strong crevasses are correctly labeled but not those that have more irregular patterns, as seen by the red ellipses corresponding to misclassified regions. 41

4-14 Per-class IA slope correction classification result including standard deviation of HH and HV as features 42

5-1 Original HH band of the 420*240 pixels regions for firn (bottom row) and crevasse (top row) taken from HDF for both EW (left column) and IW (right column). . . 46

5-2 Plotting mean magnitude from 0 to 0.04 cycles/meter for crevasse and firn. . . . 46

5-3 Four patches of 100*100 pixels from the IW crevasse region that visually show strong crevassing. 47
5-4 Spatial Fourier transform analysis for the four crevasse patches identified with strong crevassing. 48
5-5 Wavelet filter plot made from a sinusoid and gaussian envelope. 49
5-6 Crevasse wavelet analysis results. The crevasse region was convolved with a wavelet filter directed with various angles. The absolute maximum value of the whole image result was found and defines the max_score of the angle of the wavelet. 49
5-7 Firn wavelet analysis results once convolved with a directed wavelet of various angles. 50

A-1 HV backscatter intensities across 1 year. HV values are more consistent from January to April just as observed for HH values. 59
A-2 Comparison between local IA and projected local IA in geometry. The blue plane is the slant-range plane and is defined by the beam source from the satellite and the nadir point. The grid corresponds to the ground range plane and the reflector is depicted here as a brown sphere. (a) corresponds to a view in which the projected local incidence angle is not visible. In (b), the norm to the slant-range (dotted green line), the projected surface norm (red), and the projected local incidence angle are visible. 60
A-3 Comparing the GLIMS outline (red) from the year 2000 to recent optical imagery from LANDSAT-8. There is a retreat of the glaciers where it reaches the sea to around 3km on average. For the rest of the two glacier complex, the outlines matches still the glacier of today. Base map from Norwegian Polar Institute, 2017. 61
A-4 SAR results example scene with training regions position. 62

B-1 Crevasse region that was segmented into smaller patches. The red dots correspond to the centers of the 21 patches. 63
B-2 Subset crevasse patches from the larger region found in Figure B-1. 64
B-3 Spatial Fourier transform images of the crevasse patches. 64
B-4 Spatial Fourier transform images of the firn patches. 65
B-5 Spatial Fourier transform analysis of the firn patches, applying averages over the X and Y axes. 65
B-6 Spatial Fourier transform analysis of the crevasse patches, applying averages over the X and Y axes. 66
B-7 Radial frequency of spatial Fourier transforms of crevasse patches 66
B-8 Radial frequency of spatial Fourier transforms of firn patches 67

List of Tables

2-1 SAR bands with corresponding frequency and wavelength ranges. 5

4-1 Slope of HH band vs IA calculated from the different training regions and averaged. 31

Acronyms

DUT Delft University of Technology

ETH Swiss Federal Institute of Technology

RWTH Aachen University

UiT The Arctic University of Norway

NPI Norwegian Polar Institute

SAR Synthetic Aperture Radar

IA Incidence Angle

GLIA Gaussian Linear incidence angle correction algorithm

HDF Holtedahlfonna glacier complex

KNG Kongsvegen glacier complex

RTC Radiometric Terrain Correction

GPR Ground Penetrating Radar

I Surface Intensity at a given local IA

Chapter 1

Introduction

This chapter presents the context, motivation, existing research, and research objectives. Section 1-1 provides context and shows the research gaps. Then, Section 1-2 goes into a more thorough literature review to present state-of-the-art methods. The research objectives are detailed in section 1-3, and then the thesis structure is provided in section 1-4.

1-1 Motivation

The Intergovernmental Panel on Climate Change (IPCC) has been monitoring the state of the world's glaciers since 2014. It reports that in a span of 7 years, almost all worldwide glaciers have shrunk and will continue to shrink as they are out of balance with current climatic conditions. Unfortunately, the changes in glaciers due to climate change can disrupt the global water cycle, affecting human resource economics (Vaughan et al., 2013). Glaciers in the Arctic are particularly at risk as this region is warming nearly four times as much as the rest of the world, and climate models are systemically underestimating this amplification (Rantanen et al., 2022). Glaciers are an important proxy for climate change; thus, understanding their evolution in time provides key insights into climate patterns.

While fieldwork on glaciers is resource-expensive, the development and use of remote sensing have been of great value for monitoring those often inaccessible regions. For regions such as Svalbard, situated above the Arctic Circle, optical imagery is obsolete during polar winters and in cloudy weather. Using active microwave sensors solves this issue as they can penetrate clouds, and there is no need for sunlight. In the 1980s, scientists showed the use of Synthetic Aperture Radar (SAR) data for snow and glacier monitoring which sparked interest in the scientific community to further organize large-scale missions to acquire SAR data and make the data easily accessible (Rott and Mätzler, 1987).

The European Space Agency (ESA) pioneered such missions with the European Remote-Sensing Satellites (ERS) 1 and 2 in 1991 and 1995, which provided new methods for snow

mapping (Rott and Nagler, 1994, Nagler and Rott, 2000). Closely following this, other space agencies worldwide launched their missions, such as the Canadian Space Agency with RADARSAT-1, the Japanese Space Agency with ALOS-1 in 2006, and the Indian Space Agency launching RISAT in 2012. In 2014 and 2016, ESA subsequently started the Copernicus Program, which included two satellites aiming to provide SAR imagery at a high revisiting rate of 6 days with high resolution. Equipped with C-band SAR instruments, Sentinel 1A and 1B provide openly accessible data to analyse maritime and land regions, and their evolution in time.

The intensity of the backscatter received by the satellite varies according to surface and subsurface properties, including (non-exhaustively) surface roughness, snow grain size, dielectric properties, and overall the amount of water and air the target contains (Fung, 1994). In this way, glacier facies, characterised by separate physical properties, can be differentiated using SAR data. Due to the geometry of the acquisition of SAR imagery, the backscatter intensities depend on the incidence angle (IA) between the microwave signal and the norm of the surface it reaches. As the IA becomes larger for a far-range target, the backscatter intensity is even more compromised, which requires correction before any backscatter analysis can be done on the image. The distortion is increasingly hindering the final analysis for an image containing varying surface roughness. This thesis aims to implement a novel IA correction for glacier classification to study how much it improves the accuracy of glacier facies detection.

1-2 State of the art

For SAR imagery in the C-band, up to a few meters of dry snow are considered transparent (Adam et al., 1997), and SAR glacier zones can be correlated (though not directly mapped) to glacier facies (Partington, 1998). Different methods have been developed to differentiate SAR glacier zones from visual observations including k-means clustering (König et al., 2004, König et al., 2004), supervised and unsupervised algorithms using K-Wishart distributions (Doulgeris et al., 2008), contextual non-Gaussian clustering (Akbari et al., 2013), fully polarimetric support vector machine (SVM) classifier (Callegari et al., 2016, Thakur et al., 2016), and convolution neural networks (Baumhoer et al., 2019). Before analysing the backscatter intensities, the data must be corrected for terrain, noise, and backscatter intensity decay due to the incidence angle. Terrain and noise correction are done in standardized pre-processing steps. However, the widely used method to correct IA across all zones within one SAR image is using an averaged constant correction throughout (Rott and Davis, 1993). We refer to this as a *common IA slope correction*. With a constant correction, this pre-processing step, included in the *radiometric terrain correction* (RTC), will then over or under-compensate certain zones within the image.

Huang et al. (2016) have integrated local IA correction for glacier classification in an SVM algorithm by separately segmenting the image per local incidence value, followed by training and predicting the SVM within those zones using only backscatter intensities. Per-class IA slope correction was implemented for sea ice by replacing a constant mean backscatter intensity vector from a Gaussian probability density function with a linearly variable mean. This was successfully done with a supervised algorithm (Lohse et al., 2020)

and an unsupervised segmentation (Cristea et al., 2020). Contrary to the scientific literature highlighting the variation of IA for sea-ice backscatter intensities (Mäkynen and Karvonen, 2017), the per-class IA slope dependency for glacier zones has not been studied, nor has a per-class IA slope correction per glacier class been implemented. This would also be the first time a per-class IA slope correction is performed on land studies.

Glaciologists at the Norwegian Polar Institute (NPI) have annually monitored the evolution and mass balance of Kongsvegen (KNG) and Holtedahlfonna (HDF) glacier complexes. Following Benson (1960) definition, a glacier sequentially presents different facies based on its physical properties. Firn, superimposed ice, and glacier ice are such zones and have been studied on our glaciers of interest using Ground Penetration Radar (GPR) (Melvold and Hagen, 1998, Dunse et al., 2009, Langley et al., 2009). These two glacier complexes are closely located to Ny-Ålesund, a town on the West coast of Svalbard, which grew from a mining area to a permanent research settlement in 1966. This easy access to KNG and HDF motivates the extensive research conducted there and provides valuable ground truth data for glacier monitoring using satellite imagery. On KNG, König et al. (2002) have used SAR satellite imagery (ERS) to detect superimposed ice with a comparison with ice cores, and Langley et al. (2008) showed that SAR satellite data (data from the Advanced Synthetic Aperture Radar (ASAR) instrument) aligns well with GPR-derived glacier facies by visually inspecting the backscatter image. KNG and HDF are two different glacier types with varied surface complexity (low for KNG and high for HDF) - more on these two glaciers can be read in section 2-3-3.

1-3 Research objectives

We specifically focus on using Sentinel-1 data, the most up-to-date and freely available SAR satellite imagery.

Before using the per-class IA slope correction algorithm on our glacier data, we first need to understand which data type is available and what should be used for optimal results. Then, we aim to implement the correction with a linearly variable mean backscatter intensity dependent on the per-class IA slope correction algorithm developed by Lohse et al. (2020) called originally “GIA”. In this work, we want to emphasize the use of the linear assumption so that we will call it throughout this report “GLIA”. This will be done on our training regions and on HDF and KNG glacier complexes to assess the extent to which this correction improves in general glacier classification and the limitations. Finally, it will be shown that our classifier vastly mislabels the crevasse region of HDF, so we attempt to extend the algorithm to include a texture feature by exploring features involving standard deviation and spatial Fourier transforms.

In all, we can summarise our research objectives as below:

1. What data is necessary to implement the per-class IA slope correction algorithm for our glaciers of interest?

2. Implementation of the per-class IA slope correction in classification:
   (a) How well does this new IA correction improve glacier classification compared to state-of-the-art methods?
   (b) From a glaciology perspective, how does the firn line of HDF and KNG evolve over the years?
3. How to extend the GLIA algorithm to include a texture feature?

1-4 Thesis structure

In Chapter 2, the thesis explores the theoretical concepts necessary to understand the project as a whole. In section 2-1, SAR remote sensing concepts are detailed, including the type of data we can acquire with Sentinel-1 products, the acquisition geometry, and the overall understanding of the backscatter intensity. Then, as we apply Bayesian classifiers, the statistics involved are outlined in section 2-2. The final theoretical chapter includes a definition of glaciers and glacier facies and how we can relate them to SAR imagery. This chapter ends with a description of our two glacier complexes of interest (Holtedahlfonna and Kongsvegen) in Section 2-3.

In Chapter 3, the methods and data used are specified. Following our three research objectives, we have three sections: section 3-1 tackles the data selection methodology, including the glacier boundaries, digital elevation model, and SAR products used; section 3-2 outlines the methods for pre-processing, common IA slope and per-class IA slope correction, algorithm implementation and training data; section 3-3 provides the process for the texture feature extension to improve classification accuracy.

Chapter 4 presents the results in a similar 3-step fashion for data selection in section 4-1, classification in section 4-2 (on training data and then glaciers) and then including the texture extension in section 4-3.

Discussions on the results and suggestions for future work can be found in chapter 5. A summary of the project can be finally found in the conclusion in Chapter 6.

The appendices provide additional plots, data, and explanations of steps to support the research findings. Specifically, Appendix A details HV results, the difference between projected and non-projected IA, and the glacier outline comparison between what is publicly available and what was internally provided. Appendix B contains additional figures for Spatial Fourier transform patch and wavelet analysis. The final chapter C explains the management of data throughout the project.

Chapter 2

Theoretical background

2-1 SAR remote sensing concepts

This chapter presents in subsection 2-1-1 the type of SAR data and their applications are explained in subsection 2-1-2 the basics of SAR imaging and its resolution parameters, with the geometry of SAR acquisition and an explanation of the backscatter coefficient.

2-1-1 Sentinel-1 product availability

Satellite-based SAR acquires valuable data for glacier monitoring as the microwave frequency range (300 MHz to 30 GHz) penetrates through clouds and is thus minimally affected by weather (Forster et al., 1996, Braun et al., 2000). The different wavelengths of SAR systems are referred to as bands and follow the frequency and wavelength ranges as shown in Table 2-1.

Band	Frequency	Wavelength
Ka	27–40 GHz	1.1–0.8 cm
K	18–27 GHz	1.7–1.1 cm
Ku	12–18 GHz	2.4–1.7 cm
X	8–12 GHz	3.8–2.4 cm
C	4–8 GHz	7.5–3.8 cm
S	2–4 GHz	15–7.5 cm
L	1–2 GHz	30–15 cm
P	0.3–1 GHz	100–30 cm

Table 2-1: SAR bands with corresponding frequency and wavelength ranges.

The selection of the SAR band depends on the extent to which one wants the radar to interact with the target and how deep it should penetrate. For instance, if one wants subsurface information that can penetrate through vegetation and soil, then the L or P-band with its

long wavelengths would be suited. On the other hand, if detailed surface information is required for urban planning, then the X or K-band is adequate. The Sentinel-1 (S1) satellites are two C-band satellites that emit energy at a carrier frequency of 5.4 GHz (Aulard-Macler, 2011). The acquisition modes include extra-wide swath (EW) (400 km for a resolution of 20 m * 40 m) and interferometric wide swath (IW) of higher resolution (250 km for 5 m * 20 m). IW products are primarily available on land, while EW covers maritime or polar regions. This active imaging system transmits a beam on the surface and provides a 2D image of the complex radar reflectivity of a scene. The waves transmitted and received are linearly polarised, with horizontal (H) or vertical (V) polarization. This means the received signal can be a combination of HH, HV, VH, or VV polarizations, where the first letter shows the polarization of the emitted wave and the second of the received one. Typically, smoother surfaces will minimally interact with the emitted beam, so the received one will have the same polarization orientation. So, co-polarizations HH and VV will have higher intensities. On the other hand, cross-polarizations (HV and VH) will have a higher magnitude of backscattering coefficient for a stronger surface roughness as it mostly comes from volume scattering or multiple reflections.

Finally, S1 SAR products are available in different levels of in-house pre-processing by ESA. Single-look complex (SLC) products keep phase information and are widely used for interferometry analysis. In comparison, ground-range detected (GRD) products have been multilooked and only contain detected amplitudes. IW GRD products are images with 10 * 10 m pixels, while EW GRD products have 40 * 40 m pixels. We will use EW and IW GRD products throughout this work.

2-1-2 Principles of SAR imaging

This subsection is mostly based on the theory of remote sensing presented in the "Synthetic Aperture Radar: Systems and Signal Processing" textbook by Curlander and McDonough (1991).
Usually, the radar system is mounted on an aircraft or satellite, and the onboard antenna emits and receives reflected microwave signals. As spatial resolution is proportional to the ratio of the signal wavelength to the length of the sensor's antenna, SAR systems simulate a larger antenna (and thus a higher spatial resolution) by combining sequences of shorter acquisitions. It assumes that as the system moves along the flight path, i.e., in the *azimuthal direction* and the target is within the radar footprint during what is called the *Synthetic aperture length*. Received pulses from targets on the ground, also called *chirps*, are then summed, compressed, and finally processed using an inverse dispersive filter. The result of this is a much narrower pulse. Higher resolution can be obtained by increasing the bandwidth, i.e., the range of frequencies within the chirp, and reducing the antenna length in the limit of system configurations.

SAR acquisition geometry For the case of the transmitter and receiver being at the same position (i.e., a monostatic imaging system), Figure 2-1 represents schematically the geometry of the side-looking SAR system.

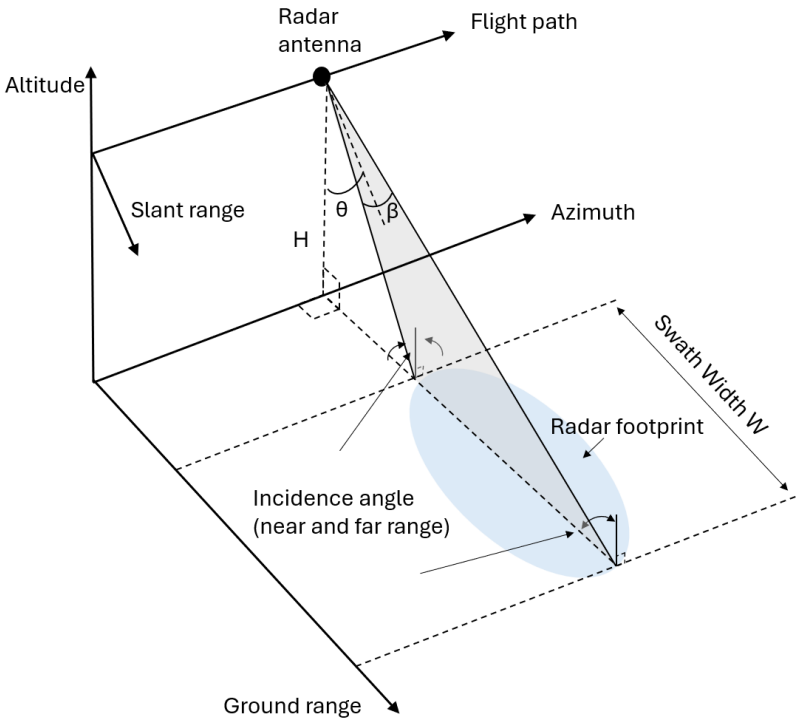
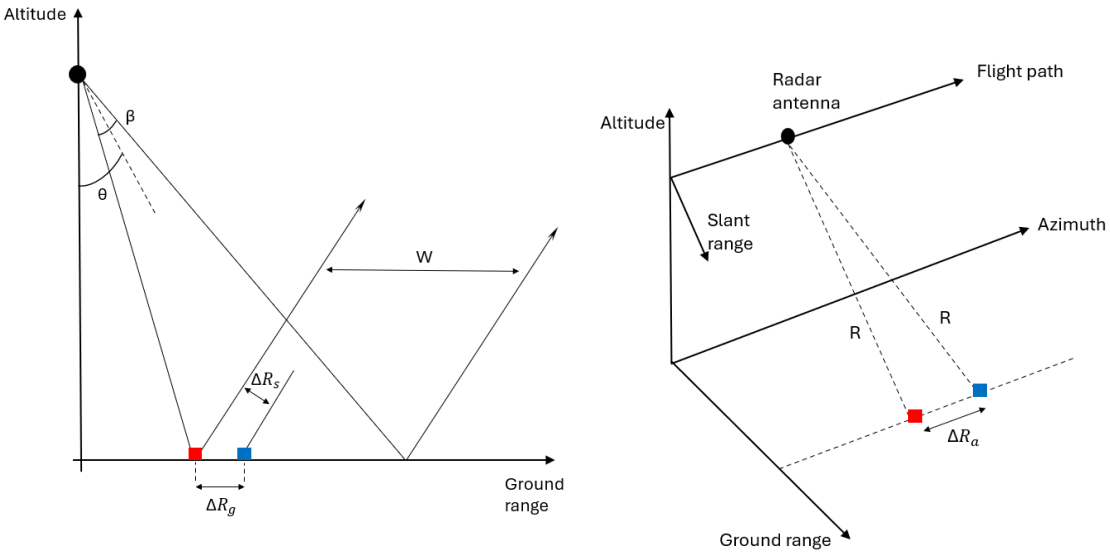


Figure 2-1: Synthetic Aperture Radar imaging geometry where H is the elevation of the radar antenna from the ground, θ is the look angle, and β is the width of the antenna beam.

As the beam is directed at an angle to the surface, it measures the backscatter in the *slant-range*; however, for interpretation purposes, it is projected to the horizontal distance of Earth's surface, which is the *ground range*. The *radar footprint* then commonly refers to the ground-range illumination of the radar beam. At the end of pre-processing, SAR products provide backscatter strengths for each pixel, where the size of each determines the range and azimuth resolution of the radar system.



(a) Illustration of ΔR_g and ΔR_s, respectively defined as the ground and slant range resolutions.

(b) ΔR_a Azimuth resolution and slant range R.

Figure 2-2: Geometry illustrating SAR resolution between two targets (red and blue rectangles). Based on Curlander and McDonough (1991).

Resolutions The resolution corresponds to the minimum distance that two points along a specific line can be differentiated. Figure 2-2 illustrates the geometry to understand ground-range, slant-range, and azimuth resolution. For a general SAR system, the resolution ΔR_g in ground range can be expressed as

$$\Delta R_g = \frac{c}{2B_c \sin\theta}, \tag{2-1}$$

where c is the speed of light, B_c is the chirp bandwidth and θ the look angle.
The resolution along the azimuth is defined as being half the antenna size L, i.e

$$\Delta R_a = \frac{L}{2}. \tag{2-2}$$

Sentinel-1 uses different acquisition modes, including strip mode (SM) and TOPSAR (acquisition mode of IW and EW products). SM products have an azimuthal resolution following equation 2-2, but due to TOPSAR's sweeping pattern, which aims to create a wider swath and have longer illumination times of a target, the resolution of IW and EW is less than that

of strip mode (De Zan and Guarnieri, 2006). From Sentinel-1 documentation, the IW and EW range*azimuth resolutions (in m) are 20 * 22 for IW and 93 * 87 for EW. The products are oversampled as the pixel size is 10 m and 40 m for IW and EW, respectively.

Backscatter coefficient The calibrated backscatter intensity σ^0, corresponding to the squared-backscatter coefficient, measures how much power is scattered back per unit area on the ground. Following the F. T. Ulaby et al. (1982) derivation, σ^0 is obtained from the radar equation by normalizing the received power to account for the radar system characteristics and area covered by the beam. The backscatter intensity depends on frequency, polarization, IA of the beam, and scattering mechanisms of the target (dielectric constant and geometry). The higher the surface roughness or small IA, the higher the backscatter intensity. Figure 2-3 presents different mechanism types. The dominating scattering mechanisms can be explained by the surface roughness at the wavelength scale, as well as the properties of the media itself.

