Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Engineering
Department of Civil and Environmental Engineering

Galina Valeska Jonat

Synthetic Aperture Radar backscatter for automated delineation of the Disko Island glaciers

Master's thesis in Engineering and ICT

Supervisor: Prof. Hongchao Fan, Prof. Irina Rogozhina, Dr. Signe Hillerup Larsen

August 2020

NTNU
Norwegian University of Science and Technology

Galina Valeska Jonat

Synthetic Aperture Radar backscatter for automated delineation of the Disko Island glaciers

Master's thesis in Engineering and ICT
Supervisor: Prof. Hongchao Fan, Prof. Irina Rogozhina, Dr. Signe Hillerup Larsen
August 2020

Norwegian University of Science and Technology
Faculty of Engineering
Department of Civil and Environmental Engineering

NTNU
Norwegian University of
Science and Technology

Abstract

Glaciers are scientifically recognised as valuable indicators of the severity of regional climate change and have received considerable attention in recent years due to their high sensitivity to changes in the climate system and their status as the Earth's largest freshwater reservoir. Monitoring the evolution of glaciers is important, both to warn of hazards resulting from glacier melt and their potentially significant impact on human infrastructure, and to inform adaptation measures for future human activities through integrated water management. However, in-situ glacier monitoring is costly and multi-spectral optical satellite imagery – the current baseline for remotely-sensed glacier delineation – is limited both by cloud cover and by illumination conditions, making it nearly impossible to monitor glaciers near the poles during the polar night.

This study assessed the potential of Synthetic Aperture Radar (SAR) imagery to detect glaciated areas independent of weather and illumination. The SAR imagery used in this study is C-band Sentinel-1 imagery and Disko Island was picked as the study region. The seasonal evolution of SAR backscatter for all of Disko Island was characterised. Two classification methods, one supervised (Decision Tree) and one unsupervised (k-means), were examined in the context of glacier detection. Different geology types interfering with the task of glacier detection were identified. Comparisons were drawn to existing glacier databases and relevant products derived from multi-spectral optical imagery.

The analysis carried out in this study revealed that there are clearly distinguishable groups of backscatter amplitudes, with variation through the seasons for different geology types. The backscatter groups were found to not necessarily be linked to a certain surface type in either season. Of the two classification methods presented, the supervised decision tree method was found to give the best results. Imagery from the summer month of June was found best suited for the task of glacier detection for both methods as the melt and snow season present a challenge to the examined methods. Nevertheless, SAR imagery was found to provide an opportunity for automated year-over-year tracking of glacial ice extent.

Sammendrag

Breer er sensitive indikatorer på klimaendringer og utgjør den største samling ferskvann på Jorda. På grunn av isbreenes direkte respons til små forandringer i klimasystemet har breene fått mye oppmerksomhet de siste årene. For å kunne varsle om farene for folk og infrastruktur som følge av bresmelting slik som flom og skred, er det viktig å kontinuerlig overvåke breenes utvikling. Dessuten kan mulighetene for klimatilpasning utforskes, spesielt når det gjelder å utnytte avrenningen fra breer i sammenheng med vannkraft eller drikkevann. Nå til dags blir breene overvåket mest in-situ eller ved hjelp av optiske multispektrale satellittbilder. In-situ observasjonene er dog ofte en logistisk utfordring og dyre. Kvaliteten av de optiske observasjonene er sterkt avhengig av været (spesielt skydekket), og siden optiske instrumenter er passive, er det ikke mulig å observere regionene i løpet av polarnatten.

Denne masteroppgaven kartlegger postensialet av bruk av syntetisk apertur-radar (SAR) bilder for formålet av breobservasjoner. I motsetning til optiske overvåkningsinstrumenter kan SAR brukes til å ta bilder selv der det ikke er sollys, for eksempel på natta eller i overskyede områder. Bildene brukt i denne oppgaven er Sentinel-1 satellittbilder, og området som ble undersøkt er Disko Island i Grønland. Signalene mottatt fra overflatene av hele Disko Island ble analysert gjennom årstidene. Deretter ble to klassifiseringsalgoritmer, beslutningstre og k-means, testet til å automatisk gjenkjenne isbre i bildene. Overflatetypene som viste lignende refleksjon som breene ble identifisert, og resultatene ble sammenliknet med multispektrale optiske bilder.