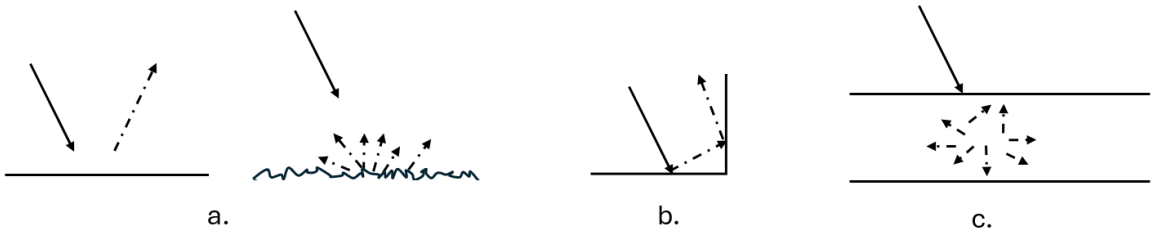


Figure 2-3: Different small-scale scattering mechanisms (dotted arrows) from an incident beam (full arrow). a) corresponds to surface scattering (for two different types of surface roughness at wavelength scale), b) to double-bounce scattering, and c) to volume scattering

IA effect As SAR collects data using a side-looking configuration, the IA between the emitted beam and the surface normal increases with distance. Thus, the backscattered energy decreases proportionally as the image is captured further from the sensor. This intensity decay, also called *IA effect*, is more pronounced the wider the swaths and, without correction, makes target differences harder to distinguish.

Empirically, it has been observed that the decay could be exponentially approximated, such that the surface intensity observed at a given incidence angle θ_i is

$$I_{\theta_i} = I_0 e^{-\theta_i/\gamma}, \tag{2-3}$$

where I_{θ_i} is the surface intensity observed at incidence angle θ_i, I_0 is the intensity (brightness) at incidence angle 0, θ_i is the wave incident angle, and γ is the constant associated with a target class (Mäkynen and Karvonen, 2017, F. Ulaby, 1980).

In the logarithmic domain, this expression becomes linear and can be written as

$$I_{\theta_i}[dB] = a + b\theta_i, \tag{2-4}$$

where $b = -1/\gamma$ and $a = \log I_0$ (Makynen et al., 2002, Gill et al., 2015, Zakhvatkina et al., 2017, Mäkynen and Karvonen, 2017, Mahmud et al., 2018). Backscatter intensities in the log-domain are commonly assumed as multivariate Gaussian distributions, characterized by their mean and covariance. These two parameters are further dependent on IA as the decay is minimal for a closer range (small IA), and thus, the mean intensity is higher than those of the far range. On the other hand, for data acquired over a broad range of IA, the mean will not be representative, and the variance will be high.

To mitigate this, a common correction across all classes can be applied, either using a constant IA slope based on the IA from the beam or by assuming the simplification of an ellipsoid earth. Generally, the latter method, which we refer to here as *common IA slope correction*, is widely used for radiometric correction as part of the pre-processing of products. This is done in a method called "Radiometric Terrain Correction", where both geometric distortions (terrain correction) and radiometric correction (common IA slope correction) are applied. However, while topography is accounted for, the common IA slope correction with a constant rate of decibels per IA degree ([dB/1°]) for all classes would overcorrect or undercorrect certain classes as the rate of decay depends mostly on roughness conditions and hence glacier facies (Wagner et al., 1999). Some prior research has investigated the possibility of manually correcting the IA based on the class; however, this method is time-consuming (Mladenova et al., 2012, Lang et al., 2016).

Different methods were implemented to account for variations within IAs in a SAR image. Karvonen et al. (2002) initiated an iterative per-class correction by normalising the backscatter intensities with the class's relevant slope while running the classification in parallel. Lohse et al. (2020) created an automated per-class IA correction by directly calculating the probability density functions for a given data point with a linearly variable mean from given slopes. At the same time, Cristea et al. (2020) developed an unsupervised approach, substituting a constant mean of backscatter intensities with a linearly variable one.

2-2 Statistics for Bayesian classifiers

In this work, we adopt the classification approach developed by Lohse et al. (2020) and Cristea et al. (2020). While one is a supervised classification and the other an unsupervised segmentation, both are based on Bayesian classification. Hence, this section introduces the concept of Bayesian classifiers and the mathematical formulations using SAR backscatter data.

Bayesian classifier Following Bayes' rule, where $p(\omega_i|\mathbf{x})$ is the posterior probability, $p(\mathbf{x}|\omega_i)$ the class conditional probability, and $p(\omega_i)$ the prior probability, we have

$$p(\omega_i|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_i)p(\omega_i)}{p(\mathbf{x})}. \tag{2-5}$$

If we assume equal prior probabilities for each class, i.e. $p(\omega_i) = p(\omega_j)$ for $\omega_i \neq \omega_j$, then from equation 2-5, maximizing the posterior probability $p(w_i|\mathbf{x})$ is equivalent to maximizing

the class conditional probability $p(\mathbf{x}|w_i)$. This is the principle behind *maximum likelihood classification.*

For classification purposes, the class conditional probability $p(\mathbf{x}|\omega_i)$ represents the likelihood of observing the feature vector $\mathbf{x}$ given that it belongs to class ω_i. This is also known as the probability density function (PDF) for sample $\mathbf{x}$ given class ω_i. Class assignment in Bayesian classification is then done by assigning to the data point the class with the highest posterior probability such that the maximum likelihood decision rule is

$$\mathbf{x} \rightarrow \omega_i \text{ if } p(\mathbf{x}|\omega_i) > p(\mathbf{x}|\omega_k) \quad \forall k \neq i. \tag{2-6}$$

Backscatter intensities in the log-domain can be approximated as multivariate Gaussian distributions. So for a class ω_i with mean vector μ_i and covariance matrix Σ_i, its PDF is

$$p_i(\mathbf{x}|\omega_i) = \frac{1}{(2\pi)^{d/2}|\Sigma_i|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\mu_i)^T \Sigma_i^{-1}(\mathbf{x}-\mu_i)}. \tag{2-7}$$

By determining from labeled data the mean and covariance corresponding to each class and assuming the features of $\mathbf{x}$ (corresponding to the pixels of the SAR image in our case) are normally distributed and independent given the class, the classifier can compute the PDF of the sample for each class and assign the one with the largest PDF. The per-class IA correction method will be explained in the Methods chapter in section 3-2-2.

2-3 Glaciology

This section summarizes the essential concepts of glaciology to use SAR data for glacier zone classification. In subsection 2-3-1, the anatomy of a glacier is described, with the key subsurface characteristics and ice differences provided. In subsection 2-3-2, we dive deeper into how to interpret SAR data for glacier zonation. Finally, a description of the two glaciers of interest is provided in subsection 2-3-3.

2-3-1 Anatomy of a glacier

While we can categorize these by the surface material's temperature and physical characteristics, the transformation of snow into ice is continuous with no clear distinction. It depends on the type of glaciers, their topography, and the general environment influencing the melt and refreezing cycles. For the study of HDF and KNG, firn, glacier ice, and superimposed ice are the most relevant glacier facies and will be the ones studied throughout this work (Personal communication, NPI).

We primarily base our understanding on the textbook by Cuffey and Paterson (2010) to understand the differences between those zones. The glacier facies differentiated below follow the definition set by Benson (1960) and are further developed and illustrated by Paterson (1994). Figure 2-4 presents a revisited simplified structure of the first few

meters of a glacier's subsurface. We describe them below from the highest elevation, where little or no melting occurs, to the lowest, where there is more melting than snow accumulation.

- **Firn**: Firn corresponds to wetted snow that has survived one summer without being transformed to ice and snow on polar glaciers where no melting occurs. There is typically an ambiguity in the distinction between firn and snow; however, in our project, as we focus on SAR imagery in winter months, we will combine the two into one class that we name "firn." Firn can be found in the *dry zone* where no melting occurs, in the *percolation zone* where there is intermittent melting and refreezing, and in the *wet zone* where there is persistent liquid water present. Firn presents some ice structures (ice glands, lenses, and general layers) due to the meltwater refreezing in the percolation and wet zone. The lower boundary between the firn to ice transition is called the *firn or snow line* and typically will not have strong annual variations. The firn line will take at least multiple years of constant net negative mass balance to retreat (König et al., 2002). On the other hand, several net positive mass balance years will lead the firn line to advance in lower altitudes of the glacier. A retreat of the firn line means that more bare ice is exposed, which in turn furthers the surface melt (Hall, 2004, Van den Broeke et al., 2008). In this way, the firn line measures long-term trends in glacier mass balance.

- **Superimposed ice (SI)**: When melting and refreezing due to temperature changes across seasons, layers of ice are formed. This happens after the wet-snow zone at the surface and right before it is buried under firn. The end of the SI zone is called the *equilibrium line*, as the glacier loses as much mass as it gains along this line. Above it, the glacier has a positive mass balance (*accumulation zone*) and under it a negative one (called the *ablation zone*).

- **Glacier Ice**: Glacier ice is formed in a different mechanism than SI. As snow accumulates and the overlying weight increases, the firn below becomes more compact, and pores filled with air and water progressively close off. Glacier ice is at the end of the gradual metamorphism of snow called *firnification*. Glacier ice still contains minimal air bubbles and is much denser than firn. This is typically the surface of glaciers in the ablation zone. At the end of summer, all snow and SI will have melted away in lower altitudes, which means the glacier ice is at the surface.

2-3-2 Glacier zonation through SAR imagery

On top of varying across local IA, backscatter intensities and general penetration depth depend on the water content and dielectric constant, density, structure of the media, and target surface roughness. Therefore, we can use SAR data to identify the types of zones the glacier presents.
The glacier facies identified earlier from Benson do not necessarily match the ones found using SAR imagery as they provide zones based on fluctuations in the physical properties of the snow, which could change in terms of days and weeks. In contrast, glacier facies are due to mechanisms built over the years (Langley et al., 2008). Therefore, we make the distinction

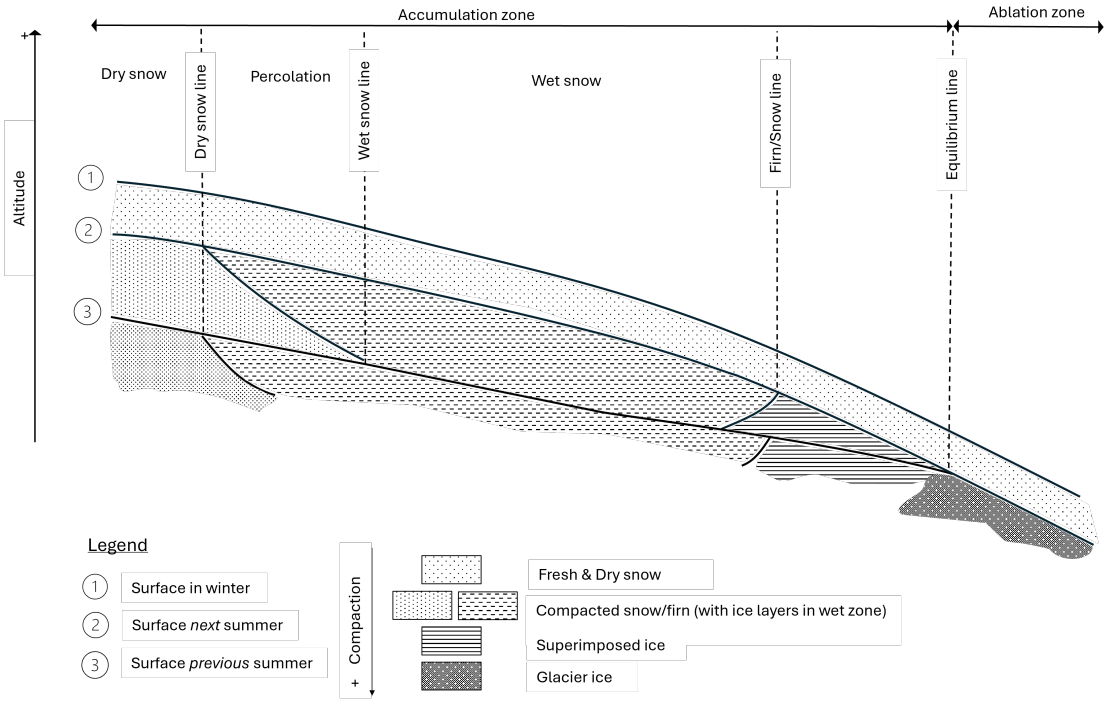


Figure 2-4: Simplified structure of a glacier's facies within the first meters of the subsurface. These facies result from the transformation of snow to ice, which is influenced by altitude (/temperature variations) and compaction levels. Diagram based on Cuffey and Paterson (2010).

between glacier facies, which are a glaciology interpretation of snow and ice transformation following Benson's definition, with radar glacier zones (also called "zones" throughout this work) (Forster et al., 1996). Acknowledging that SAR classes do not correspond directly to glacier facies, they can be mapped by knowing the glacier's structure. In this way, as we are guided by glaciologists at the NPI, glacier "zones", "facies", and SAR "classes" will be words interchangeably used in the rest of this work.

Relatively dry snow zone : A dry-snow zone can be defined as a function of its volume water content (VWC) or of the temperature at which the snowpack is - below 0°, the zone remains dry. This zone is free of liquid water, and the only scattering contribution is due to the snow grains. Snow grains are much smaller (millimeter scale) than the microwave wavelength (centimeter scale for C-band), so the SAR signal of a C-band system will pass through the snow crystals without major disruptions. For dry *non-compact* snow lying on the glacier's surface, it will be essentially transparent to SAR imagery.

Percolation zone In the percolation zone where meltwater is present and refreezes, the dominant scattering is volume, which adds to the surface scattering from the ice structures created by meltwater refreezing. So we can correspond a high backscatter intensity zone with a *compacted and wet* firn zone.

Wet snow zone and ablation zone Backscatter intensities and penetration depths in wet zones are lower due to strong attenuations from liquid water. Until 1995, it was thought impossible to distinguish between glacier ice and SI until (Marshall et al., 1995) reported using surface roughness with SAR data. Since then, König et al., 2002 has shown that SAR imagery can distinguish between SI and glacier ice because SI has a higher air bubble content than homogeneous ice. It is noted that surface layers of SI tend to have smaller bubbles, which have a lower backscatter intensity. This study found that SI thinner than 10 cm was not detected using SAR imagery. However, 25 to 100 cm thick layers were detected, so the thickness of the SI layer should be considered. The crystal size and orientation were reported to differ from SI to glacier ice on a Canadian ice cap, which could also provide differences in the SAR imagery; however, this was not tested or researched in Svalbard (Koerner, 1970). Because SI presents more air bubbles, it will have higher backscatter intensities.

As explained later, we will only take data from winter months to train our algorithm, so we assume that our glaciers are covered in snow. However, as this snow is dry and has limited to no melting, the fresh snow covering the entire glacier will be transparent to SAR imagery, thus only revealing snow that has undergone compaction and melting mechanisms from the previous summer surface.

Applied to KNG & HDF Müller (2011) studied microwave penetration in snow and ice with KNG as one of the main case studies. It was found that from low to high altitudes of KNG, C-band σ^0 gets stronger, which could be explained by varying scattering mechanisms along the glacier's slope. By investigating the distance to the dominating scattering source (called

phase center), they found that with altitude gain along KNG, the dominating source goes from surface to volume scattering, which explains the high backscatter intensity. No studies have been conducted to compare HDF facies; hence, a ground truth was not found.

2-3-3 Glaciers of interest

We are interested in two glacier complexes on Svalbard's northwestern coast. The HDF (Holtedahlfonna) complex (HDF) comprises of three glaciers: Kronebreen/Holtedahlfonna, Infantfonna, and Fatumbreen for a total area of 385 km^2. The KNG (Kongsvegen system) covers around 172.8 km^2 and includes Kongsvegen and Sidevegen glaciers (RGI, 2023). The detailed breakdown of each glacier within the complexes can be found in Figure 2-5. The two complexes join one another 5 km before the head of Kongsfjorden, and a large moraine separates them.

Description of HDF: Kronebreen, which corresponds to the lower reach of the HDF complex is a steady fast-flow tidewater glacier and one of the most persistent fast-flowing glaciers in Svalbard (Kääb et al., 2005, Nuth et al., 2012, Van Pelt and Kohler, 2015). Kronebreen presents large crevasses at the surface as shown in the photography 2-7a. Holtedhalfonna sits at a higher elevation and is the principal region where snow and ice accumulate. Infantfonna is a small cirque and, just like Fatumbreen, feeds into the more extensive glacier system.

Description of KNG: In our work, we include Sidevegen as part of Kongsvegen as they share the same glacier tongue and valley. Kongsvegen is a surge-type glacier characterized by cyclical periods of instability and rapid increase of ice flow. Known surges were in 1800, 1869, and the latest in 1948, and are currently at the end of their quiescent phase (Liestøl, 1988a, Woodward et al., 2002). As its bottom is partly to fully frozen, transport of any surplus is not possible down the firn area (Liestøl, 1988b), which supports why Melvold and Ove Hagen, 1998 has concluded that the glacier was heading up towards a new surge as it is building up a reservoir area of ice. At the surface, the glacier is characterized by its smoothness (slopes ranging from 0.5 to 2.5°) with limited crevasses (J. O. Hagen et al., 1993). This glacier's mass does not redistribute due to those limited slopes, so any loss or gain within the accumulation and ablation zones (surface mass balance) correspond directly to elevation changes (Melvold and Ove Hagen, 1998, J. Hagen et al., 1999).

In 1987, the NPI placed nine stakes along the center line of Kongsvegen to annually investigate its surface mass balance (J. Hagen et al., 1999). Their position can be visualised in Figure 2-6. Photographs of the main glacier facies within Kongsvegen are presented in Figure 2-7, ordered from lowest to highest elevation (Personal communication with Jack Kohler). Photograph 2-7b shows the glacier front of KNG covered by a thin layer of fresh snow. The impact of the 1948 surge can be observed by the oblique lines in the glacier ice that correspond to relics of crevasses. Patterns perpendicular to those old crevasses correspond to current drainage channels. Photograph 2-7c corresponds to undulating glacier ice (sub 10 m resolution) surface, which differs mainly from the flat surface found in superimposed ice as seen at stake 6 in photograph 2-7d. This depicts the process of making superimposed ice, as it is clear that meltwater is being refrozen. Finally, firn is observed at stake eight, which is close to the highest elevation of Kongsvegen. Photograph 2-7e shows the relatively smooth

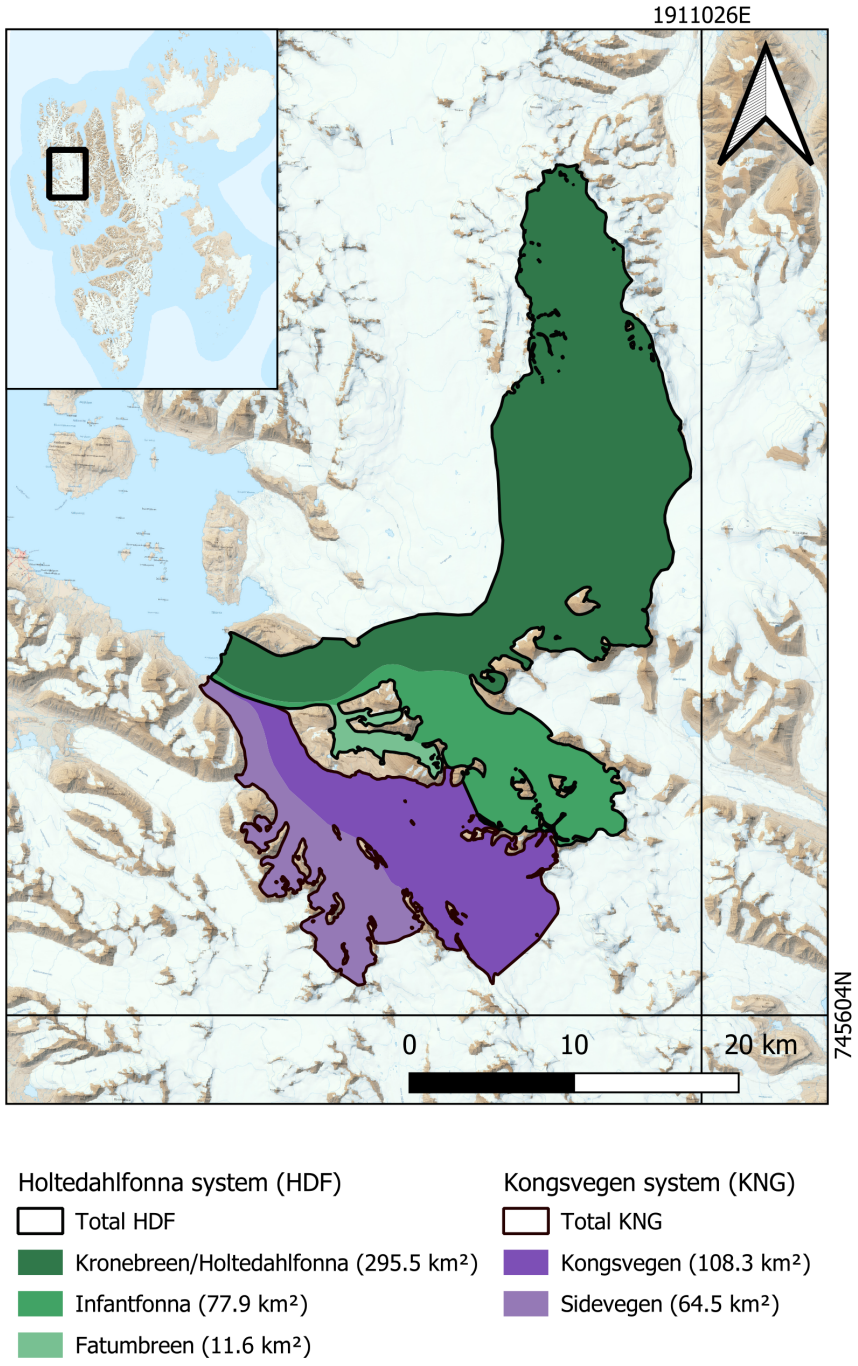


Figure 2-5: HDF and KNG glacier complexes. Each complex comprises multiple glaciers; the areas found through QGIS are detailed in the figure. Base map from Norwegian Polar Institute (2017).

surface of firn covered with some fresh snow. Before 2018, Kongsvegen did not present any crevassing; however, the acceleration of the surge from then onward was fast enough to have them formed, as seen in the photography.

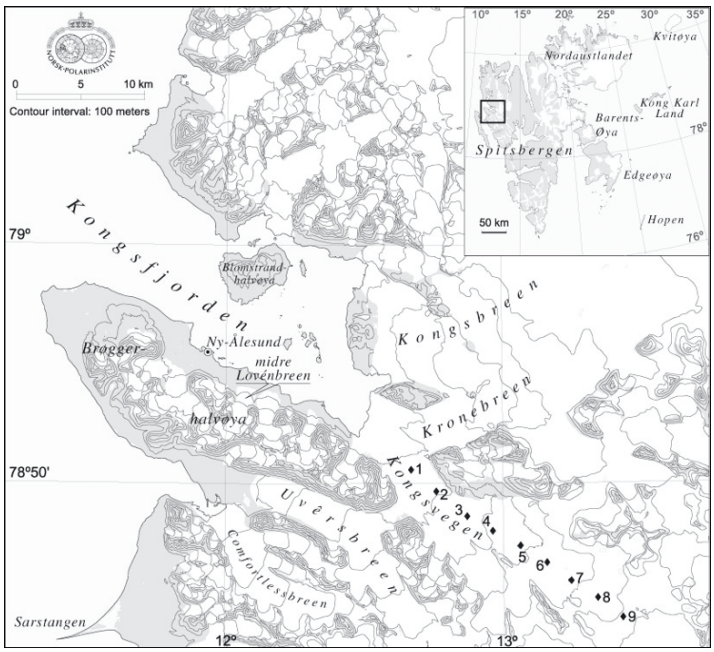
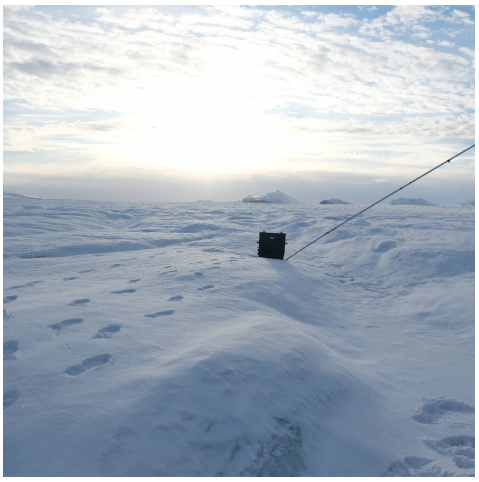


Figure 2-6: Position of the nine stakes placed along the centerline of KNG for annual mass balance monitoring. Source: Hawley et al. (2008).

Depths of glacier zones The depths to which one zone transitions to the other vary according to the glacier and its topography. For KNG, a firn core has been collected at around 11 m of depth at stake KNG8 (Hawley et al., 2008). Using GPR, Brandt et al. (2008) have shown that the depths of SI vary significantly from 0 to 20 m and are shallower next to drainage channels. Anything below firn and superimposed ice will be glacier ice, and this spans surface level at the ablation zone to around 300 m, which is the total depth of the glacier. No depth information was found from previous literature for HDF.

(a) KRB front with heavy crevassing

(b) KNG-KRB front showing crevasse relic and drainage channels



(c) Stake KNG1: Glacier ice

(d) Stake KNG6: Superimposed Ice (SI) being formed

(e) Stake KNG8: firn with newly formed crevasses

Figure 2-7: Photographs at different elevations of the surface of KRB and KNG to visualize glacier facies (Personal communication, Jack Kholer).

Chapter 3

Data and method

3-1 Data selection

In this section, we present the data used throughout the project: subsection 3-1-1 details the glacier boundaries, subsection 3-1-2 the Digital elevation model, and 3-1-3 the SAR products.

3-1-1 Glacier boundaries

The most recent and officially published glacier boundaries can be found in the Randolph Glacier Inventory (RGI) 7.0 dataset compiled by the Global Land Ice Measurements from Space (GLIMS) initiative (RGI, 2023). However, they only have outlines that have not been updated since 2000. Hence, the glacier boundaries used in this work are a draft based on the 2020 Sentinel 2 Mosaic, internally communicated by NPI. A comparison between the two outlines can be further seen in Appendix A-3.

3-1-2 Digital Elevation Model

A Digital Elevation Model (DEM) must be used for terrain correction and calculating local incidence angles.

To avoid capturing sub-pixel variation and outdated fine-scale glacier variations, the DEM from the German Aerospace Center (DLR) was used (Wessel et al., 2018) as it is of lower resolution compared to an internally provided 20m resolution DEM by NPI. If a high-resolution DEM is used, it would be necessary to smoothen it over several pixels to get a pixel-level terrain slope required for the IA effect. TanDEM-X at 90m of resolution was used as it is open access and of lower resolution (Rizzoli et al., 2017). The cells N79E012 and N78E012 were merged using QGIS to cover our glacier region. As per the TanDEM-X data guide, no data values were set to -32767.

3-1-3 SAR products

The objective is to understand the data availability of Sentinel-1 SAR products to choose the optimal dataset for training purposes. Scenes of interest were counted by directly searching the Copernicus database using its API. Downloads of the scenes were done using Creodias, a cloud service for Copernicus, as the Copernicus interface was being updated, which complicated direct downloads.

HDF and KNG glaciers are located close enough to always be within the same Sentinel-1 scene, so the analysis of scene count is equal and relevant for both glaciers. Initially, we prioritized looking at IW data, which offers a higher resolution of backscatter intensities (10*10 m) compared to EW (40*40 m). GRD products were used as they have already been pre-processed in-house by ESA.

All product information covering our glaciers (total from 2014 to 2024) was found and put into a dataframe in subsets to bypass the API maximum limit of 1000 products at a time. Then, they were all concatenated into one main dataframe. The number of scenes was counted and analyzed after creating the appropriate labels for the type of satellite used, the orbit direction, the polarization channel, and the year they were acquired.

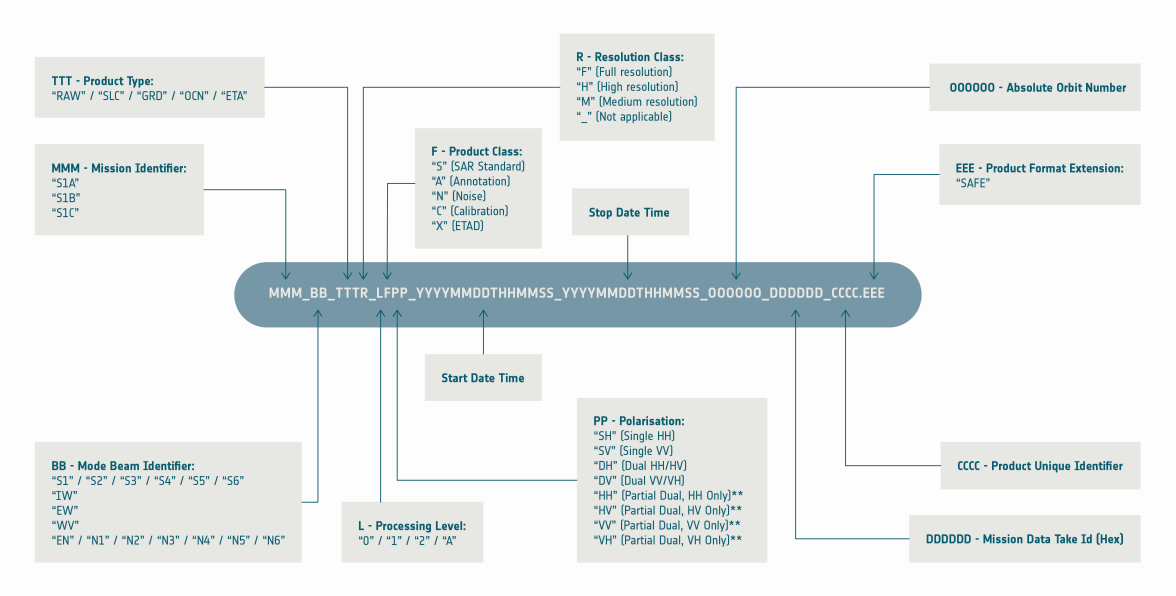


Figure 3-1: Naming conventions of S1 products useful to filter the scenes per satellite, mode of acquisition, polarisation, time of acquisition, and orbit number. Source: Copernicus, 2024

To group the existing scenes, the naming convention of Sentinel-1 products was used to categorize if the scene is from satellite A or B and the specific polarization channel, as they can easily be extracted from the string in the name. Figure 3-1 shows the definition for each string in the product's name. Orbit direction is, however, only included in the metadata, but it can be noticed that descending passes for this region only take place at 6 am (around

0615) and ascending passes at 3 pm (generally 1545). Thus, they were labeled according to the timestamp and included in the product's name.

Relative orbits As we want to observe the per-class IA dependency from the backscatter intensity, we also need to acquire data from a single polarization that contains different relative orbits and covers the same area of glacier zones at various IAs. The product name from Sentinel-1 data provides the absolute orbit number, i.e., the unique count for each complete orbit around the Earth. On the other hand, a relative orbit corresponds to the distinct path the satellite covers, and products of the same relative orbit have similar IA and look directions. In that way, different relative orbits provide a wider range of IA for the glaciers studied. Indeed, different relative orbits imply wider fluctuations in IA, which is necessary to understand the relationship between backscatter and IA. To then find the relative orbit of the different scenes, we can calculate them from their absolute orbit number found in each product's file and use the following calculation (SentiWiki, 2024):

$$\text{For Sentinel-1A:} \quad \text{relative\_orbit} = \quad \text{mod (absolute orbit number} - 73,175) + 1 \tag{3-1}$$

$$\text{For Sentinel-1B:} \quad \text{relative\_orbit} = \quad \text{mod (absolute orbit number} - 27,175) + 1 \tag{3-2}$$

3-2 Classification method

This section provides the methodologies for classifying the training data and the glacier images. Subsection 3-2-1 presents which pre-processing steps on SAR products were applied; subsection 3-2-2 details the method used to evaluate the backscatter dependency on IA and thus apply common and per-class IA slope correction; subsection 3-2-3 describes the supervised learning algorithm and training data used.

3-2-1 Data pre-processing

The SAR products were batch pre-processed using ESA's Sentinel Application Platform (SNAP) tailored explicitly to Sentinel products and SAR processing using the standard flow depicted in Figure 3-2. The pre-processing was applied to both backscatter intensities HH and HV for all products. Range-Doppler Terrain correction was applied with a pixel spacing differing on whether EW or IW was used (40m and 10m, respectively), and the subset of TanDEM-X 90m covering the glaciers of interest was loaded for the correction and calculation of local IAs. No radiometric correction was implemented in the pre-processing stage to understand better the impact of the different corrections on the final classifications.

The terrain correction operator in SNAP had further the choice to output the projected or non-projected local IAs. The non-projected "local incidence angles" values are the IA used throughout this project. For further reading on the differences between projected and non-projected IAs, please refer to section A-2.

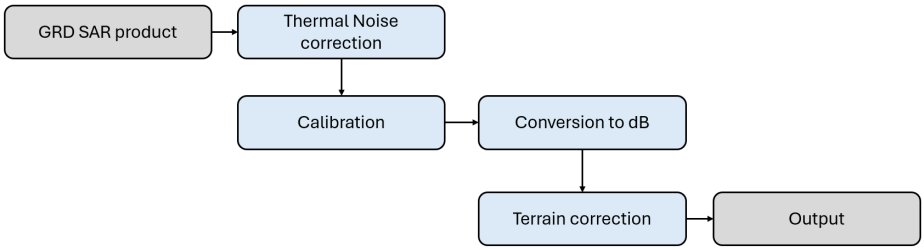


Figure 3-2: Pre-processing graph for all the SAR IW and EW products used in this project.

3-2-2 Backscatter dependency on IA

Common IA slope correction Common IA slope correction is the most widely used approach, which involves applying a mean IA rate correction to the entire backscatter image for IA angles either defined with ellipsoidal angles or local IA. The choice of angles depends on the application and whether topography would be relevant to correct for (e.g., mountainous regions) or not (e.g., sea-ice). While this is practical, it also implies that this correction will over or under-compensate zones with more or less surface roughness within this backscatter image.

Per-class IA slope correction To account for different class roughness and IA-slopes (and so dependency on IA), Lohse et al. (2020) integrated IA as a class property. Using the fact that a linear function can approximate the backscatter intensity in the log domain, the mean for class i can be defined as

$$\mu_i(\Theta) = a_i + b_i \cdot \Theta. \tag{3-3}$$

Substituting this in equation 2-7, we get the following PDF accounting for per-class IA correction

$$p_i(\mathbf{x}|\omega_i) = \frac{1}{(2\pi)^{d/2}|\Sigma_i|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-(a_i+b_i\cdot\Theta))^T\Sigma_i^{-1}(\mathbf{x}-(a_i+b_i\cdot\Theta))}. \tag{3-4}$$

Slopes and intercepts of each training data zone were found by taking the winter months, HH+HV polarization of some data points (4 to 7) of three years of EW data: 2015, 2021, and 2023. Each year was fitted using linear regression, and then the slopes and intercepts for all three years were averaged to get an IA slope associated with each class. The code for the three classifiers can be found on Github here.

3-2-3 Algorithm implementation

The GLIA's fitting phase uses the training regions to determine the class-specific slope and intercept, b_i and a_i defined in 3-3. When designing the "predict" phase of the algorithm, it was found to be more computationally efficient first to remove the variable mean to calculate a zero mean deviation, compared to directly computing equation 3-4. In other words, for

class i, deviation d_i from the mean for $\mathbf{x}$, and for the class-dependent slope and intercept b_i and a_i respectively, we have

$$d_i = \mathbf{x} - \mu_i = \mathbf{x} - (a_i + b_i \cdot \Theta).$$

This could be done as a block array calculation. Then, the standard PDF of the multivariate normal distribution of d is calculated, and the decision rule presented in equation 2-6 is applied to assign a label to the pixel.

Training data The training data from which we calculate the IA slopes correspond to winter EW products of 2015, 2020, and 2023 from January to April. The labeled pixels correspond to the regions defined in Figure 3-3. All regions except firn have been outlined in the Kongsvegen glacier complex as the different glacier zones are easier to differentiate than Holtedahlfonna. Indeed, possibly due to lower elevations, the firn training region of KNG had too much variation in the backscatter intensity, so we decided to use the IA slope from the HDF firn. In the georeferenced coordinate system, these training regions are each 1 km*1 km. For EW products, this corresponds to 25*25 pixels.

The classification accuracy was analysed through confusion matrices, which provide a percentage of correctly classified classes from the true labels to the predicted ones.

To monitor the firn line while preserving edges, the classified products were smoothed with a median filter of 5*5 pixels. Then, the separation between the classes corresponding to firn and superimposed ice facies was traced by visual inspection of the results in QGIS. To quantitatively understand the extent to which there was a decrease or increase in pixels labeled in the SAR class corresponding to firn, two outlines were traced in QGIS to cover the firn region of Holtedahlfonna and Konsvegen. The tool "zonal histogram" was used to get the proportion of labels of pixels in the regions, and the decrease percentage was calculated over the years.

3-3 Texture feature

This section provides information on the methodology used to incorporate a texture feature in the algorithm.

Crevasses are zones of stretching and compression that result in large fractures in glacier ice. Due to these severe roughness and slope variations, the backscatter in the crevasse regions has small geographical scale variations that do not correspond to the glacier ice signature that would be found elsewhere. Since the mean backscatter intensity of the crevassed glacier ice is similar to that of firn, additional information is needed to separate the two. Visually, we can distinguish the two facies in the training images because of their different textural signatures. Hence, adding some texture features to each pixel of the glacier image would provide a simplistic way to consider it.

One of the most straightforward and computationally fastest ways to incorporate textural information is to use the standard deviation of the backscatter intensity within a given

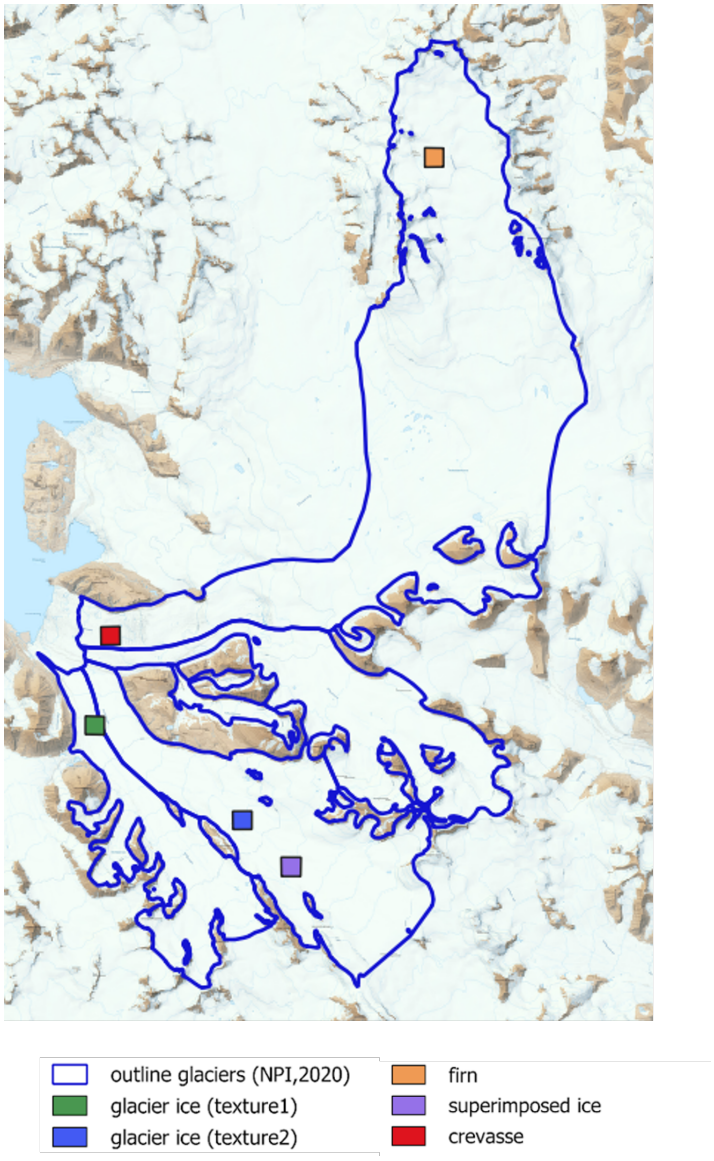


Figure 3-3: Training regions' locations on KNG and HDF.

window. In order to include standard deviation in the GLIA algorithm, we compute the standard deviation of HH and HV backscatter intensities across a window over the whole image (Figure 3-4). The mean and standard deviation of the standard deviation of the central pixel were calculated for window sizes on each glacier class training region for both EW and IW. Both standard deviations are then added to the features of the central pixel. In this way, now pixel images contain four channels:

pixel_channels = [HH, HV, std_HH, std_HV].

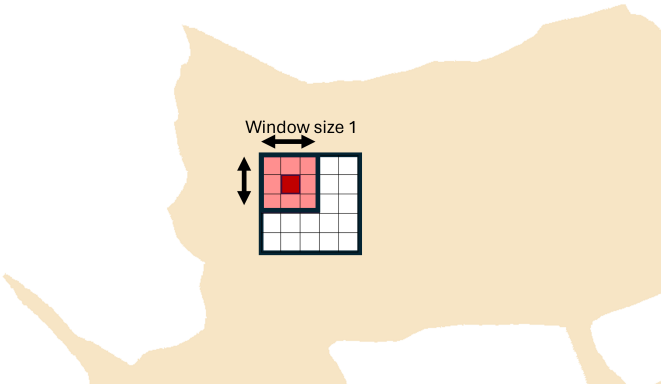


Figure 3-4: Visualisation of the standard deviation calculation. A window (here of size 1) travels through the glacier image, calculating the standard deviation for the center pixel. To get reliable results of standard deviation metrics (mean and std) for each training region, this process is done multiple times over a larger grid and then averaged.

By plotting the HH_mean vs the HH_std for each class, we can observe whether the crevasse training data points overlap significantly or not with other classes and then decide accordingly on the optimal window size, i.e., the window size, which is the minimum in size to be accurate while still allowing crevasse training points to be significantly different from other classes. This was done for both EW and IW training regions for windows of sizes 3, 5, 7, 11, 15, and 21 (pixel count), as seen in Figure 4-11a and Figure 4-11b respectively. The standard deviations were calculated for EW on a product of 16/01/2023 and for IW on a product taken on the 10/01/2023 product.

Once we can observe if adding the standard deviation (std) of the backscatter as an additional feature improves the separability of classes, we check if this measure satisfies the linear assumption with respect to the local IA to be able to include it in the GLIA algorithm directly. For this, we take the 27 EW products of varying relative orbits in winter months to get a wide local IA spread. Those products cover the years 2015, 2020, and 2023, as these are the only years available in EW mode over this area. We run the optimal window size found previously through the training regions and calculate the std of the mean HH backscatter intensity of the center pixel. Then, the std of all pixels are averaged per training region to finally get per scene a mean value of the std of the HH mean, named “mean_HH_mean_of_std”. By plotting this mean value for different scenes of different years as a function of the mean IA, we can observe whether the std metric fits the linear assumption necessary for the use of GLIA.

Chapter 4

Results

4-1 Data selection

This section provides the results of the data search. Subsection 4-1-1 counts the amount of available data per product type. Subsection 4-1-2 shows the backscatter variability over the months to select a time window of product acquisition. Finally, subsection 4-1-3 shows the data available in terms of relative orbits for both IW and EW, which motivates our final selection of SAR product types.

4-1-1 Data availability

For GRD products, we can filter scenes by the satellite used (S1A, S1B), polarization channel (single: HH, VV, dual: VV+VH, HH+HV, and orbit direction (ascending or descending). Any files available on Creodias finishing by '.COG' were excluded from future analysis as SNAP cannot pre-process them. For EW data, 1472 products are available from 2014 to 2024, covering our glaciers with a non-zero count for all years. 1452 are HH+HV; from that, 928 are from January to April. Those 928 products were distributed over 2015, 2020, 2023, and 2024, but downloads using the Creodias platform for 2024 products were unsuccessful. For IW data, we found 660 HH+HV products. IW products generally span from 2015 to 2024; however, HH+HV and winter products are only available from 2017 to 2024, excluding 2019. However, again, 2024 products were not available to download using Creodias.

4-1-2 Months

Due to snow melt and variation of liquid water content within the snowpack and on the ice surface, the backscatter intensities will vary between seasons. To train our algorithm on the most consistent dataset, we have analyzed to what extent the seasons affect the separability of glacier regions when looking at the backscatter intensities vs. the local IAs. Our study

focuses on the local IA from the glacier's subsurface, meaning we want the most negligible impact of surface changes.

NPI mentioned that from September onwards, we should expect the least melting and fresh snow, and that is when they generally take their reference field measurements. Looking in more depth at pre-processed backscatter intensities for one year (arbitrarily assumed to be 2021 here) in Figure 4-1, we can observe that the winter months from January to April offer minor variability, so we decide to filter out any other months for our study (HV values have the same trend and can be found in the appendix A-1).

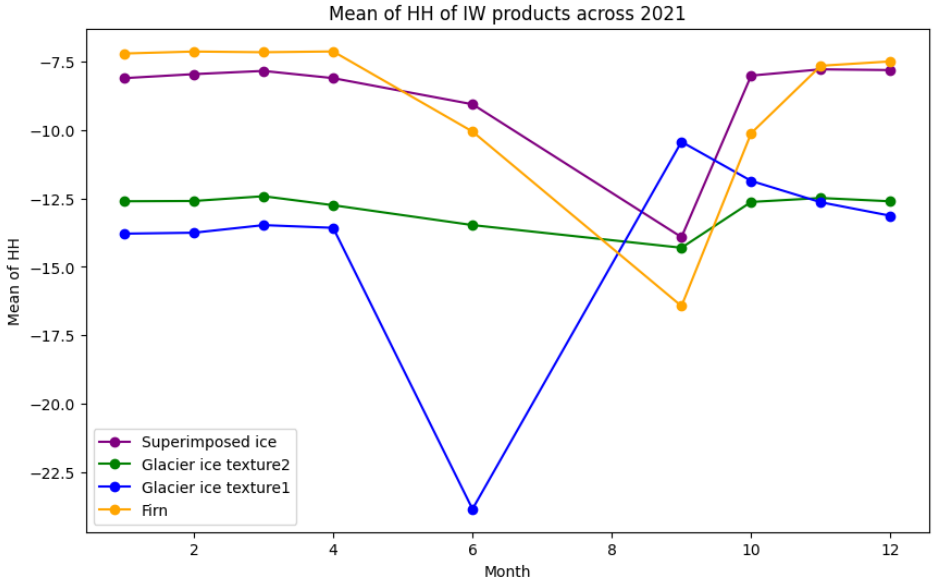


Figure 4-1: Mean backscatter intensity for the HH channel for IW products spanning our glacier complexes across 2021. All intensities are stable within 1 dB from January to April when the least melt and refreezing occurs.