Det ble identifisert forskjellige grupper av refleksjonsamplituder som varierer gjennom årstidene for de forskjellige overflatene, som imidlertid ikke kunne tilordnes en spesiell overflatetype. Beslutningstre-metoden viste seg til å egne seg best til automatisk gjenkjenning av is. Bildene fra sommermånedene var enklest å håndtere for algoritmen siden smeltingen om våren og gjenfrysing av breoverflatene om høsten, samt snødekket om vinteren gjorde det mer vanskelig for algoritmen å skille breene fra de andre overflatene. Uansett ble SAR identifisert som en mulighet for år-til-år gjenkjenning og overvåkning av breoverflater.

Acknowledgements

With this thesis, a long chapter is coming to an end. It was not always an easy ride, but I am happy and grateful for all the lessons learnt on the way and not least for all the people I met during this journey leading to my master's degree.

First, I would like to thank Professor Hongchao Fan for facilitating my thesis work by agreeing to supervise me on a topic he was not very familiar with. I would also like to express my appreciation to my co-supervisor at NTNU, Professor Irina Rogozhina, for volunteering to support a student out of her field of responsibility, for all the discussion and scientific insight given on the topic and on the whole thesis endeavour and pushing me towards high-quality research. Finally, I would like to thank Dr. Signe Hillerup Larsen from the Geological Survey of Denmark and Greenland for providing me with local insight as well as motivation and ideas through the course of my thesis.

Furthermore, I could not leave Trondheim before expressing my appreciation to the people who supported me throughout the journey of this master's degree; the NTNUI Diving group and the Capoeira groups I got to know both in Trondheim and Copenhagen, who always made me feel like I have a second home and place of warmth to return to wherever I went and who supported me through several exam periods and other challenges.

Lastly, I would like to thank my family for their positivity and their support throughout all these years, and for their proofreading of this thesis, hunting for each and every comma and period.

Galina Jonat
Trondheim, August 2020

Contents

List of Figures

List of Tables

Acronyms

$P(cloud)$ cloud probability. 25, 26

$P(snow)$ snow probability. 26

DEM Digital Elevation Model. 8, 12, 21, 27, 30, 66, 89

ECMWF European Centre for Medium-Range Weather Forecasts. 3

ESA European Space Agency. 3

FCC False Color Composites. 7

GLIMS Global Land Ice Measurements from Space. 2, 21, 27, 31, 42, 53, 71, 74

GRD Ground Range Detected. 22, 23, 29

GTN-G Global Terrestrial Network for Glaciers. 2

InSAR Interferometric SAR. 10, 11

IPCC Intergovernmental Panel on Climate Change. 1

IW Interferometric Wide Swath Mode. 21–23

MSI Multi-Spectral Instrument. 23

NDSI Normalised Differential Snow Index. 6, 7, 25, 26

NIR Near Infrared. 23, 26

NSIDC US National Snow and Ice Data Center. 2

RGI Randolph Glacier Inventory. 2

SAR Synthetic Aperture Radar. I, 3, 7–11, 13, 21, 69, 70, 72, 75, 77, 79, 81

SLC Single Look Complex. 22, 23

SNAP Sentinel Application Platform. 28

SWIR Short Wave Infrared. 23, 25, 26

TIR Thermal Infrared. 6

TOPSAR Terrain Observation with progressive Scans SAR. 21

VIR Visible Infrared. 23, 25

VNIR Visible to Near Infrared. 6

WGMS World Glacier Monitoring Service. 2

WKT Well-Known Text. 30

WV Wave Mode. 21

Chapter 1

Introduction

1.1 Motivation

Around 10 % of the Earth's land area is covered by glaciers or ice sheets, which support unique habitats and are interconnected with other components of the climate system through the global exchange of water, energy, and carbon.

Over the last decades, global warming has led to global-scale glacier mass loss, permafrost thaw, and decline in snow cover and Arctic sea ice extent. In their *Special Report on the Ocean and Cryosphere in a Changing Climate*, the Intergovernmental Panel on Climate Change (IPCC) expects this trend to continue in the near-term (2031–2050) due to surface air temperature increase (Pörtner et al., 2019). Accordingly, there will be an increase in disaster risk to human settlements and livelihoods in high mountain areas and in the Arctic due to future hazards, such as floods, fires, landslides, avalanches, unreliable ice and snow conditions, and increased exposure of people and infrastructure. To adapt to these changes and to potentially profit from them, context-specific monitoring and forecasting of the cryosphere is needed to inform adaptation decision-making, planning and implementation (Pörtner et al., 2019).