For this data, the HH and HV bands for the EW data of each training region are plotted in Figure 4-2. The two different textures of glacier ice heavily overlap in both bands and in general, the separation in both bands is roughly equivalent (removing the outliers).

4-1-3 Relative orbits

Unfortunately, only one relative orbit number is available for HH+HV polarization for our IW data covering our region of interest. This insufficient spread of IA prevents us from observing the relationship between IA and backscatter values.
For this reason, we use EW data, which has a lower spatial resolution but more variability in its relative orbit over our glaciers. As mentioned above, only the years 2015, 2020, and 2023 can be used for the HH+HV winter products, and all have at least 11 different relative orbits.

Ideally, we would be able to train the classifier on EW data as it contains the different relative orbits, but to test the classifier using IW data as it has higher resolution. For this, we would need to ensure that IW corresponds to EW data, which we would theoretically expect as they are calibrated to do so. Figure 4-3 plots the IW data over EW one. When plotting the HH

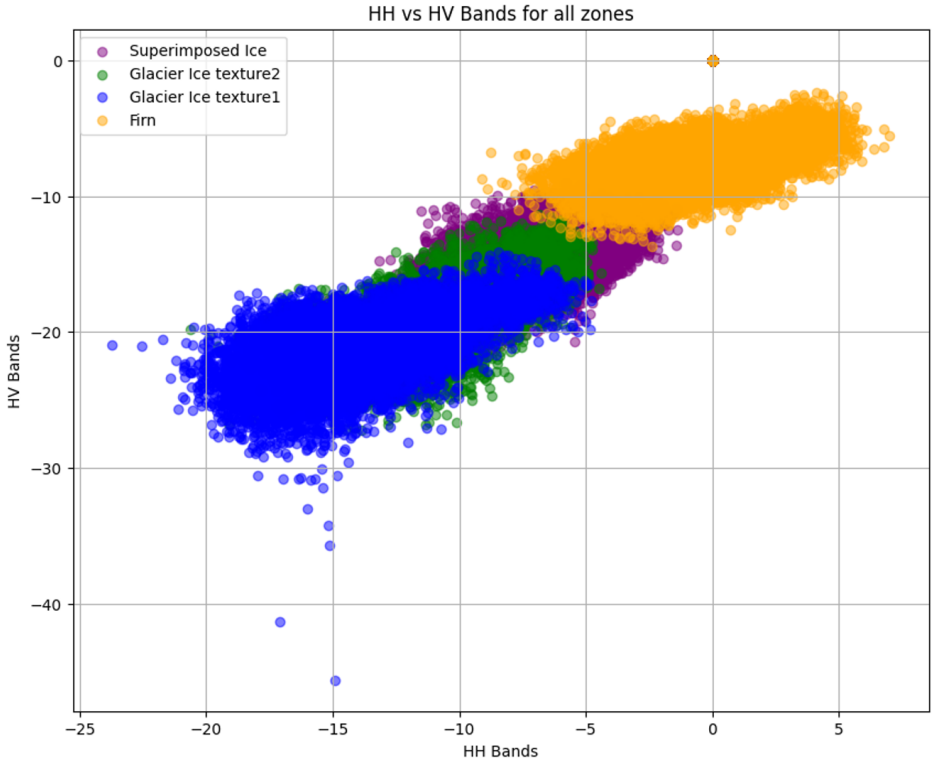


Figure 4-2: Backscatter intensity spread of HH and HV bands of all training regions.

backscatter intensities per class, IW data falls in the same range as EW, which points to what we expect. However, the similarity between the backscatter values of IW and EW should be more rigorously checked in future research.

4-2 Classification

This section presents the results of classification on the impact of integrating local IA for separating the glacier zones (subsection 4-2-1), classification results on training data and training accuracy calculations (subsection 4-2-2) and finally classification on our glacier complexes and glaciological interpretation (subsection 4-2-3). In Figure A-4 in the Appendix, one can see an example SAR scene where SAR zones can be distinguished by their backscatter intensities.

4-2-1 Backscatter dependency on IA

Looking back to the range of backscatter intensities in Figure 4-2, we can observe that some backscatter values have non-unique associated zones, specifically a large overlap between superimposed ice and glacier ice. If we add the other dimension of IA and fit a line as seen in Figure 4-4, we can observe a linear relationship between the backscatter values and IA spread. We also observe that 1) the linear relationship of each class is more separable than

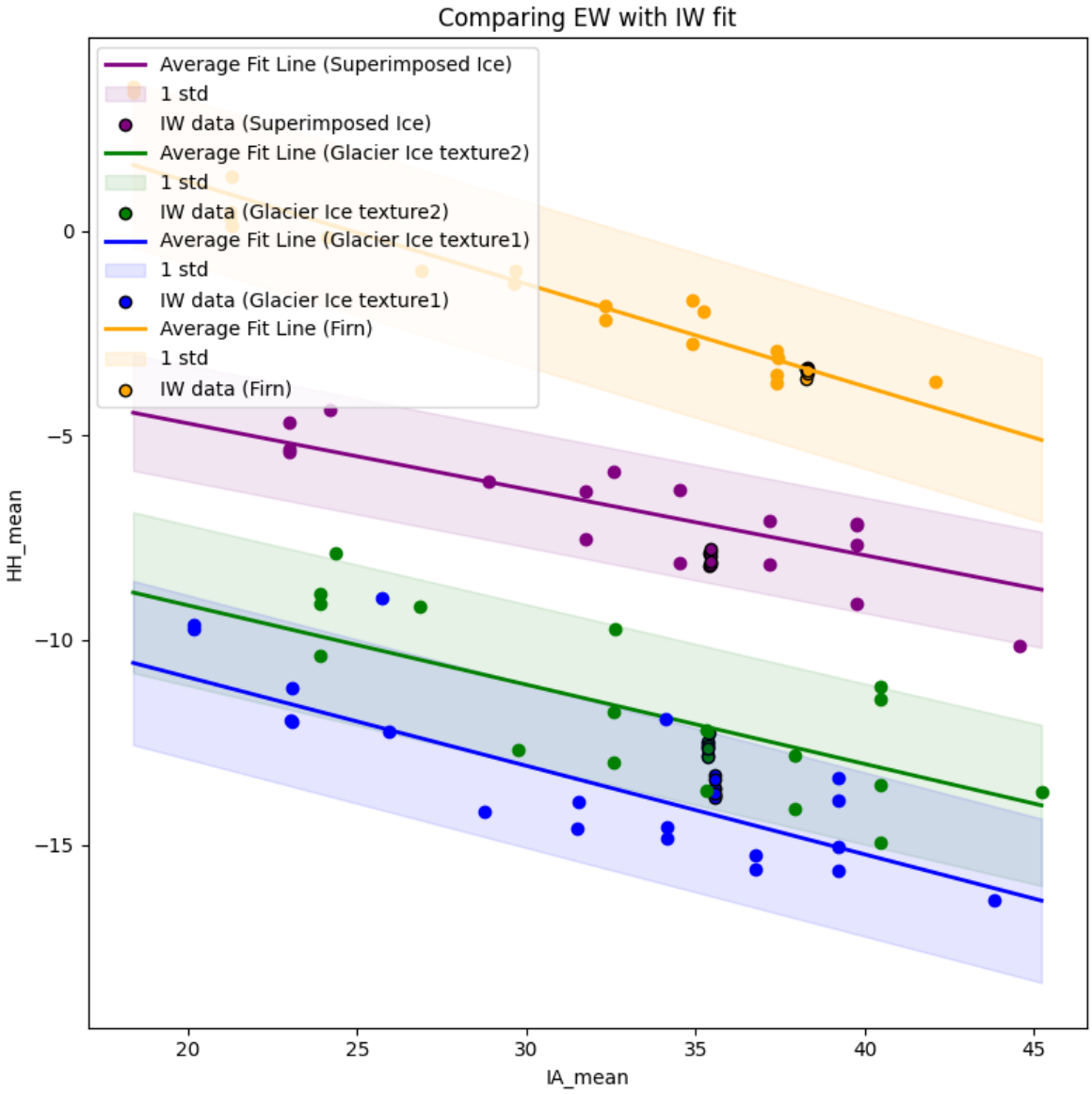


Figure 4-3: Plotting the backscatter intensity of the HH band versus the mean local IA for both EW and IW over the whole image. IW data does not have a large variation in mean IA as there is only one relative orbit, but we can observe that IW data for each zone fall into the corresponding EW fitted line.

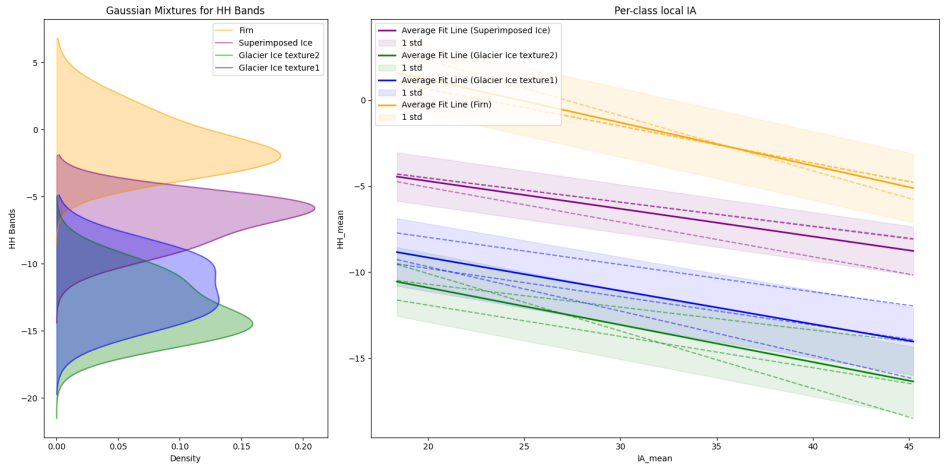


Figure 4-4: Comparison of HH backscatter distributions of each class (left) with their 2D separability when plotted as a function of incidence angle (right).

without IA, where the Gaussian mixtures significantly overlap, and 2) the classes are not strictly Gaussian, which will also affect classification accuracy. The steepness of the slope is affected by two key factors: the dominant scattering mechanism and the presence of small-scale surface roughness. The slope will be the least steep for a volumetric scattering as the diffusive reflection will occur independently of the IA. This contrasts with surface scattering, where a larger IA will imply a larger separation between the main reflection and the received beam. For high surface scale roughness, there will also be a similar observation that for smoother surfaces (low small-scale roughness), the reflection will be stronger further than the received beam for a large IA (see in the theory section Figure 2-3 for visualizations). The IA slopes are given as shown in Table 4-1.

Zone	**Slope [dB/°]**
Firn	-0.25
Glacier Ice texture1	-0.20
Glacier Ice texture2	-0.19
Superimposed Ice	-0.14

Table 4-1: Slope of HH band vs IA calculated from the different training regions and averaged.

Slopes and intercepts of each zone were found by taking the winter months, HH+HV polarization of some data points (4 to 7) of three years of EW data: 2015, 2021, and 2023. Each year was fitted using linear regression, and then the slopes and intercepts for all three years were averaged. Figure 4-4 indicates that the average of each zone can be distinctly identified and separated from the other zones through this linear model. There is, however, a significant overlap when taking into account 2 or 3 standard deviations, which would mean that the large variability within each zone will hinder the accuracy of a classifier.

Overall, our zones are more separable when looking at the backscatter values paired with IA than when looking at backscatter values alone. This motivates adding IA per class dependency

in supervised learning, as we would expect a higher accuracy compared to without.

Once the relevant data was selected, we trained and tested the three algorithms to understand how much GLIA would improve the classification accuracy. This section presents these results. In subsection 1, we compare the initial data used and classifiers on the training data and then on the glaciers of interest. In subsection 2, we refine our results by adding the crevasse class, which was previously not considered.

4-2-2 Classification of training regions

Figure 4-5 presents the confusion matrix of the three algorithms run on the training regions. The matrices present the true labels in the i-th row and predicted labels in the jth column. The values correspond to the number of predicted labels/true labels, and thus, the better the accuracy, the closer to 1 the ratio becomes, and the more the matrix tends to be a unitary one.

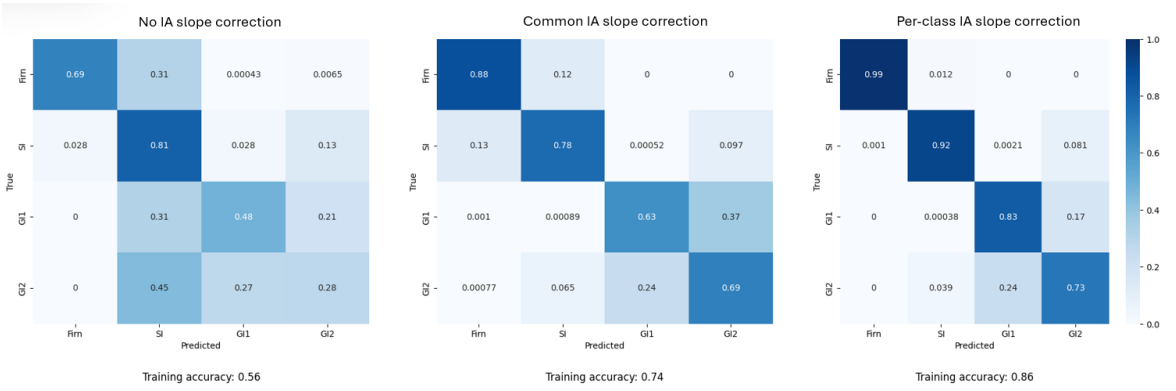
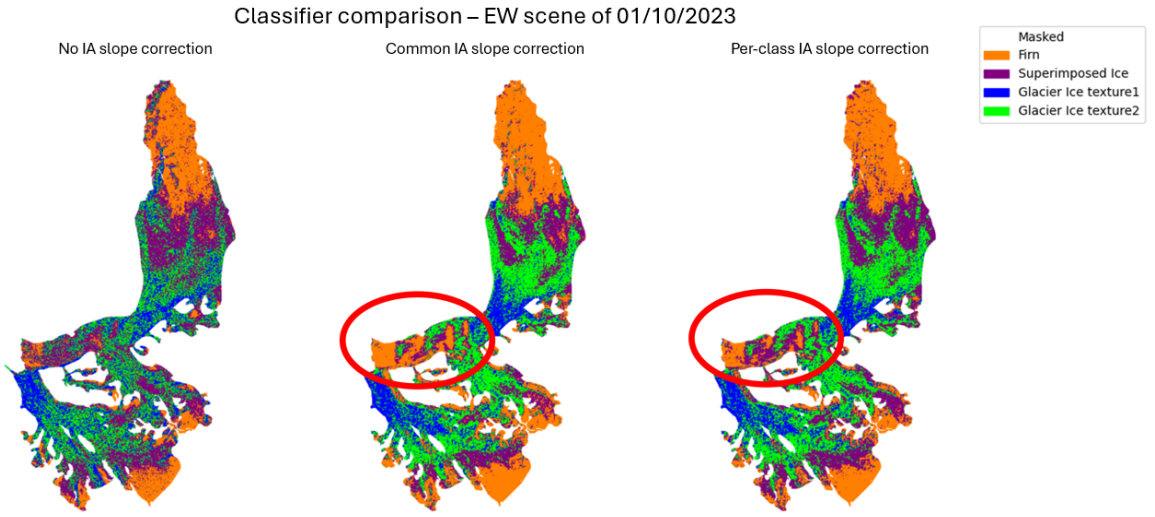


Figure 4-5: Confusion matrices for training regions comparing Gaussian classifier (left), common IA slope correction (center), and per-class IA slope correction (right). SI abbreviates the Superimposed Ice zone, and GI1 & GI2 correspond to the zones Glacier Ice with texture1 and Glacier Ice with texture2, respectively.

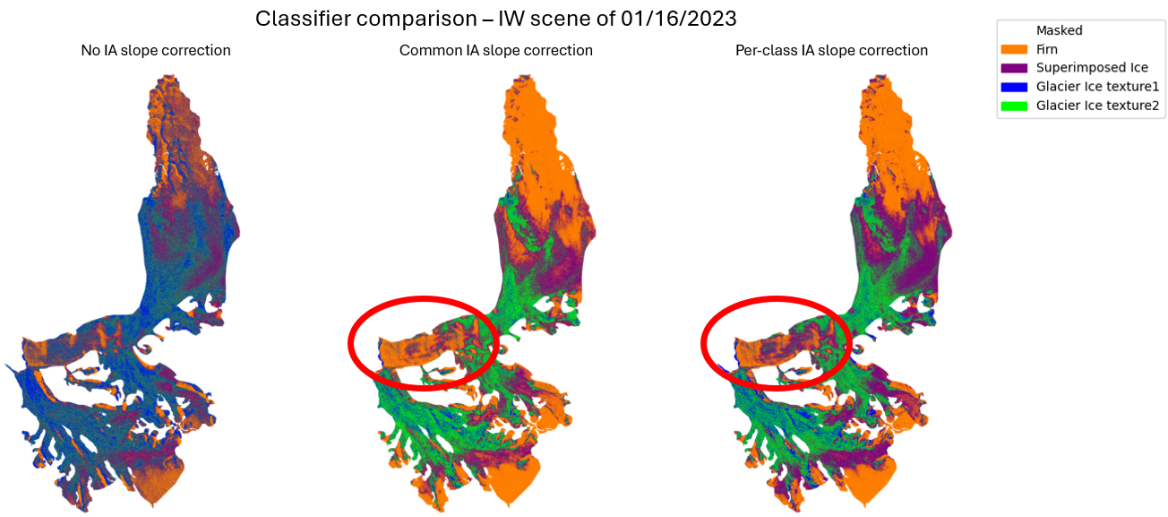
The matrices show a generalized improvement from the Gaussian classifier (average training accuracy of 56%) to the one integrating common IA slope correction (74%) and the one integrating per-class IA slope correction (GLIA) (86%). In particular, the common IA slope correction improves the separability between superimposed ice and glacier ice. The Gaussian classifier seems more biased to classify superimposed ice as firn, while the misclassification between those two classes is balanced in the common IA slope algorithm. GLIA can significantly reduce this misclassification thanks to the firn class having a steeper slope than the other classes, as seen in Figure 4-4. With this algorithm, the two different glacier ice textures are still misclassified as one another, which is expected from the scatter plot of backscatters, as these classes significantly overlap (Figure 4-4) and do not have significant slope differences (Figure 4-4).

4-2-3 Classification of glaciers

Figure 4-6 shows the SAR classification mapped to glacier zones on HDF and KNG.



(a) EW



(b) IW

Figure 4-6: Comparison of Gaussian classification without any correction (left), common IA slope correction (center), and per-class IA slope correction (right) for both EW (top) and IW (bottom) products.

From those results, the common IA slope correction and the per-class IA correction algorithms provide a more realistic and accurate zonation of glacier facies, where the firn zone is present at the highest altitudes, then superimposed ice, then glacier ice towards the lower ends of the glacier complexes. Common IA slope and per-class IA corrections provide similar clustering with more distinct zones for GLIA. In this study, there is no stronger advantage in using GLIA for the glaciology interpretation of zones.
Between the EW and IW results of common IA slope and per-class IA slope correction, the two different glacier ice textures are misclassified, as we should expect zonations like those seen in EW. The difference could be explained because IW is higher resolution, so potentially, the slopes identified early on EW data do not adequately match those found in IW as seen in Figure 4-6b.
We have checked whether two SAR products close in time have similar classifications to motivate reliable results for time series. Figure 4-7 compares one scene from the 4th of February 2017 and one 12 days later. The short timespan between them makes it highly unlikely that glacier classification results should differ, and indeed, we can observe matching labeling of firn and superimposed ice. Intensity change in the HH band can be visually observed, pointing to some physical change in the properties of the subsurface. Looking at the weather data from Yr at Ny-ålesund base station, between the 5th and the 12th of February, there have been positive temperatures (0°-4°C) and a cumulative of 79.4 mm of rain (**yr˙2017**). It is, hence, possible that the glacier ice regions have also been affected by rain on these hotter days.

Firn line monitoring 2017 to 2023 The firn line for years 2017, 2020, and 2023 were mapped and visualised in Figure 4-8. For HDF in 2017 and 2020, feature 1 in the figure corresponded to pixels classified with the zone corresponding to firn and directly connected feature 2, creating a firn line that reached around 3 km lower than where the firn line can be observed in 2023. Additionally, Feature 3 has basically disappeared, with only a few sparse pixels classified as firn compared to a much denser cluster for 2017 and 2020. For KNG, its firn line has retreated more homogeneously by around 500m along the perpendicular to the slope, with the largest retreat on its North side with around 2 km of retreat.

Regarding pixel label count, there is a 20 % decrease in firn labels for KNG and 25 % for HDF. This accounts for all pixels labeled as firn within the defined region of interest as specified in the methods section and hence also would consider misclassified pixels. However, while one cannot take those values as an explicit quantitative measure of firn reduction, they highlight a substantial decrease in both glacier regions.

Crevasse region An issue for both algorithms is that we do not expect firn at the south tip of Holtedahlfonna (lower altitude), which could be explained by crevasses being there. We further refined the data on which the algorithm was trained to improve the accuracy from a glaciology perspective. To account for the mislabeling of the crevasse region, we added a training class of the same size, 1 km * 1 km, as the others and included it. Figure 4-9 shows that IA dependency of the HH bands for crevasse would be averaged as a flat slope (at $+0.01$ dB/°), which indicates volume scattering being dominant in such heavy large-scale roughness surface. However, it will still be helpful as it differs from the negative slope found for both firn and superimposed ice, so we can still expect an improvement in accuracy.