1.2 Glaciers and their role in the context of climate change

A glacier is defined as "multi-year surplus accumulation of snowfall in excess of snowmelt on land and resulting in a mass of ice at least 0.1 km^2 in area that shows some evidence of movement in response to gravity" (*NASA Earth Observatory Glossary*). A glacier may terminate on land or in water. Storing about 69 % of the Earth's freshwater, glacier ice is the largest reservoir of freshwater on Earth (Paul et al., 2015).

The spatial distribution of glaciers depends on

- climate (e.g., air temperature and precipitation),
- terrain (influencing direct and indirect solar radiation, ice flow, supply of rock debris and accumulation of snow from indirect sources),
- deposition of albedo-lowering aerosols and dust, and
- deposition of suspended atmospheric aerosols,

which leads to wide variations in regional and global behaviour of glaciers (Kargel et al., 2014). This high sensitivity to changing conditions positions glaciers as key indicators of climate change (Paul et al., 2015); as glacial ice is at its pressure melting point, the glacier needs to adjust its extent and surface elevation to adapt to changes.

Between 2006 and 2015, mass loss from Arctic glaciers, $(-47 \pm 16)\,\mathrm{Gt\,yr^{-1}}$, contributed to global mean sea level rise at a similar rate, $(0.6 \pm 0.1)\,\mathrm{mm\,yr^{-1}}$, as the melt of the Greenland Ice Sheet. With new extremes in the Arctic climate system, such as the 6 °C winter anomaly recorded from 2014–2018, and atmospheric circulation changes in the polar regions, glacier mass balance dynamics will vary across the Arctic, including different rates of retreat between eastern and western glaciers in Greenland's periphery (i.e., outside the Greenland Ice Sheet). This will again induce an increased surface melt of the Arctic glaciers due to the positive feedback from lower surface albedo (Meredith et al., 2019).

1.3 Monitoring glacier dynamics in a changing climate

Glacier outlines are the base for any glacier-specific calculations and modelling (e.g., length/volume/mass changes, run-off, ice thickness, future glacier development), giving rise to a high demand for a globally complete glacier inventory (Pfeffer et al., 2014 and Radić and Hock, 2013).

As early as 1990, at the Second World Climate Conference, the need for a worldwide monitoring service for climate data has been identified. Since 2009, the Global Terrestrial Network for Glaciers (GTN-G) has joined the efforts of worldwide glacier monitoring services such as the World Glacier Monitoring Service (WGMS), the US National Snow and Ice Data Center (NSIDC), and the Global Land Ice Measurements from Space (GLIMS) initiative. These and similar services collect glacier data, such as glacier inventories (e.g. the Randolph Glacier Inventory (RGI) or the GLIMS inventory), thickness measurements, and glacier fluctuations derived from in-situ and remotely sensed observations (Global Terrestrial Network for Glaciers, 2018).

Existing databases, though providing a good overview of the worldwide situation,

often contain relatively old data (e.g., on Greenland's Disko Island from 2001, 2006, 2016; see *GTN-G Viewer*) as they require considerable coordination and standardisation effort and also serve varying purposes, such as climate monitoring and hydrological modelling. To monitor and model glacier dynamics in a continuously changing climate, more frequent and recent observations are needed. However, glacier delineation often relies on manual work and thus glaciologist expertise, as not all glacial boundaries are easy to spot and can be hard to distinguish from their surroundings. This is especially the case for debris-covered glaciers. Accordingly, there exists a demand for technical solutions to reliably detect glaciers and glacier changes.

In the past, glacier outlines have mainly been derived from multi-spectral imagery (Kääb et al., 2014). The use of Synthetic Aperture Radar (SAR) has also proven to be useful in this context, especially due to its independence from weather and daylight. Both of these data sources are readily provided by the Copernicus programme, the European Union's Earth Observation programme, which is coordinated and managed by the European Commission and implemented and operated by actors such as the European Space Agency (ESA) and the European Centre for Medium-Range Weather Forecasts (ECMWF). Copernicus represents an independent and operational information capacity to the EU "to warrant environment and security policies and to support sustainable economic growth" (European Commission, 2020). The programme is served by a set of dedicated satellites, called Sentinels, and other contributing missions; these spaceborne observations are complemented by information collected from in-situ systems (e.g., ground weather stations). With the radar Sentinel-1A and -1B and the multispectral Sentinel-2A and -2B data freely available, and climate change monitoring being one of the main objectives of the programme, Copernicus offers a good opportunity for glacier monitoring.

1.4 Objectives

This study examines the potential of SAR imagery characterising glaciers in the region of Disko Island, Greenland, to evaluate a weather- and illumination-independent method of glacier monitoring in this region. With the presence of a large number of temporarily fast-moving glaciers (i.e., surge-type glaciers), inter-seasonal observations can improve the monitoring of surge events and predict their development and effect on livelihoods in the region. Extending the established method of multi-spectral observations by using SAR can potentially contribute to reliable and automated glacier tracking in the challenging environment that Disko Island presents.

1.4.1 Research Questions

The goal of this study is to assess the potential of SAR imagery for recognition of glaciers among different surface types. Here, the following questions are addressed:

1. Can unsupervised or supervised learning algorithms successfully and confidently detect glaciated areas? How well do they perform? Does the performance vary with the seasons?
2. Can unsupervised or supervised learning algorithms successfully and confidently detect other surface types and rule these areas out as non-glaciated?
3. Which types of surface are most commonly interfering with the recognition of glacial ice from radar?
4. Is there a pattern for SAR backscatter amplitudes from clean ice? Is there a pattern for SAR backscatter amplitudes from debris-covered ice? How does this pattern evolve over time (throughout the seasons and long-term)? If there are such patterns, are these unique for glaciers?
5. How well are the unsupervised and supervised learning methods performing compared to glacier detection using multi-spectral optical imagery?

Chapter 2

Background: Glacier monitoring with satellite remote sensing data

To date, continuous monitoring of glaciers is mostly done in-situ, assisted by remote sensing data. However, in-situ observations are costly, due to many glaciers being in remote and inaccessible locations. Glacial structures (e.g., crevasses) often make in-situ observations even harder, if not impossible.
Responding to these challenges has motivated the launch of several satellites with the observation of the cryosphere being one of the mission goals (e.g., CryoSat, GRACE, Sentinel-1, Sentinel-2, Envisat, ERS-1 and -2); thanks to this alternative data source, near real-time monitoring of glaciers worldwide is now possible. Figure 2.1 shows the different spectral ranges in which imagery can be obtained and how much of the energy in this spectrum is able to pass through the atmosphere, and marks the wavelength ranges for a number of common Earth observation missions. They will be discussed in more detail in the following sections.

Data collection, however, is only part of the glacial tracking process; the data must be converted into outlines before it can be scientifically useful.

Accurate glacier outlines are used

- for glacier area change analysis,
- as masks to analyse glacier velocity fields,
- for volume change estimation (in connection with altimetry and/or gravimetry), and
- as input and validation datasets within glacier modelling

(see e.g., Radić and Hock, 2013, Bamber and Rivera, 2007).

This chapter presents the different remote sensing data sources and techniques

for glacier monitoring, the most popular being multi-spectral and radar imagery, with additional benefits provided by thermal mapping and gravimetry.