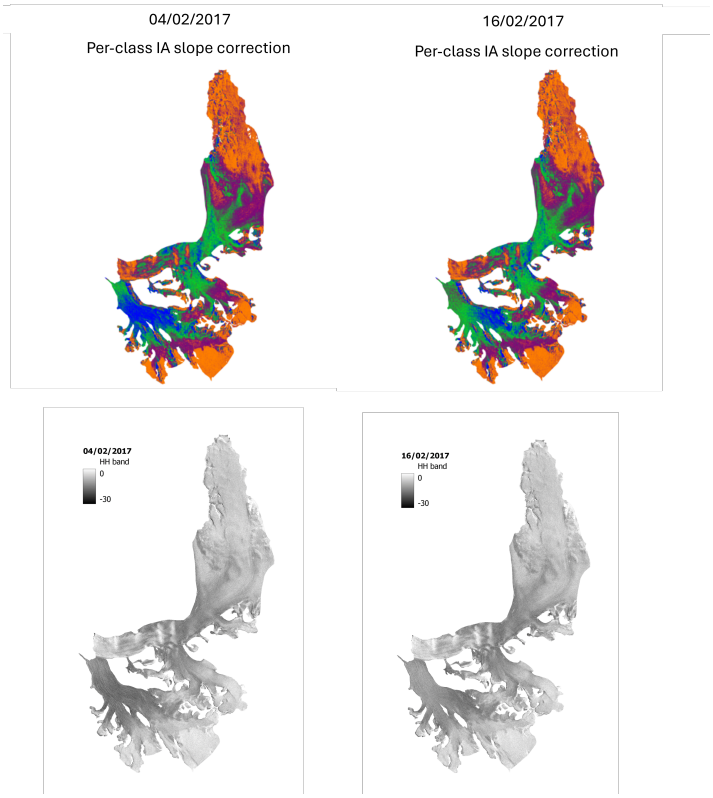


Figure 4-7: Comparison between two SAR products close in time to show that the classification is reliable over time. For the 12-day time difference, we do not expect the glacier facies to drastically change position, and this is what we observe for firn (orange) and superimposed ice (purple). The two glacier ice textures (green and blue) show the most differences between the two product classifications. This correlates with the two texture ice distributions greatly overlapping and some visual differences observable in the backscatter intensities.

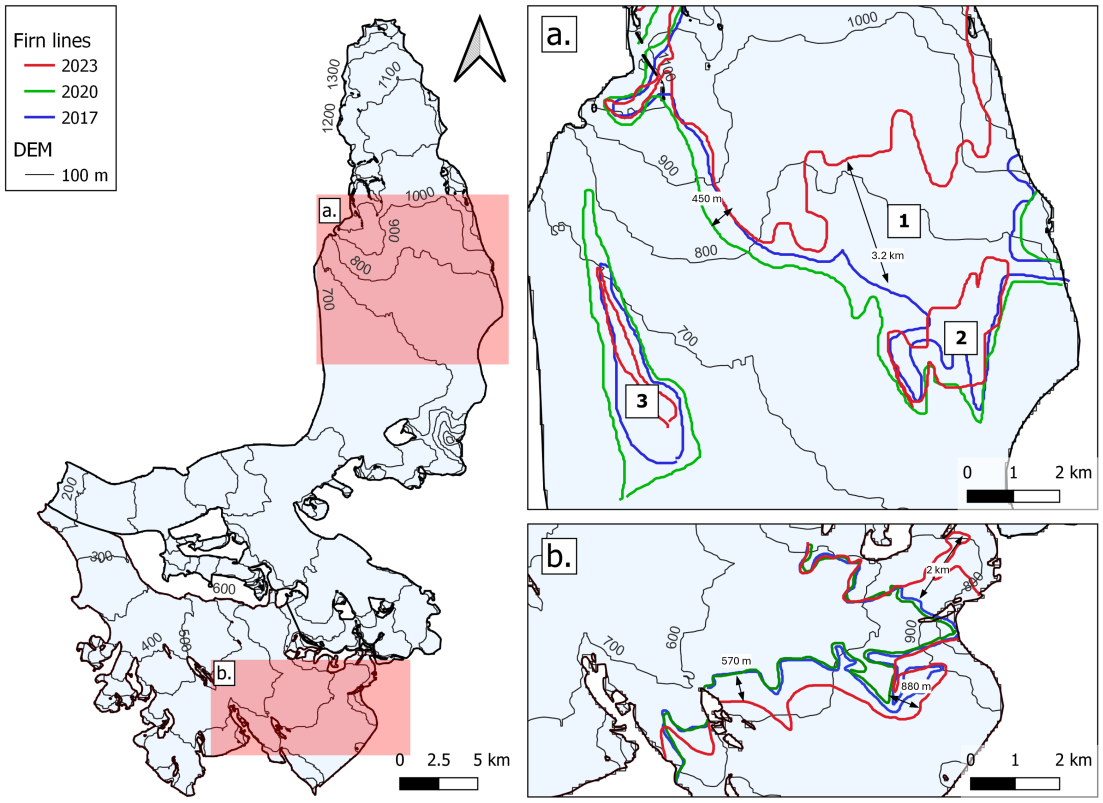


Figure 4-8: Firn line comparison from 2017 to 2023 for HDF (a.) and KNG (b.). A homogeneous retreat of around 570m can be observed for KNG perpendicular to its slope. For HDF, feature 1 is not classified as firn in 2023, and feature 2 & 3 are more sparse

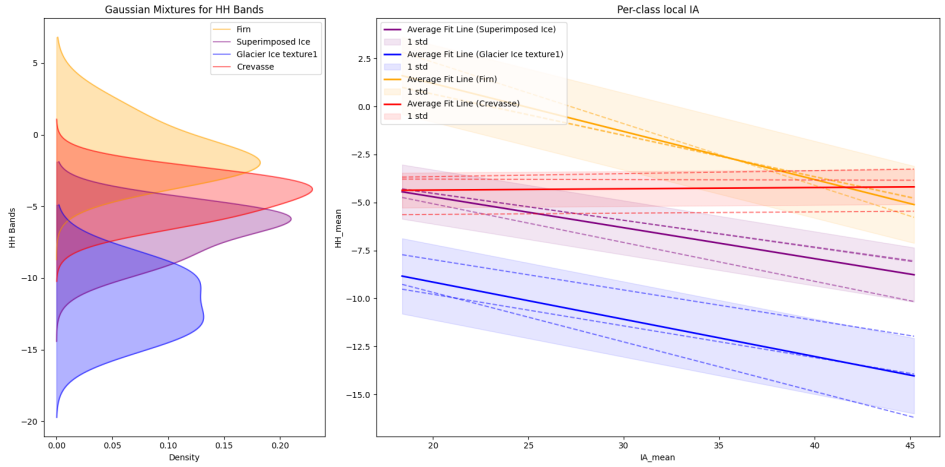


Figure 4-9: Separability of the four regions firn, SI, glacier ice, and crevasse with (right) and without (left) IA considerations. The crevasse distribution significantly overlaps with the firn and SI distributions, considering the backscatter intensities. However, its IA slope is positive, which makes it more separable when correcting for per-class IA.

The training accuracies using the same three classifiers are calculated and shown in Figure 4-10a. From these results, the overall training accuracy diminishes due to the crevasse region backscatter data overlapping with firn and superimposed ice, as seen in Figure 4-9.

(a) Adding crevasses **(b)** Removing GI2

Figure 4-10: Confusion matrices over the training regions comparing training accuracy for a differently fitted algorithm.

It is also evident from the previous subsection that the two textures of glacier ice need to be more distinct in our backscatter data to provide separable classes. To keep the Gaussians between zones with the least overlap in distributions, glacier ice of texture2 was removed as it overlaps the most between glacier ice of texture 1 and superimposed ice. This outputs the confusion matrix presented in Figure 4-10b.

As the backscatter data is not enough to characterize crevasse and separate it from the other classes (specifically firn and superimposed ice), we investigate how we can extend our algorithm to include a texture feature.

4-3 Texture feature

Finally, this section presents the results of incorporating the texture feature using standard deviation into the classification algorithm. Subsection 4-3-1 presents the results from the attempt to find the optimal window size to perform the standard deviation calculation; subsection 4-3-2 checks if this new feature fits the linearity assumption between the metric and the IA spread, necessary to apply the GLIA algorithm; subsection 4-3-3 provides the classification training results; and finally, subsection 4-3-4 shows the new classification on the glacier complexes.

4-3-1 Window size analysis

We can observe from the graphs of Figure 4-11 that in EW, there is no significant overlap in general between classes and that increasing the window sizes effectively increases the distance between the center of the clusters. On the other hand, in IW, there is a significant overlap of distributions for firn and crevasse. The greater the window size, the greater the standard deviation of the center of the crevasse cluster compared to the one of firn. Firn and Superimposed ice clusters become tighter when adding this standard deviation dimension. Even if the maximum window size is reached, the standard deviation feature does not entirely separate firn and crevasse. However, it provides a better separation for crevasse and superimposed ice. Compared to EW, a larger standard deviation can also be observed for the IW mode. This can be explained by the products being acquired and pre-processed by ESA with different amounts of multilooking. For the number of looks per range * azimuth, IW has 5*1 looks and EW 6*2. As EW has more significant amounts of multilooking, the averaging will lower the standard deviation compared to IW products.

Note also that the two products compared are from the same training region; however, they were acquired at different times and days and would have two different radar IAs.

4-3-2 Linearity with local IA

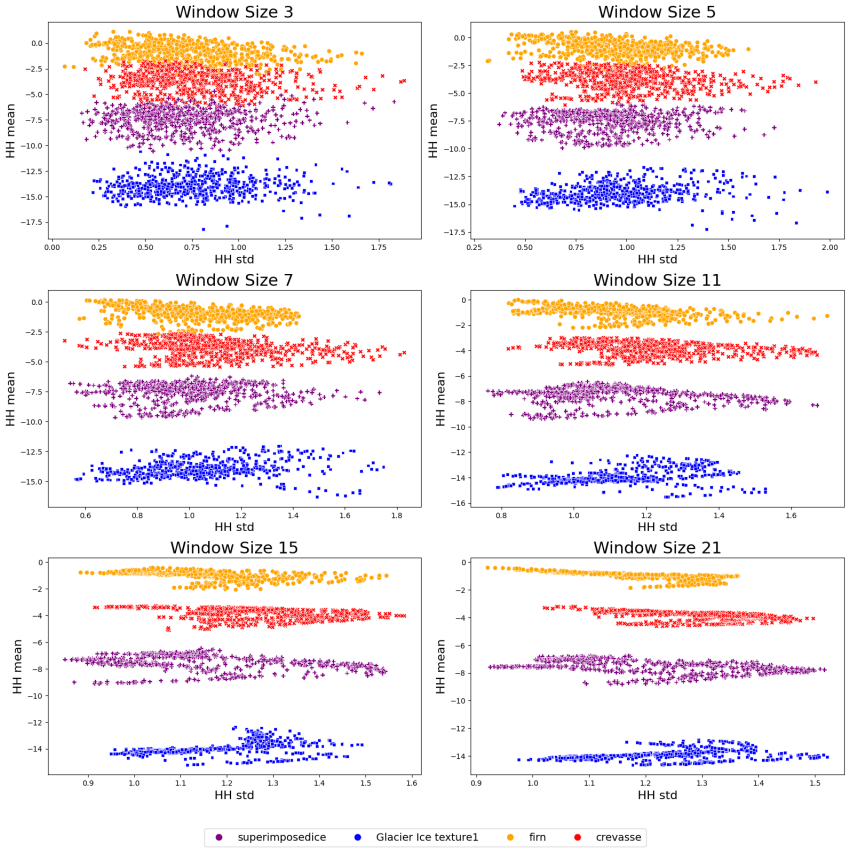
Observing the slopes in Figure 4-12, the data for 2015 and 2020 fits the linear assumption. The lack of lower IA values hinders the linear regression applied to find the line of 2023, and thus, the lines seem to not follow the same trends as the 2015 and 2020 data. However, it can be noticed that if we assume a similar intercept as the global average line for the 2023 fitted one, then the data points would fall within the plausible range of the globally averaged line. The final IA slopes used to correct the data in the learning phase of the algorithm for the std feature were taken to be the average slope between the 2015 and 2020 data.

4-3-3 Classification result on training data

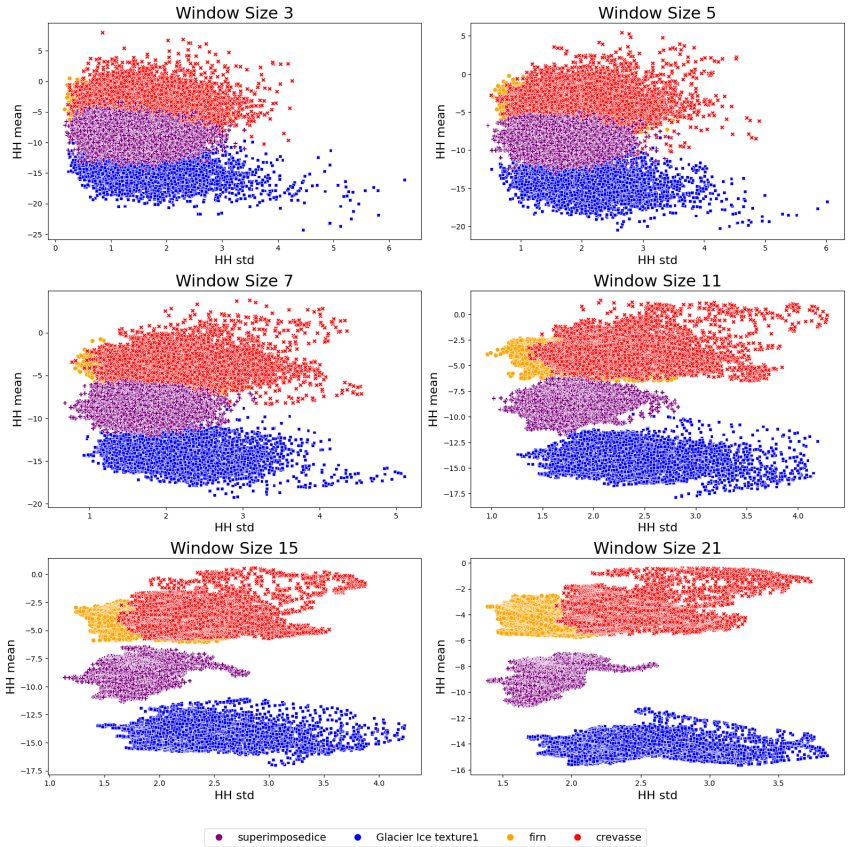
We can observe from comparing Figure 4-13b, which includes the standard deviations features, and Figure 4-10a, which only has the backscatters, that there is an increase in overall accuracy by 10% for the classification of the crevasse training region. This is due to decreased mislabelling of crevasses into firn and SI and less glacier ice being labeled as firn.

4-3-4 Classification results on glacier

Runtime in classifying EW products with the std feature is 2 min, and around 25 min for IW products. When applying the fitted algorithm on our glaciers, the firn region is mislabelled with crevasses, even more so than without the standard deviation feature. This can be understood by the distribution overlap between those two classes in backscatter and their standard deviation.



(a) Window comparison EW



(b) Window comparison IW

July 29, 2024

Figure 4-11: Window size comparison of the std calculation for EW and IW. The objective is to find the minimum size window so that each class is the most separable.

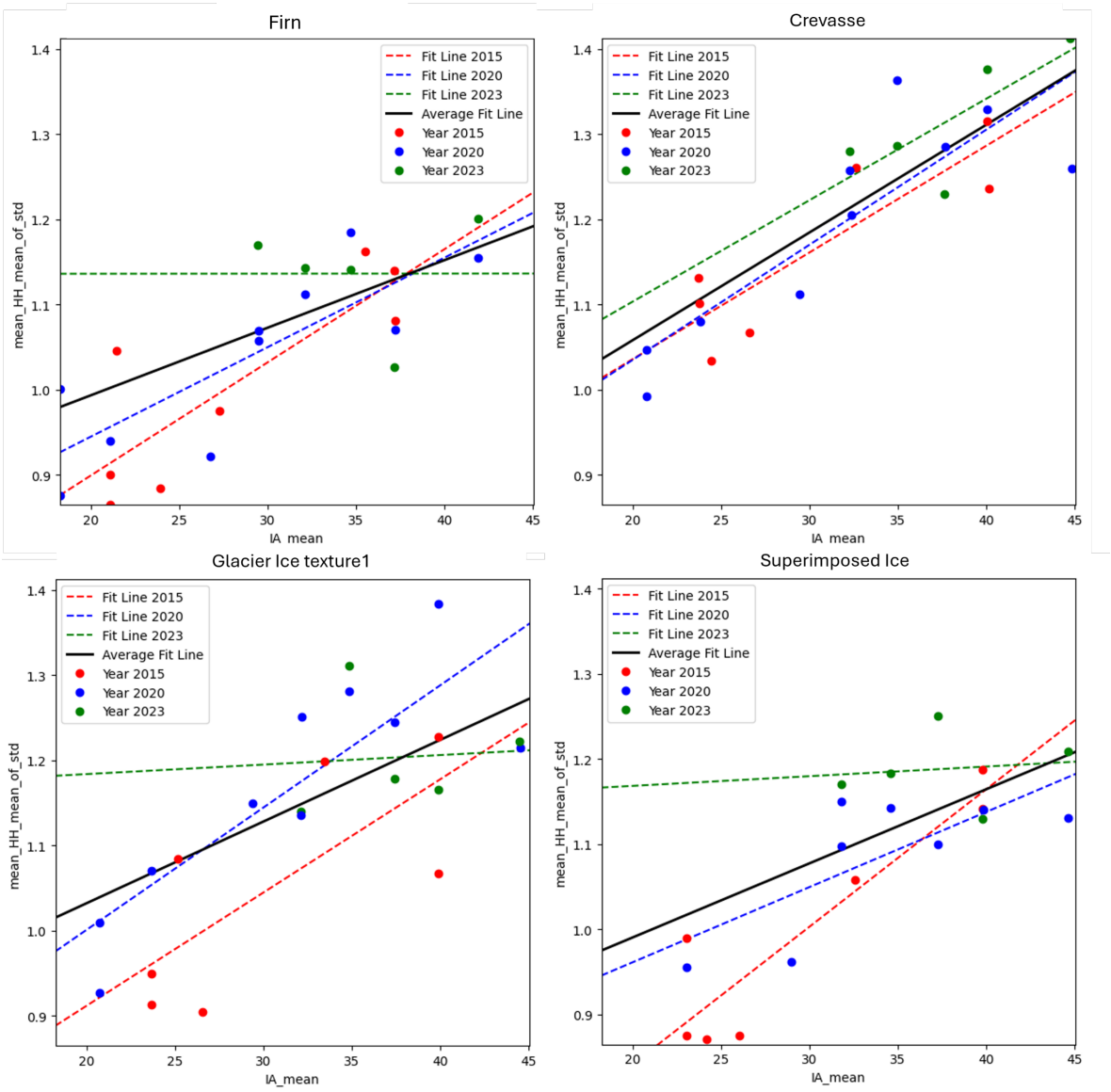


Figure 4-12: Calculation of the mean, standard deviation of the HH band backscatter intensity averaged across training regions versus the local IA for years 2015 (red), 2020 (blue), and 2023 (green).

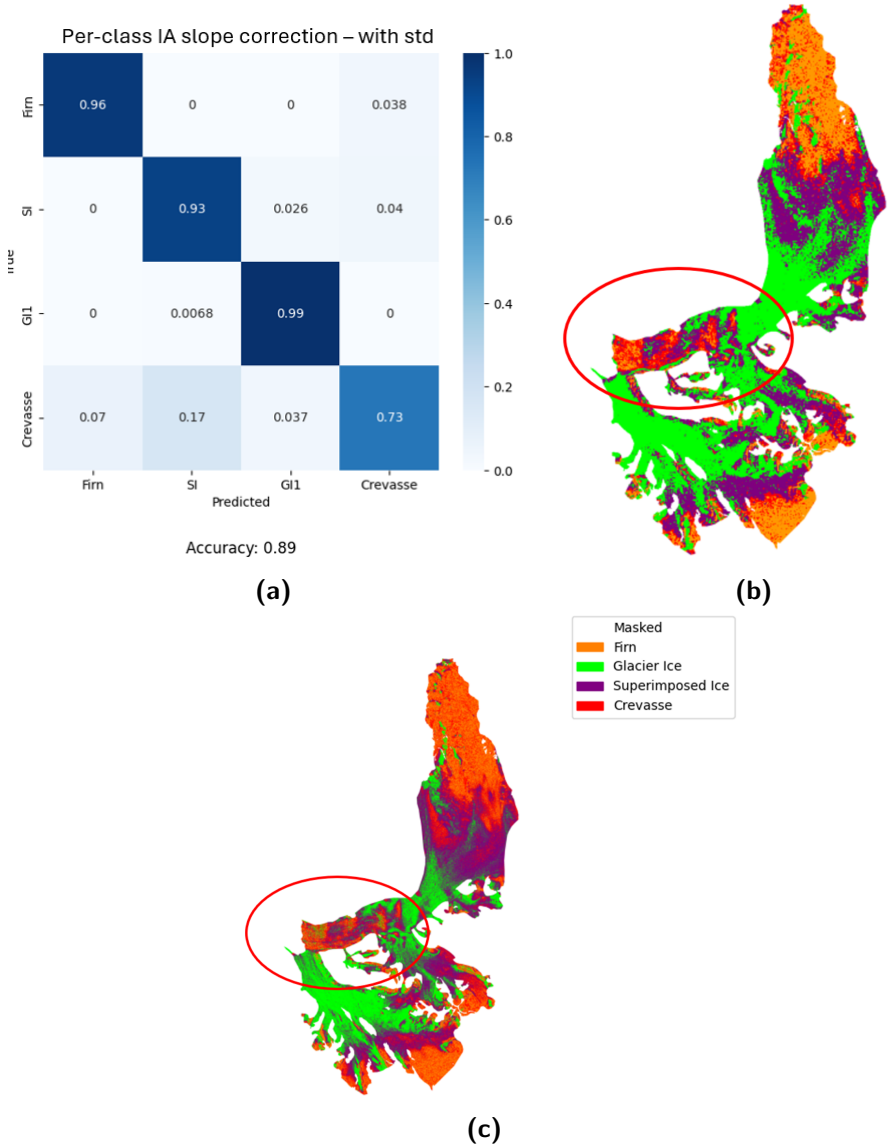
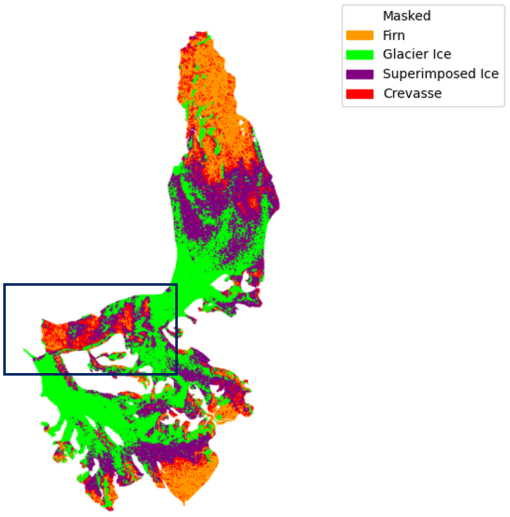
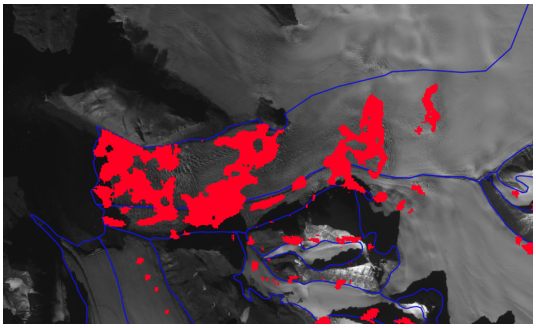


Figure 4-13: Mapping the labeled crevasses on optical imagery when using standard deviation as a pixel feature. Strong crevasses are correctly labeled but not those that have more irregular patterns, as seen by the red ellipses corresponding to misclassified regions.

Furthermore, if we compare the crevasse position on optical imagery to the labeled one, as seen in Figure 4-14, significant crevasses are mislabelled. After discussing with glaciologists, this specific crevasse region, which is not labeled as such, is a region with crushed crevasses and highly irregular patterns, possibly due to the glacier movement. This could explain why the standard deviation texture feature measures a similar random spread to firn.



(a) Classification result of 10th of January 2023



(b) Optical imagery of Sentinel-2 from 26th of August 2022. The red zones correspond to pixels labeled as crevasses.

Figure 4-14: Per-class IA slope correction classification result including standard deviation of HH and HV as features

There is also a notable difference in class labeling in mid-altitudes of Holtedahlfonna between glacier ice, firn, and superimposed ice.

Chapter 5

Discussion & future work

This section discusses our work with a summary of the interpretations of our results, from the data availability, to the classification of the glacier complexes. We then highlight potential improvements and extensions to this research.

Data availability For Sentinel-1 SAR data covering the HDF and KNG glacier complexes, there is a larger scene count for HH+HV polarizations, and thus we decided to only take such scenes for our classification. Scenes should be taken in winter months, from January to April, as it presents less variability in backscatter intensities due to minimal melting, and thus water inteference in SAR measurement. Whilst IW's higher resolution is preferred, there was an insufficent amount of data covering different relative orbits to allow us to calculate the per-class slopes of backscatter intensities vs local incidence angle. Therefore, for these glaciers, EW data was necessary for training purposes of the algorithm. This should not, however, be a problem for glaciers in other regions where there is a more comprehensive IW data coverage. Note also that for glaciers which are not in polar regions, the months which provide the most stable backscatter intensities could vary. Whilst EW and IW data are calibrated and should be directly comparable, it is left to confirm that EW slopes would correspond to IW ones, assuming the same amount of surface roughness resolution.

General implementation of GLIA on glaciers GLIA assumes a linear relationship between the backscatter intensities and the local incidence angles. This assumption holds true for these glaciers for the four classes identified, namely firn, superimposed ice, glacier ice, and crevasses. As a supervised algorithm, it requires the user to input the slopes per-class. For glacier zones corresponding to different surface roughness, these slopes are expected to vary between different glaciers, which makes direct transfer of the method trained classifier between glaciers challenging. This can be seen for HDF and KNG which have different structures and topography leading to different surface roughness. Aside from being misclassified with high roughness areas, the firn class seems to be transferable from HDF to KNG and possibly other glaciers as this zone sits in the highest elevations of the glacier (and thus more likely to have

less melt), but also has a smoother surface. However, the firn class showed the steepest slope out of all classes, which would mean that it has the most surface scattering. Müller (2011) has, however, pointed out that firn would have the highest backscatter intensity due to the most volumetric scattering. This disagreement could be understood by the firn region in HDF possibly containing some ice structures as it could be in a wet zone. If such ice structures exist close to the surface, then surface scattering would take place and this would explain such a steep slope. The GLIA algorithm allows for a clear separation between superimposed ice and firn as they have have different IA slopes. This makes it a useful algorithm to monitor the firn line over time. In general, GLIA requires prior analysis of the glacier's zones and IA slopes in order to be implemented. Misclassification of texture is observed for both common IA slope and per-class IA slope correction as the distribution of the backscatter intensity of crevasse regions significantly overlaps with firn and superimposed ice.

Classification of HDF and KNG These two glacier complexes show a relatively low variability in IA slope vs backscatter intensity, therefore as expected the improvement from common IA slope to per-class IA slope correction do not drastically change the zonation and general glaciology interpretation gained from classification. GLIA improves the tightness of the distributions and the zones are more defined. The initial classes chosen with two different glacier ice texture and without crevassing were inadequate as the glacier textures were resolved differently from EW to IW, and whilst crevasses are technically glacier ice, they can be considered an independent glacier zone due to the high surface roughness. This further supports that GLIA classification should consider the zones of glaciers individually as topography and generally glacier dynamics renders a firn or glacier ice of one zone to have different physical properties and surface roughness compared to another glacier. Finally, SAR data with GLIA classification has been used to monitor the evolution of the firn line over space and time, from 2017 to 2023. Firn line requires many years of net negative mass balance to retreat and this is observed on HDF and KNG. 2017 and 2020 data show similar firn line positions for KNG and HDF. 2023 show significant retreats and the isolation and/or loss of further firn zones in lower altitudes. As these are winter scenes, they represent the mass balance of the glaciers from the end of the previous summer. This agrees with meteorology data that summer 2020 and 2022 were particularly hot all over Svalbard (NMI and NRK, 2020, NMI, 2020, Copernicus, 2022), and shows quantitatively the effect of a recent rise of temperatures on these glaciers.

Texture feature

The glacier tongue of HDF showed the most misclassification due to crevasses and, thus, strong large-scale (meters to tens of meters) surface roughness. As both backscatter intensities and local incidence angles are widely variable in this region, it was necessary to include some texture feature to aid classification of this region. The standard deviation of the backscatter intensities differ largely between crevasses and superimposed ice which helps the differentiation, however integrating this new feature does not provide as much an improvement to separate firn and crevasses as they provide similar variations in roughness in relation to the resolution of the backscatter intensity data.
We further investigate here a texture feature that would include patterns by analyzing the spatial Fourier transform and similarity with a given directed wavelet. If SAR data over the

crevasse region has spatial patterns, then we should observe dominant spatial frequencies in the transform. On the other hand, for a given period of pattern that we expect crevassing to have, we should expect a convolution with the directed wavelet to yield higher magnitudes (and thus similarities) than a more randomly distributed backscatter like what is expected for the firn region.
Due to a lack of time this was not implemented in the GLIA algorithm, however it provides insights for future research.

Spatial Fourier transform Initially, two larger crevasse and firn regions from Holtedahlfonna of 420*240 pixels were delineated in the HH band on IW data of 16/01/2023 and on EW data from 10/01/2023. Figure 5-1 shows clearly in IW mode the crevasses forming inclined linear objects. We cannot observe in the EW mode, however the darker region present in the crevasse image for both EW and IW corresponds to a crevasse region on Holtedalhfonna where the crevasses were crushed due to glacier movement, and therefore they are more randomly spread then compared to the ones at the bottom left of the image, which corresponds to the sea front (personal communication, NPI).
It is clear from the EW data that linear crevasses are not observable. Within the large regions, smaller patches were segmented of a size of 100*100 pixels with a step size of 50. All 21 HH band patches can be observed in Appendix B Figure B-2 for crevasses. The 2D spatial frequencies were calculated for all patches, shifted and the central frequency removed (B-3,B-4).
We identified patches 7,8,14 and 15 as having strong crevasses in the HH band from the crevasse region selected as seen in Figure 5-3. In the 2D Fourier image, these crevasse patches present some elliptical shape that cannot be observed for firn data or weaker crevassing. The 2D Fourier transform was averaged over Y and X separately for each patch, and their difference calculated. We can observe from Figure 5-4a that these strong crevassing patches have frequencies with high magnitudes around 0.015 cycles/meter when the FT is averaged over the Y axis. This could be explained by crevasses in this case being slightly inclined in the y-axis. A potential feature could then be the mean magnitude of low frequencies (say within 0 to 0.015 cycles/meter). To get results independent of the crevasse direction, we further looked at the radial average (B-7 for whole crevasse patches and B-8 for firn ones). As seen in Figure 5-4b, dominant frequencies are not as easily identifiable as when averaging over the Y and X axis. However, they still present a larger magnitude for crevasse patches than for firn ones for low frequencies. This aligns with the visual observation of their FT that crevasses have a specific direction (elliptical shape). Figure 5-2 presents that indeed the mean magnitude of frequencies for crevasses is on average higher than that of firn for the majority of patches, even those not visually showing strong crevassing. This shows the potential for a texture feature based on a magnitude threshold of radial frequencies of spatial Fourier transforms.

Wavelet analysis As including a texture feature based on spatial fourier transforms to all pixels is computationally expensive, we also looked into wavelet analysis. Through the previous spatial FT analysis of the crevasse region of interest, we can identify a dominant frequency of around 0.015 Hz which corresponds to a period of around 70 m to train or input. This period result aligns with the crevasse peak distance of around 80 m we can observe on the backscatter intensities and measure through QGIS. From this, we can design a wavelet filter from a directed sinusoid and a gaussian envelope as plotted in Figure 5-5. For different

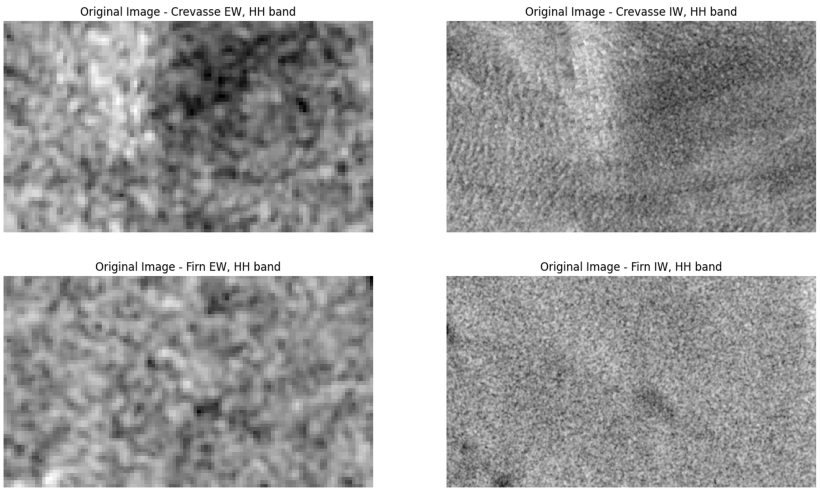


Figure 5-1: Original HH band of the 420*240 pixels regions for firn (bottom row) and crevasse (top row) taken from HDF for both EW (left column) and IW (right column).

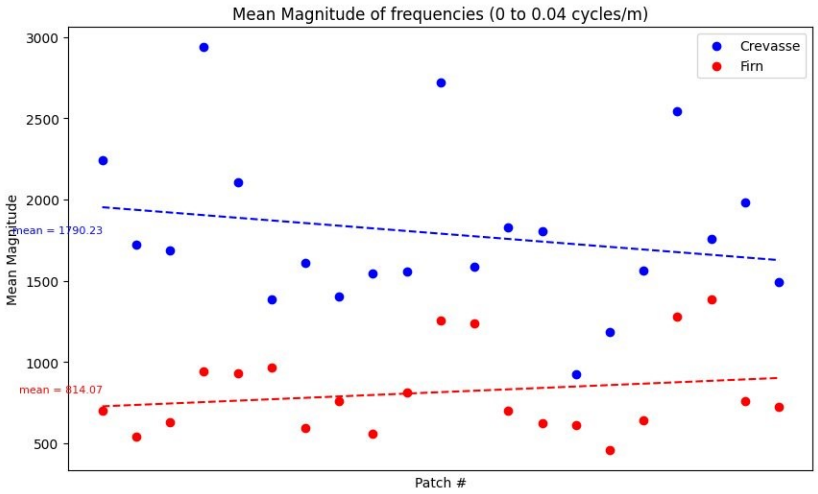


Figure 5-2: Plotting mean magnitude from 0 to 0.04 cycles/meter for crevasse and firn.

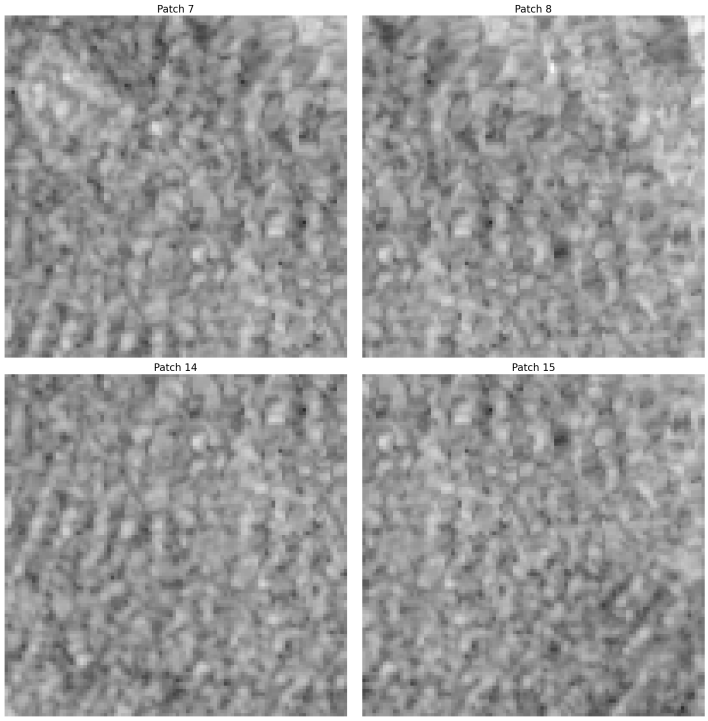
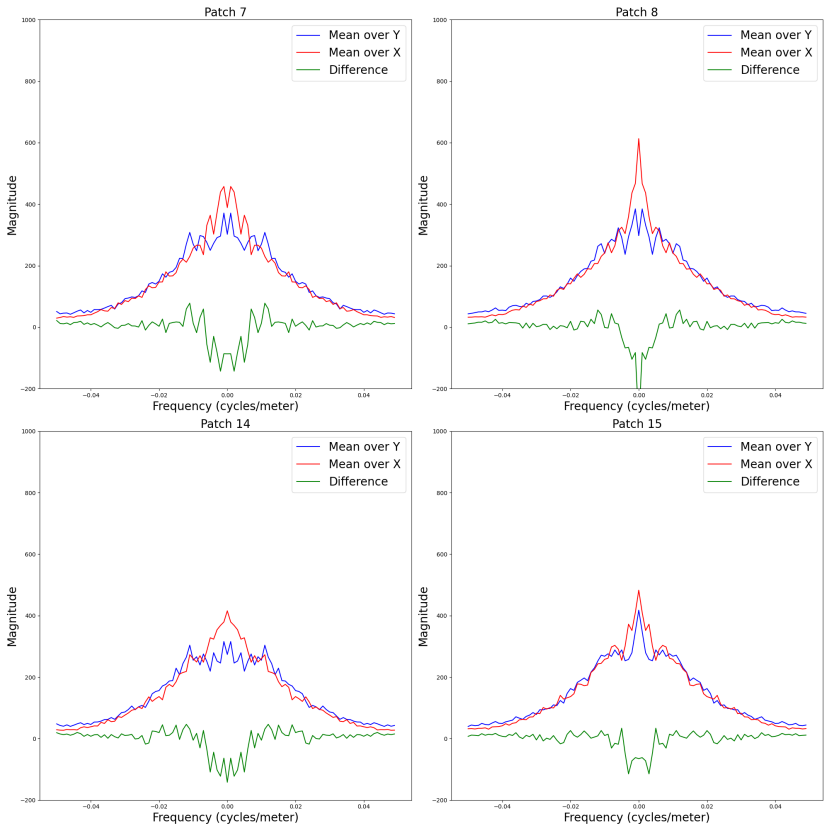


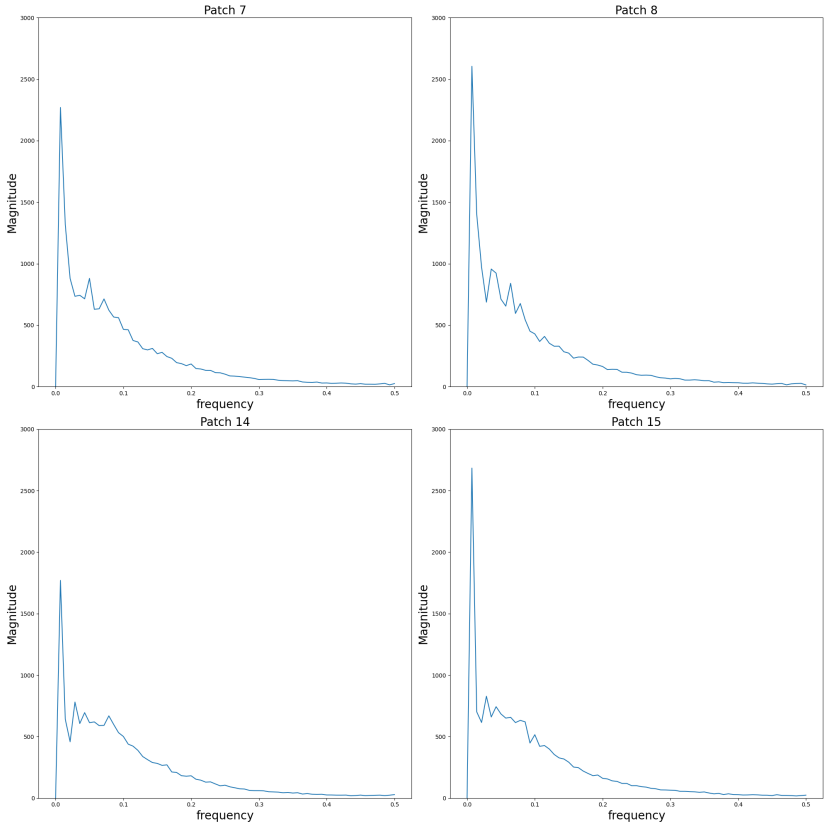
Figure 5-3: Four patches of 100*100 pixels from the IW crevasse region that visually show strong crevassing.

angles, we can convolve the wavelet filter with the regions of crevasses and firn to compare the highest absolute values, i.e the areas which align the most with a directed sinusoid and hence corresponding to a spatial pattern that would indicate crevassing. Figure 5-6 and 5-7 show examples of the convolution operation of the wavelet filter for different angles from 0° to 180° for both crevasses and firn areas on IW data. The size of the filter was taken to be twice the size of the period. We can observe that the highest absolute value of the image is on average much higher than for firn for given angles, which would indicate another type of feature possible to include in our algorithm to be able to differentiate between firn and crevasse classes. Based on this, for future research to look into a threshold classification of the maximum absolute value result from a wavelet filter.

GLCM Another way to include a texture feature could be by using the gray level co-occurrence matrix (GLCM) proposed by Haralick et al. (1973). This model calculates a range of correlated features from reference pixels and their neighbors - a clear explanation can be found in Hall-Beyer (2000). It has been widely applied in remote sensing and for glacier classification (Zhang et al., 2011, Wu et al., 2011, Sharma et al., 2024). It has further been implemented by Lohse et al. (2021) in sea-ice classification using the GLIA algorithm and this research would provide a starting point for glacier classification with per-class IA slope correction. Optimal features and combination of features for GLCM will have to be studied further to see which ones are the most relevant to improve crevasse detection.



(a) 2D Spatial Fourier transform averaged over the X (red) and Y axis (blue) and difference calculated (green).



(b) Radial frequency of the spatial Fourier Transform.

Figure 5-4: Spatial Fourier transform analysis for the four crevasse patches identified with strong crevassing.

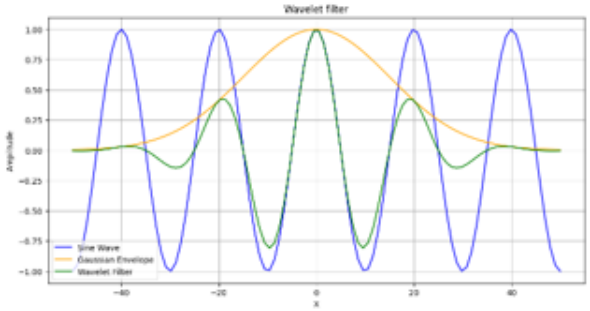


Figure 5-5: Wavelet filter plot made from a sinusoid and gaussian envelope.

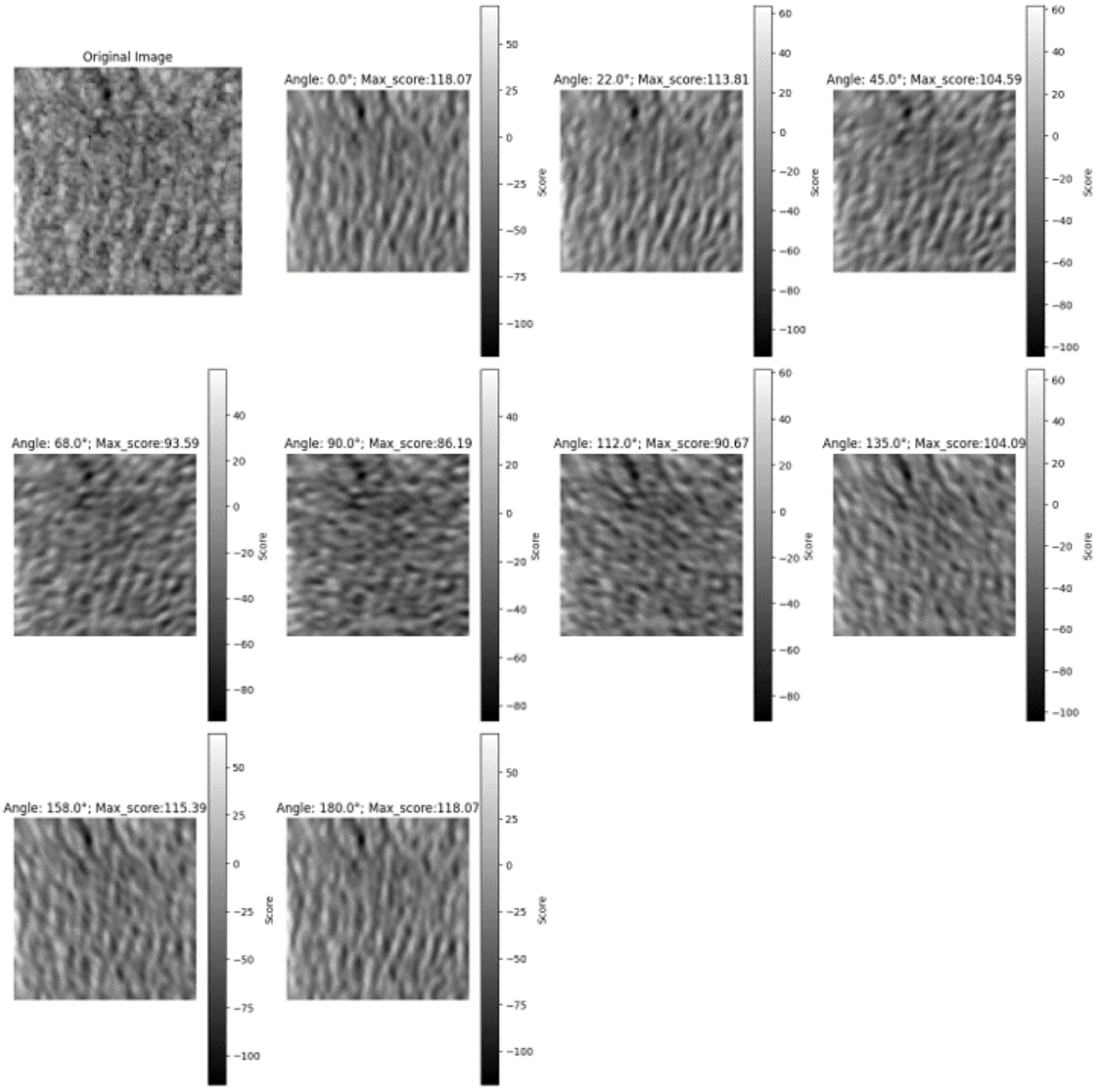


Figure 5-6: Crevasse wavelet analysis results. The crevasse region was convolved with a wavelet filter directed with various angles. The absolute maximum value of the whole image result was found and defines the max_score of the angle of the wavelet.

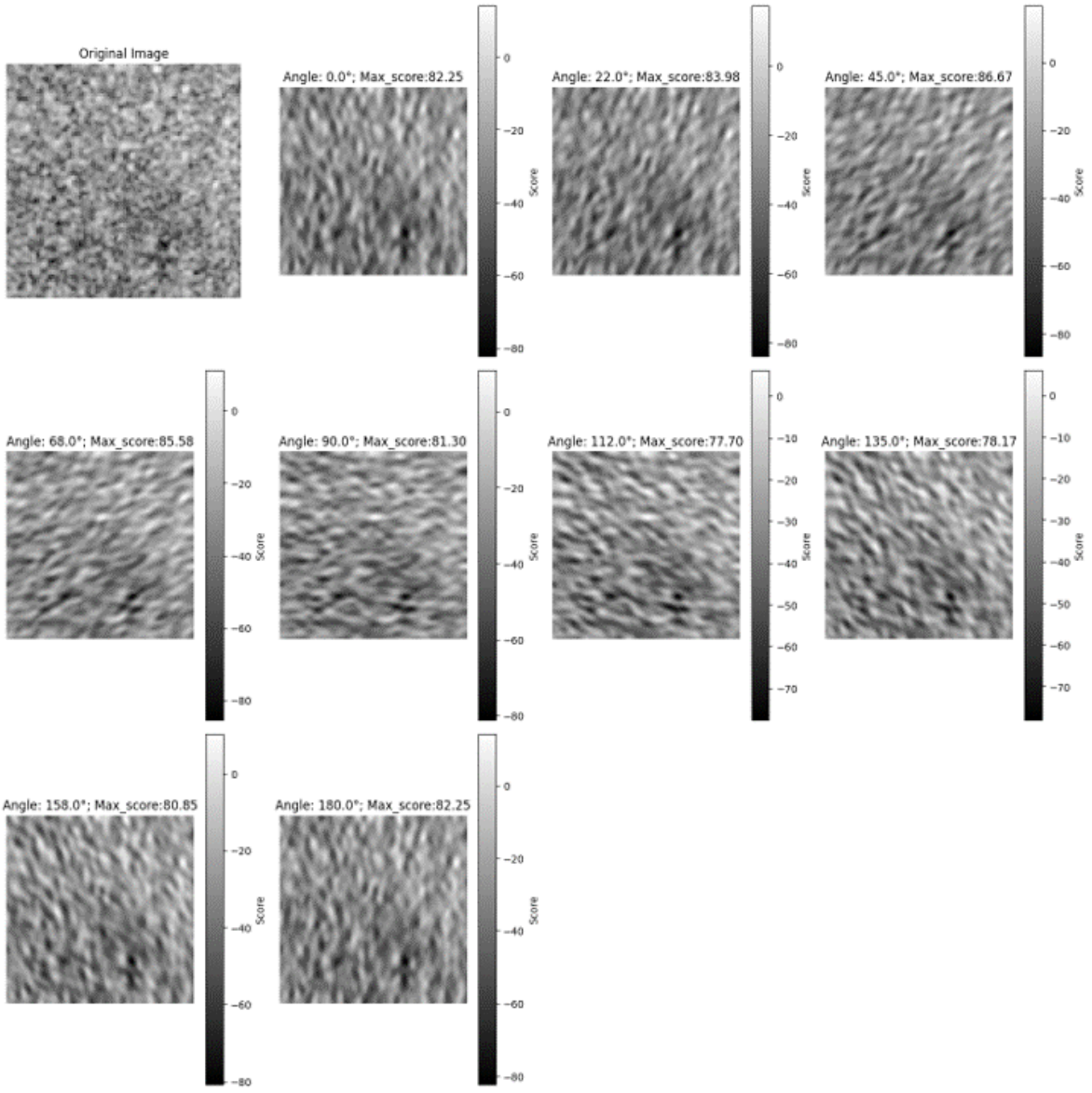


Figure 5-7: Firn wavelet analysis results once convolved with a directed wavelet of various angles.

Chapter 6

Conclusion

In this work, we applied a novel classification approach for SAR imagery to accurately map glacier surfaces, essential to monitor glaciers' mass balance in a changing climate. We presented the first case study that used per-class incidence angle slope correction over land areas, specifically glaciers, and directly incorporated SAR backscatter variation with local IA into the classification process. Applying a per-class radiometric correction enabled SAR products to be more precise and reliable regarding their backscatter intensities. While the algorithm was applied to sea ice previously, it had not been applied in a context of more topographical variations. Our case studies presented here corresponded to two glaciers Northwest of Svalbard: Holtedahlfonna and Kongsvegen glaciers. These two glacier complexes had been extensively studied during fieldwork led by the Norwegian Polar Institute and thus had a reliable ground truth for our satellite results.

The SAR images were taken from the Copernicus database and pre-processed (without radiometric correction) using ESA's SNAP toolbox.

Our first research objective was to understand the amount of SAR data available for classifying our chosen glaciers and to apply this algorithm. We labeled scenes according to their acquisition time, polarization, and relative orbit for this. Winter scenes from January to April presented the most stable backscatter intensities due to the least melting-refreeze cycle. While interferometric-wide (IW) SAR data was preferred for its high resolution, those glaciers and the Svalbard region, in general, had a higher count of products and a wider range of relative orbits for the extra-wide (EW), a lower resolution acquisition mode, possibly because most science projects in this polar region covered maritime regions. As the per-class IA slope correction required finding the slope of backscatter intensities vs local IA, we needed data covering a large range of IA, and this was only possible for EW data. The amount of SAR data available through Sentinel-1 satellites limited the study to being done fully in high resolution and consistent in the data used.

The second research objective involved testing the novel algorithm on those glaciers and understanding the benefits and limitations. We tested three Bayesian classifiers: one without radiometric correction, one using a common IA slope correction, and the final using the per-class IA slope correction. Comparing the latter with the uncorrected case and a common IA

slope correction, which was the most widely used method in remote sensing and glacier classification, it increased the classification performance over our training regions by around 30% and 10%, respectively. Glacier ice had much lower backscatter intensities than superimposed ice and firn, and firn had a more varying IA slope than the two other classes. This made the three classes more separable when integrating per-class IA slope correction. When applying to our data, the firn line could be distinctly detected for both glacier complexes and mapped to observe its evolution from 2017 to 2023. A retreat of the firn line was observed, which correlated with the meteorological data of summers in Svalbard which were particularly warm in this timeframe.

Our third research question arose from the observation that lower altitudes of HDF, which presented heavy crevassing, were largely misclassified. The high backscatter intensity distribution of crevassing also heavily overlapped with the one of firn and superimposed ice. It was likely that the calculated IA slope of crevassed areas was misleading due to the limited data available; however, its approximately null slope showed the extent to which the volume scattering present in heavy crevassing was not affected by an IA decay. To integrate a texture feature in the algorithm, we initially looked into a simplistic standard deviation metric that marginally improved the zonation of crevasses. We then explored alternatives with spatial Fourier transforms and wavelet analysis, which indicated other quantitative differences between crevasses and firn. Another alternative could be the grey co-occurrence matrix (GLCM); however, a deeper literature review for parameter optimization would be necessary. These methods showed potential in detecting crevassing within the limits of the SAR and DEM resolution, and we recommend future research to integrate a texture feature in GLIA to increase its accuracy for glacier classification.

Bibliography

Adam, S., Pietroniro, A., & Brugman, M. M. (1997). Glacier snow line mapping using ers-1 sar imagery. *Remote Sensing of Environment*, *61*(1), 46–54.

Akbari, V., Doulgeris, A. P., & Eltoft, T. (2013). Monitoring glacier changes using multitemporal multipolarization sar images. *IEEE transactions on geoscience and remote sensing*, *52*(6), 3729–3741.

Aulard-Macler, M. (2011). Sentinel-1 product definition. *Document Reference MPC-0240*.

Baumhoer, C. A., Dietz, A. J., Kneisel, C., & Kuenzer, C. (2019). Automated extraction of antarctic glacier and ice shelf fronts from sentinel-1 imagery using deep learning. *Remote Sensing*, *11*(21). https://doi.org/10.3390/rs11212529

Benson, C. S. (1960). *Stratigraphic studies in the snow and firn of the greenland ice sheet* [Doctoral dissertation, California Institute of Technology].

Brandt, O., Kohler, J., & Lüthje, M. (2008). Spatial mapping of multi-year superimposed ice on the glacier kongsvegen, svalbard. *Journal of Glaciology*, *54*(184), 73–80.

Braun, M., Rau, F., Saurer, H., & Gobmann, H. (2000). Development of radar glacier zones on the king george island ice cap, antarctica, during austral summer 1996/97 as observed in ers-2 sar data. *Annals of Glaciology*, *31*, 357–363.

Callegari, M., Carturan, L., Marin, C., Notarnicola, C., Rastner, P., Seppi, R., & Zucca, F. (2016). A pol-sar analysis for alpine glacier classification and snowline altitude retrieval. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(7), 3106–3121. https://doi.org/10.1109/JSTARS.2016.2587819

Copernicus. (2022). *Arctic temperature - european state of the climate 2022* [https://climate.copernicus.eu/esotc/2022/arctic-temperature, Accessed: 2024-07-08].

Copernicus. (2024). Sentiwiki - s1 products [Accessed: 2024-07-09]. https://sentiwiki.copernicus.eu/web/s1-products

Cristea, A., Van Houtte, J., & Doulgeris, A. P. (2020). Integrating incidence angle dependencies into the clustering-based segmentation of sar images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 2925–2939.

Cuffey, K. M., & Paterson, W. S. B. (2010). *The physics of glaciers*. Academic Press.

Curlander, J. C., & McDonough, R. N. (1991). *Synthetic aperture radar* (Vol. 11). Wiley, New York.

De Zan, F., & Guarnieri, A. M. (2006). Topsar: Terrain observation by progressive scans. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(9), 2352–2360.

Doulgeris, A. P., Langley, K., & Eltoft, T. (2008). Analysis and classification of high arctic glaciers with asar data. *IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium*, *4*, IV–181.

Dunse, T., Schuler, T. V., Hagen, J. O., Eiken, T., Brandt, O., & Høgda, K. A. (2009). Recent fluctuations in the extent of the firn area of austfonna, svalbard, inferred from gpr. *Annals of Glaciology*, *50*(50), 155–162.

Forster, R. R., Isacks, B. L., & Das, S. B. (1996). Shuttle imaging radar (sir-c/x-sar) reveals near-surface properties of the south patagonian icefield. *Journal of Geophysical Research: Planets*, *101*(E10), 23169–23180.

Fung, A. K. (1994). Microwave scattering and emission models and their applications. *(No Title)*.

Gill, J. P., Yackel, J. J., Geldsetzer, T., & Fuller, M. C. (2015). Sensitivity of c-band synthetic aperture radar polarimetric parameters to snow thickness over landfast smooth first-year sea ice. *Remote Sensing of Environment*, *166*, 34–49.

Hagen, J., Melvold, K., Eiken, T., Isaksson, E., & Lefauconnier, B. (1999). Mass balance methods on kongsvegen, svalbard. *Geografiska Annaler: Series A, Physical Geography*, *81*(4), 593–601.

Hagen, J. O., Liestøl, O., Roland, E., & Jørgensen, T. (1993). *Glacier atlas of svalbard and jan mayen*.

Hall, A. (2004). The role of surface albedo feedback in climate. *Journal of climate*, *17*(7), 1550–1568.

Hall-Beyer, M. (2000). Glcm texture: A tutorial. *National Council on Geographic Information and Analysis Remote Sensing Core Curriculum*, *3*(1), 75.

Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6), 610–621.

Hawley, R. L., Brandt, O., Morris, E. M., Kohler, J., Shepherd, A. P., & Wingham, D. J. (2008). Techniques for measuring high-resolution firn density profiles: Case study from kongsvegen, svalbard. *Journal of Glaciology*, *54*(186), 463–468.

Huang, L., Li, Z., & Tian, B.-s. (2016). Local incidence angle referenced classification on polarimetric synthetic aperture radar images in mountain glacier areas. *Journal of Applied Remote Sensing*, *10*(2), 025015–025015.

Kääb, A., Lefauconnier, B., & Melvold, K. (2005). Flow field of kronebreen, svalbard, using repeated landsat 7 and aster data. *Annals of Glaciology*, *42*, 7–13.

Karvonen, J., Simila, M., & Makynen, M. (2002). An iterative incidence angle normalization algorithm for sea ice sar images. *IEEE International Geoscience and Remote Sensing Symposium*, *3*, 1524–1527.

Koerner, R. (1970). Some observations on superimposition of ice on the devon island ice cap, nwt canada. *Geografiska Annaler: Series A, Physical Geography*, *52*(1), 57–67.

König, M., Wadham, J., Winther, J.-G., Kohler, J., & Nuttall, A.-M. (2002). Detection of superimposed ice on the glaciers kongsvegen and midre lovénbreen, svalbard, using sar satellite imagery. *Annals of Glaciology*, *34*, 335–342.

König, M., Winther, J.-G., Kohler, J., & König, F. (2004). Two methods for firn-area and mass-balance monitoring of svalbard glaciers with sar satellite images. *Journal of Glaciology*, *50*(168), 116–128.

Lang, W., Zhang, P., Wu, J., Shen, Y., & Yang, X. (2016). Incidence angle correction of sar sea ice data based on locally linear mapping. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(6), 3188–3199.

Langley, K., Hamran, S.-E., Hogda, K. A., Storvold, R., Brandt, O., Kohler, J., & Hagen, J. O. (2008). From glacier facies to sar backscatter zones via gpr. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(9), 2506–2516.

Langley, K., Lacroix, P., Hamran, S.-E., & Brandt, O. (2009). Sources of backscatter at 5.3 ghz from a superimposed ice and firn area revealed by multi-frequency gpr and cores. *Journal of Glaciology*, *55*(190), 373–383.

Liestøl, O. (1988a). The glaciers in the kongsfjorden area, spitsbergen. *Norsk Geografisk Tidsskrift-Norwegian Journal of Geography*, *42*(4), 231–238.

Liestøl, O. (1988b). The glaciers in the kongsfjorden area, spitsbergen. *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography*, *42*(4), 231–238. https://doi.org/10.1080/00291958808552205

Lohse, J., Doulgeris, A. P., & Dierking, W. (2020). Mapping sea-ice types from sentinel-1 considering the surface-type dependent effect of incidence angle. *Annals of Glaciology*, *61*(83), 260–270.

Lohse, J., Doulgeris, A. P., & Dierking, W. (2021). Incident angle dependence of sentinel-1 texture features for sea ice classification. *Remote Sensing*, *13*(4), 552.

Mahmud, M. S., Geldsetzer, T., Howell, S. E., Yackel, J. J., Nandan, V., & Scharien, R. K. (2018). Incidence angle dependence of hh-polarized c-and l-band wintertime backscatter over arctic sea ice. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(11), 6686–6698.

Makynen, M., Manninen, A. T., Simila, M., Karvonen, J. A., & Hallikainen, M. T. (2002). Incidence angle dependence of the statistical properties of c-band hh-polarization backscattering signatures of the baltic sea ice. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(12), 2593–2605.

Mäkynen, M., & Karvonen, J. (2017). Incidence angle dependence of first-year sea ice backscattering coefficient in sentinel-1 sar imagery over the kara sea. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(11), 6170–6181.

Marshall, G., Rees, W., & Dowdeswell, J. (1995). The discrimination of glacier facies using multi-temporal ers-1 sar data. *Sensors and Environmental Applications of Remote Sensing, Balkema, Rotterdam, The Netherlands*, 263–269.

Maussion, F., Hock, R., Paul, F., Raup, B., Rastner, P., Zemp, M., Andreassen, L., Barr, I., Bolch, T., Kochtitzky, W., McNabb, R., & Tielidze, L. (2023). The Randolph Glacier Inventory version 7.0 User guide v1.0. https://doi.org/10.5281/zenodo.8362857

Melvold, K., & Hagen, J. O. (1998). Evolution of a surge-type glacier in its quiescent phase: Kongsvegen, spitsbergen, 1964–95. *Journal of Glaciology*, *44*(147), 394–404.

Melvold, K., & Ove Hagen, J. (1998). Evolution of a surge-type glacier in its quiescent phase: Kongsvegen, spitsbergen, 1964–95. *Journal of Glaciology*, *44*(147), 394–404. https://doi.org/10.3189/S0022143000002720

Mladenova, I. E., Jackson, T. J., Bindlish, R., & Hensley, S. (2012). Incidence angle normalization of radar backscatter data. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(3), 1791–1804.

Müller, K. (2011). Microwave penetration in polar snow and ice: Implications for gpr and sar.

Nagler, T., & Rott, H. (2000). Retrieval of wet snow by means of multitemporal sar data. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(2), 754–765.

NMI. (2020). *Svalbard: The state of the cryosphere 2020* [https://cryo.met.no/en/svalbard-cryosphere-2020, Accessed: 2024-07-08].

NMI & NRK. (2020). *Ny-ålesund - statistics as a graph* [https://www.yr.no/en/statistics/graph/2837778/Norway/Svalbard/Svalbard/Ny-lesund?q=2020, Accessed: 2024-07-08].

Norwegian Polar Institute. (2017). S100 Topographic Raster Data for Svalbard. https://doi.org/10.21334/npolar.1990.44ca8c2a

Nuth, C., Schuler, T. V., Kohler, J., Altena, B., & Hagen, J. O. (2012). Estimating the long-term calving flux of kronebreen, svalbard, from geodetic elevation changes and mass-balance modeling. *Journal of Glaciology*, *58*(207), 119–133. https://doi.org/10.3189/2012JoG11J036

Partington, K. (1998). Discrimination of glacier facies using multi-temporal sar data. *Journal of Glaciology*, *44*(146), 42–53.

Paterson, W. S. B. (1994). *Physics of glaciers*. Butterworth-Heinemann.

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., & Laaksonen, A. (2022). The arctic has warmed nearly four times faster than the globe since 1979. *Communications earth & environment*, *3*(1), 168.

RGI. (2023). Randolph Glacier Inventory - A Dataset of Global Glacier Outlines, Version 7.0. https://doi.org/10.5067/f6jmovy5navz

Rizzoli, P., Martone, M., Gonzalez, C., Wecklich, C., Tridon, D. B., Bräutigam, B., Bachmann, M., Schulze, D., Fritz, T., Huber, M., et al. (2017). Generation and performance assessment of the global tandem-x digital elevation model. *ISPRS Journal of Photogrammetry and Remote Sensing*, *132*, 119–139.

Rott, H., & Davis, R. E. (1993). Multifrequency and polarimetric sar observations on alpine glaciers. *Annals of Glaciology*, *17*, 98–104. https://doi.org/10.3189/S0260305500012672

Rott, H., & Mätzler, C. (1987). Possibilities and limits of synthetic aperture radar for snow and glacier surveying. *Annals of Glaciology*, *9*, 195–199.

Rott, H., & Nagler, T. (1994). Capabilities of ers-1 sar for snow and glacier monitoring in alpine areas. *European Space Agency-Publications-ESA SP*, *361*, 965–965.

SentiWiki. (2024). *S1 products* [Accessed: 2024-07-09]. https://sentiwiki.copernicus.eu/web/s1-products

Sharma, A., Gupta, M., & Sharma, N. (2024). Glacier facies characterisation in transboundary west sikkim himalaya from terrasar-x; glcm based classification approach. *Journal of Spatial Science*, *69*(1), 43–59.

Thakur, P. K., Dixit, A., Chouksey, A., Aggarwal, S., & Kumar, A. S. (2016). Ice sheet features identification, glacier velocity estimation, and glacier zones classification using high-resolution optical and sar data. *Land Surface and Cryosphere Remote Sensing III*, *9877*, 172–187.

Ulaby, F. (1980). Vegetation clutter model. *IEEE Transactions on Antennas and Propagation*, *28*(4), 538–545.

Ulaby, F. T., Moore, R. K., & Fung, A. K. (1982). *Microwave remote sensing active and passive-volume ii: Radar remote sensing and surface scattering and enission theory* (tech. rep.). Addison-Wesley Publishing CompanyAdvanced Book Program/World Science Division.

Van den Broeke, M., Smeets, P., Ettema, J., Van der Veen, C., Van de Wal, R., & Oerlemans, J. (2008). Partitioning of melt energy and meltwater fluxes in the ablation zone of the west greenland ice sheet. *The Cryosphere*, *2*(2), 179–189.

Van Pelt, W., & Kohler, J. (2015). Modelling the long-term mass balance and firn evolution of glaciers around kongsfjorden, svalbard. *Journal of Glaciology*, *61*(228), 731–744.

Vaughan, D. G., Comiso, J. C., Allison, I., Carrasco, J., Kaser, G., Kwok, R., Mote, P., Murray, T., Paul, F., Ren, J., et al. (2013). Observations: Cryosphere. *Climate change*, *2103*, 317–382.

Wagner, W., Lemoine, G., Borgeaud, M., & Rott, H. (1999). A study of vegetation cover effects on ers scatterometer data. *IEEE Transactions on Geoscience and Remote Sensing*, *37*(2), 938–948.

Wessel, B., Huber, M., Wohlfart, C., Marschalk, U., Kosmann, D., & Roth, A. (2018). Accuracy assessment of the global tandem-x digital elevation model with gps data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *139*, 171–182.

Woodward, J., Murray, T., & McCaig, A. (2002). Formation and reorientation of structure in the surge-type glacier kongsvegen, svalbard. *Journal of Quaternary Science: Published for the Quaternary Research Association*, *17*(3), 201–209.

Wu, H., Zhang, Y., Zhong, W., & Sun, G. (2011). A new method for extracting glacier area using sar interferometry. *2011 International Symposium on Image and Data Fusion*, 1–3. https://doi.org/10.1109/ISIDF.2011.6024207

Zakhvatkina, N., Korosov, A., Muckenhuber, S., Sandven, S., & Babiker, M. (2017). Operational algorithm for ice–water classification on dual-polarized radarsat-2 images. *The Cryosphere*, *11*(1), 33–46.

Zhang, Y., Zhong, W., Sun, G., et al. (2011). A new method for extracting glacier area using sar interferometry. *2011 International Symposium on Image and Data Fusion*, 1–3.

Appendix A

Additional plots - Data selection

A-1 Months HV

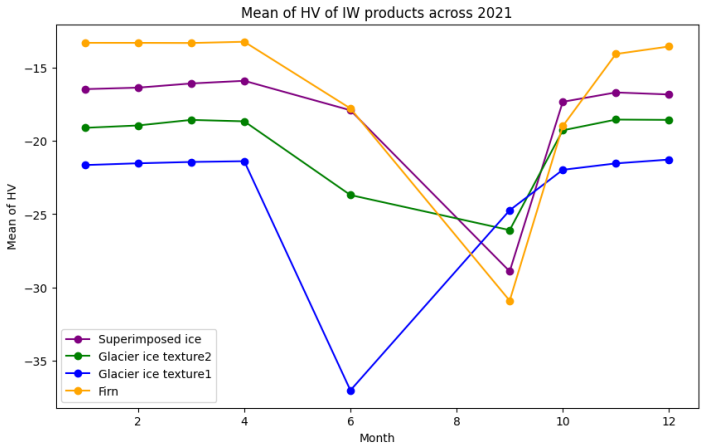
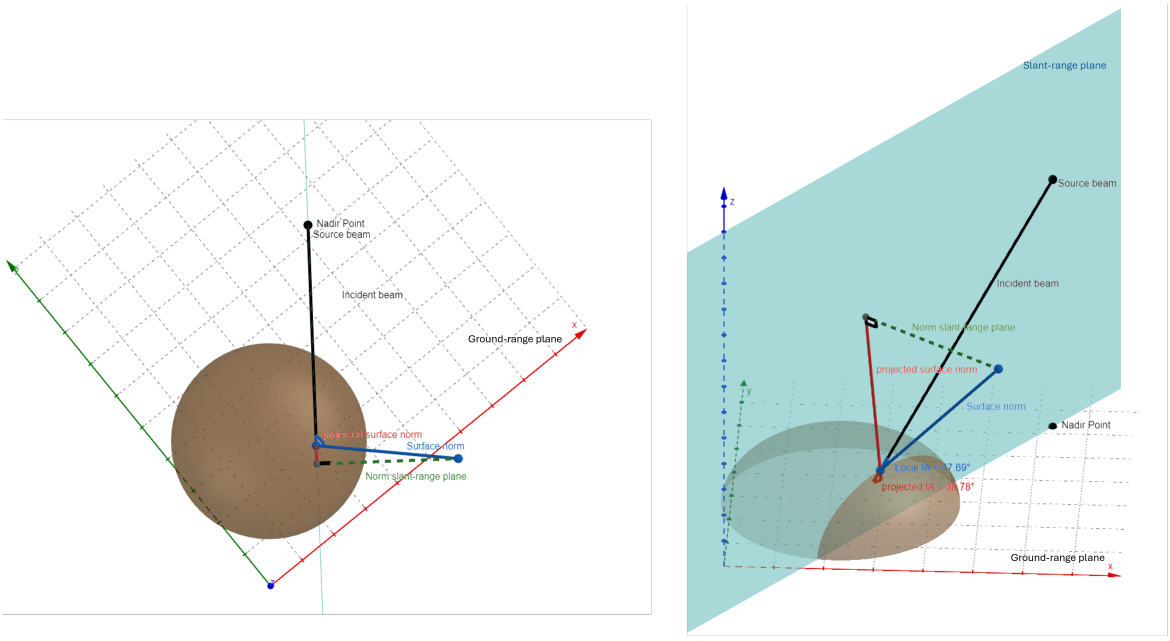


Figure A-1: HV backscatter intensities across 1 year. HV values are more consistent from January to April just as observed for HH values.

A-2 Local Incidence angles (projected or not)

The difference between the two is illustrated with Figure A-2. The projected local IAs are done so on the slant-range plane which means they would only include the range component of the angle. On the other hand, the full angle representation calculated by "non-projected local IA" capture both range and azimuthal components which will better capture the assumed symmetric decay of the diffuse scattering. The full angle behaviour is then the most relevant for our use so the IAs calculated and used throughout the rest of this work correspond to the non-projected values, simply called "local incidence angles" in SNAP.



(a) (b)

Figure A-2: Comparison between local IA and projected local IA in geometry. The blue plane is the slant-range plane and is defined by the beam source from the satellite and the nadir point. The grid corresponds to the ground range plane and the reflector is depicted here as a brown sphere. (a) corresponds to a view in which the projected local incidence angle is not visible. In (b), the norm to the slant-range (dotted green line), the projected surface norm (red), and the projected local incidence angle are visible.

A-3 Glacier outlines

We present here the comparison between the openly accessible glacier outlines found through the GLIMS dataset and the one internally communicated by NPI. As defined in the GLIMS user guide by Maussion et al. (2023), the region code 07 was downloaded as it corresponds to Svalbard and Jan Meyen, and the glacier complex boundaries and centerlines were mapped using QGIS (version 3.30.1).

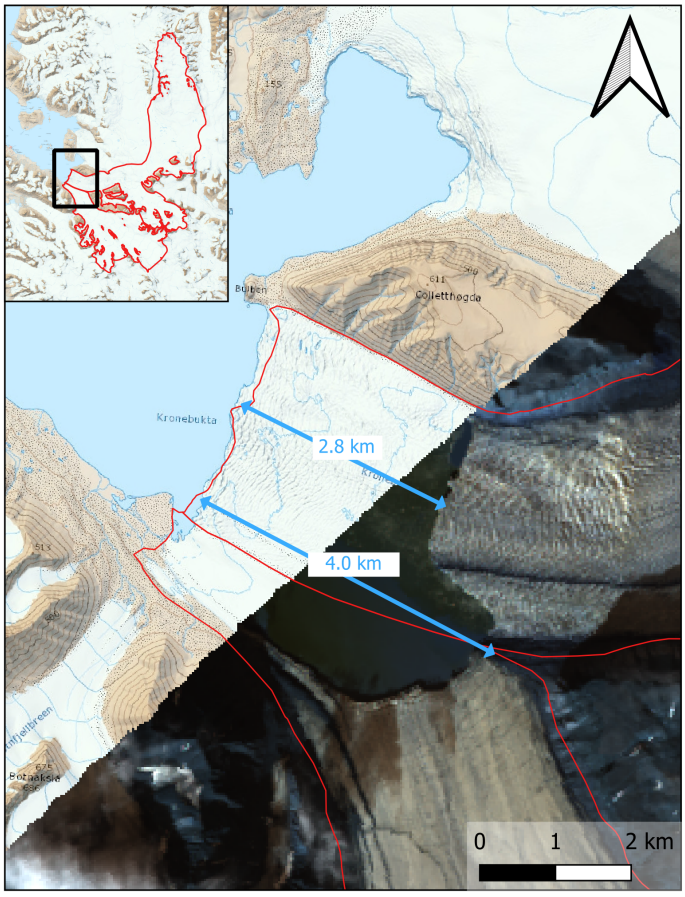


Figure A-3: Comparing the GLIMS outline (red) from the year 2000 to recent optical imagery from LANDSAT-8. There is a retreat of the glaciers where it reaches the sea to around 3km on average. For the rest of the two glacier complex, the outlines matches still the glacier of today. Base map from Norwegian Polar Institute, 2017.

fig:optical˙vs˙outline

As the GLIMS database estimated those outlines for the year 2000, they were compared with recent optical imagery to identify if the outlines are not too different from the present day. We superposed using QGIS the GLIMS boundaries with a 2023 summer (25th of August) scene of LANDSAT-8 and found that the outline matches the glacier in most areas except where the glacier complex reaches the sea. Here we observe a retreat of the glaciers of 2.8 km to 4 km. Towards the edges of Infantfonna and Fatumbreen, we also observe some melting

but reach a maximum of 1 km. The NPI is currently revising the boundaries without having a confirmed outline for the glaciers. However, they internally gave us this draft based on a 2020 Sentinel 2 mosaic, which corresponds better to the LANDSAT imagery of 2023. These are the glacier outlines we use for our study.

A-4 SAR results example scene

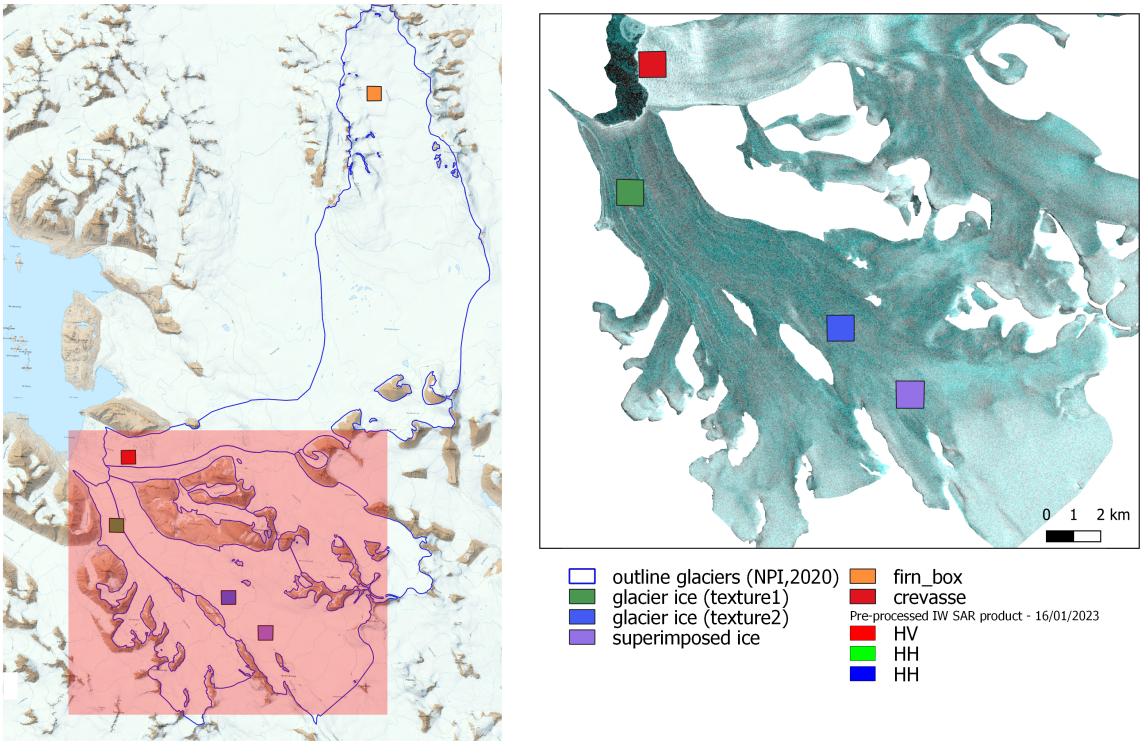


Figure A-4: SAR results example scene with training regions position.

Appendix B

Spatial Fourier transforms

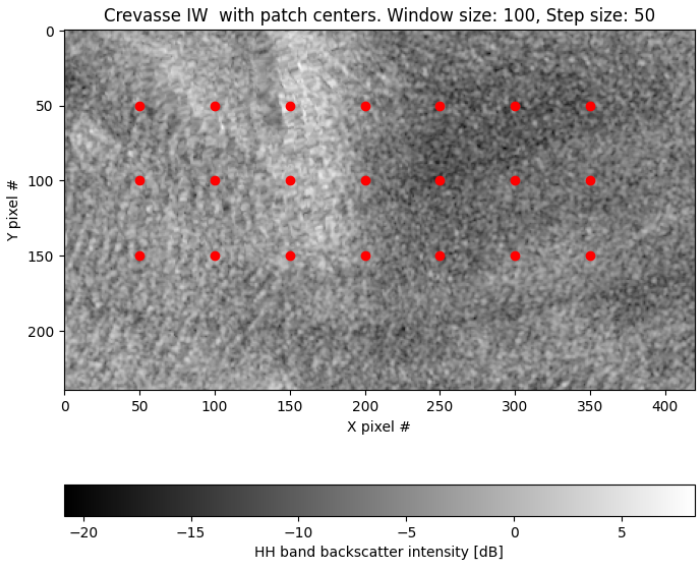


Figure B-1: Crevasse region that was segmented into smaller patches. The red dots correspond to the centers of the 21 patches.

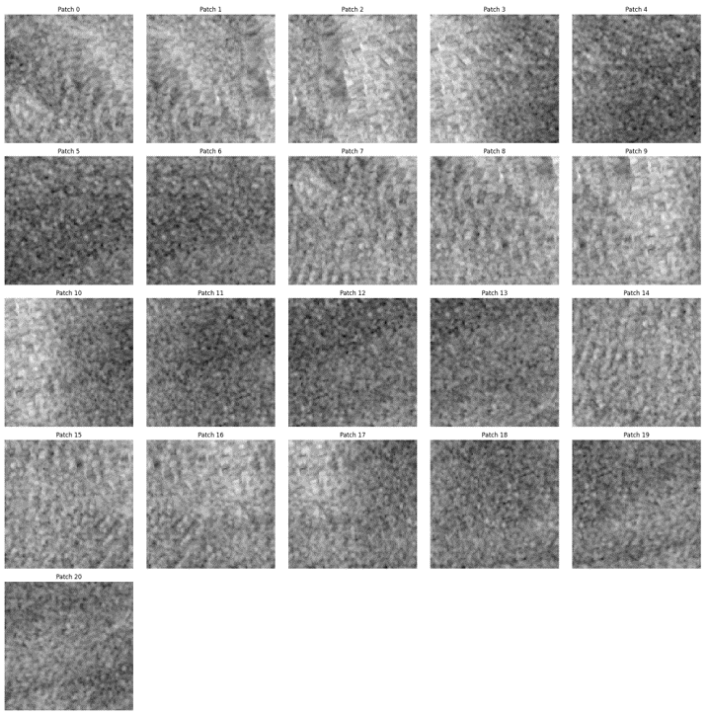


Figure B-2: Subset crevasse patches from the larger region found in Figure B-1.

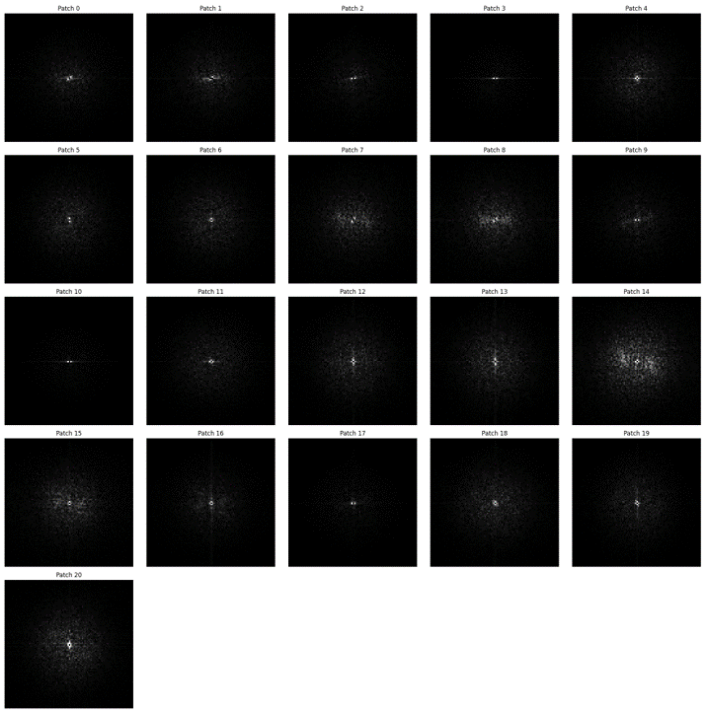


Figure B-3: Spatial Fourier transform images of the crevasse patches.

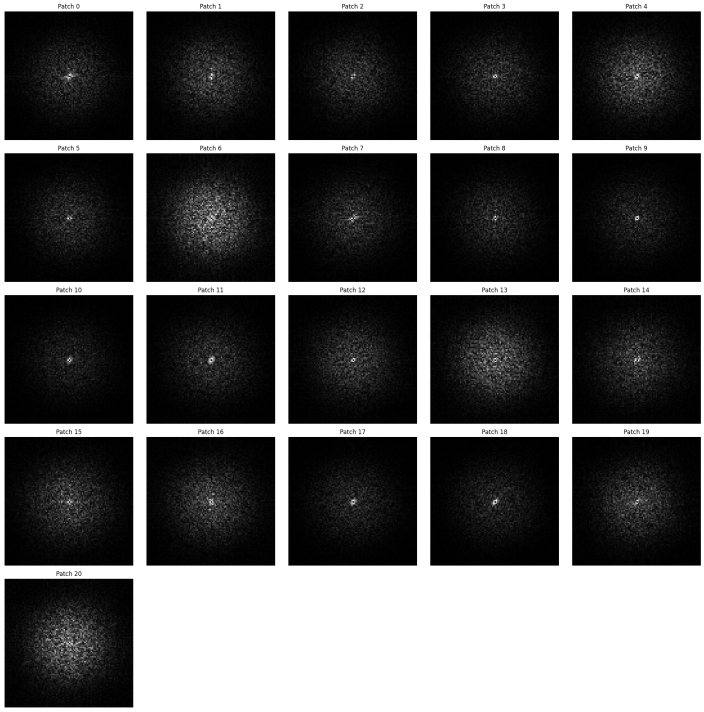


Figure B-4: Spatial Fourier transform images of the firn patches.

Figure B-5: Spatial Fourier transform analysis of the firn patches, applying averages over the X and Y axes.

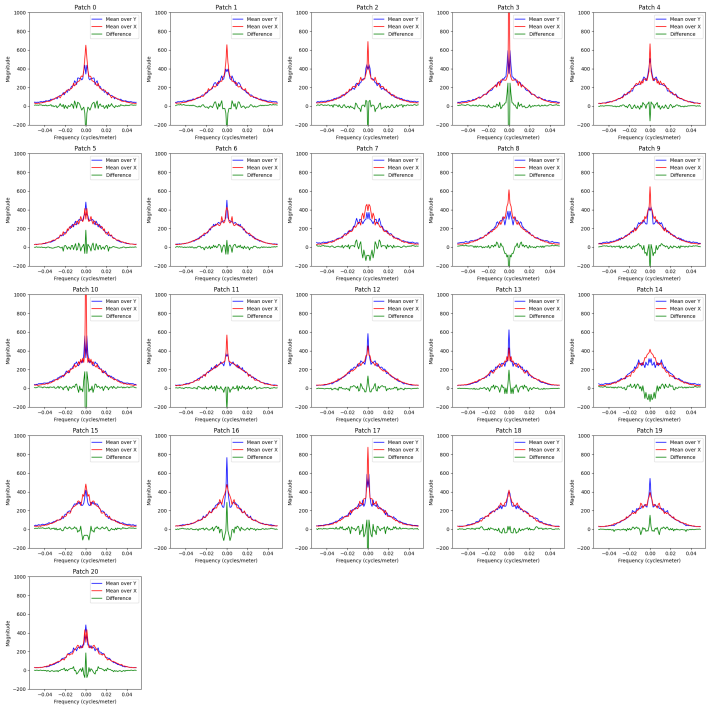


Figure B-6: Spatial Fourier transform analysis of the crevasse patches, applying averages over the X and Y axes.

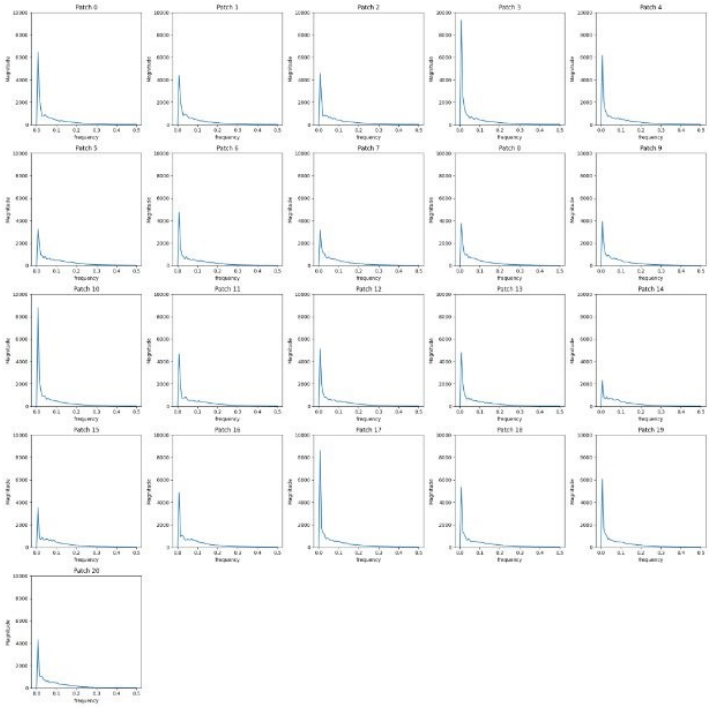


Figure B-7: Radial frequency of spatial Fourier transforms of crevasse patches

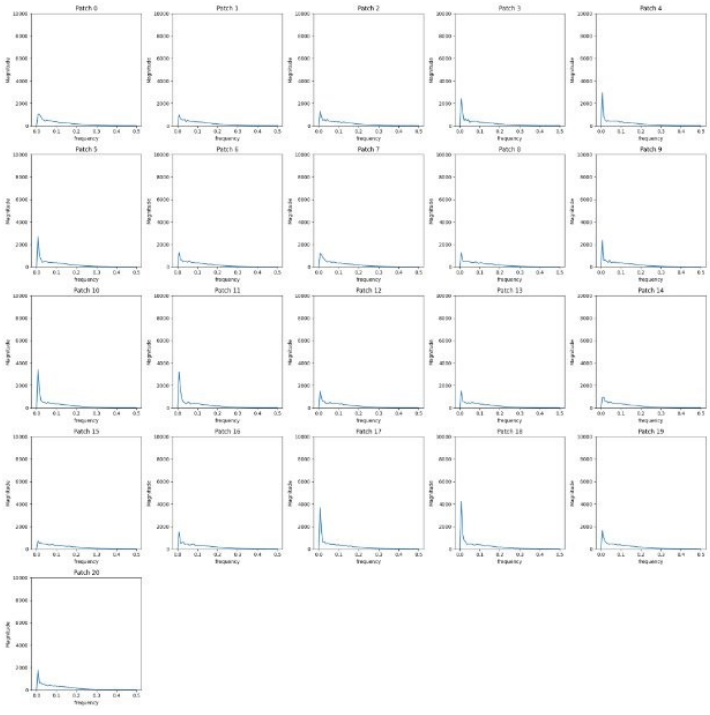


Figure B-8: Radial frequency of spatial Fourier transforms of firn patches

Appendix C

Data management

This chapter explains where the data was acquired and how it was used. All software and data used are free and of open access.
The availability of products was searched through the API of Copernicus. For the products of interest, they were downloaded using the Creodias platform onto the UiT Earth Observation's group server. Pre-processing of those products were done using the SNAP toolbox.
The main classification code can be found on Github here. The folder "classifiers" contains the three different Bayesian classifiers: one without correction, one with a common IA slope correction and the one with per-class correction. The folder "data" contains the necessary functions to load, pre-process and calculate statistics on the products. The folder "results" saves the classifications on the training data and are named according to the classes chosen. The file "main.py" is finally the main python file running the whole classification and requires for input the directory containing the training regions, the product's path we want to classify and the classes chosen.