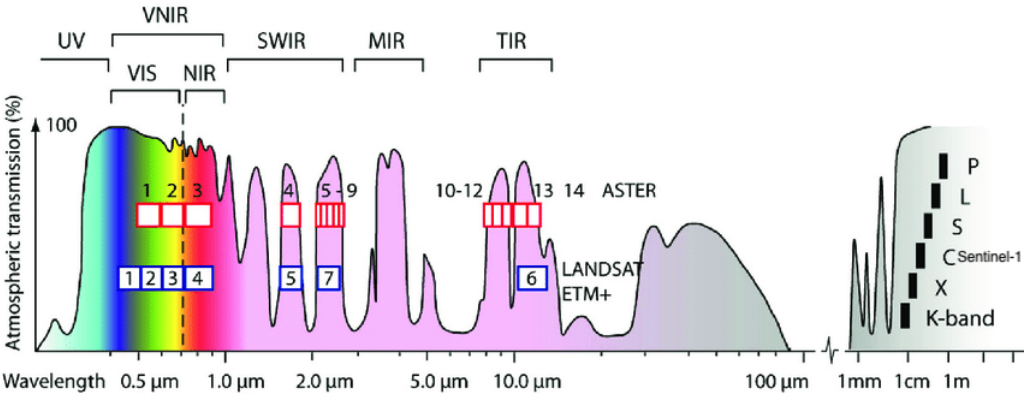


Figure 2.1: Spectral ranges and their atmospheric transmission Spectra: VIS: Visible Infrared; NIR: Near Infrared; SWIR: Short-wave Infrared, MIR: Middle Infrared, TIR: Thermal Infrared; Radar bands (Kääb et al., 2014)

2.1 Multi-spectral imagery

Multi-spectral optical imagery is the most common data source used to delineate glaciers. Such imagery is characterised by wavelengths from the Visible to Near Infrared (VNIR) up to Thermal Infrared (TIR), collecting radiation reflected and emitted from the observed surfaces. The spectral response recorded in multi-spectral imagery is used to describe and distinguish surface types and conditions. These wavelengths are more sensitive to changing atmospheric conditions, especially cloud cover. Nevertheless, multiple satellite missions serve as optical data sources for glacial delineation, including Landsat and Sentinel-2.

Manual delineation is still the most common method and is useful for highly complex delineation where expert knowledge is needed. The quality of the manual products depends highly on the mapper's knowledge, experience and the use of complex logical rules in the decision process. Manual correction is also often needed to complement or correct automated classification results. This process is time-consuming and the results are hard to reproduce (see Paul et al., 2013).

One popular method facilitating glacier mapping is *Image Algebra*, which calculates surface reflectance products from a set of bands profiting from reflection properties of the different surfaces. One such surface reflectance product is the Normalised Differential Snow Index (NDSI), indicating the presence of snow in a pixel. It is cal-

culated using two bands from multi-spectral imagery; for NDSI from Sentinel-2, see Section 3.2.2. As glaciers with intact firn accumulation are one of the few examples of summertime snow cover – at least at low altitudes – NDSI can be used as a starting point to identify glaciated areas, e.g., by applying a threshold-based approach.

Complementing NDSI are *False Color Composites (FCC)*, which display differences in reflectance of landscape features. Outlining glaciers from FCCs works well for clean ice and snow and can both facilitate manual delineation or be used for (reliable) automatic classification. It has also been proven successful for large-area mapping of ice flow over cloud-free regions (Fahnestock et al., 2016).

Finally, *spectral transforms* can be used to extract the most significant features from the different observation bands, e.g., performing a principal component analysis.

Challenges connected to glacier detection with optical data are

- frequent cloud cover and seasonal non-glacial snow,
- debris-covered glaciers (often needing expert post-processing),
- definition of glacier entity (purpose-based), and
- comparison with former inventories or field measurements

(Kääb et al., 2014).

2.2 Synthetic Aperture Radar

Synthetic Aperture Radar (SAR) is an active, side-looking illumination system. Radar pulses are transmitted to the side and reflected from different objects at different times, allowing for object discrimination.

In contrast to the short waves of multi-spectral optical imagery, the much longer SAR waves penetrate through clouds, allowing for weather- and illumination- independent observations. Due to the SAR operating at frequencies between 230 MHz to 300 GHz, radar waves can travel through dry, cold snow such that the received backscatter provides information about the material below the dry snow (e.g, glacier or bedrock). The most used bands on Earth observation satellites are C- and X-band, operating at frequencies between 4–8 GHz and 8–12.5 GHz, respectively, corresponding to wavelengths of 7.5–3.75 cm and 3.75–2.4 cm (see Figure 2.1).

A SAR system can operate in different polarisation modes. Polarisation describes the direction (horizontal (H) or vertical (V), see Figure 2.2) in which the microwave signals are transmitted and received. Using the echoes from multiple polarisation modes adds information to the reflectivity of different surfaces and can facilitate,

e.g., the generation of Digital Elevation Models (DEM).

Single polarisation transmits and receives the microwave signal in the same direction, i.e., HH or VV.

Co-polarisation operates in two pairs, i.e., HH and HV, VV and HV, or HH and HV.

Multi-polarisation systems both transmit and receive in both directions, i.e. HH, VV, HV, and VH.

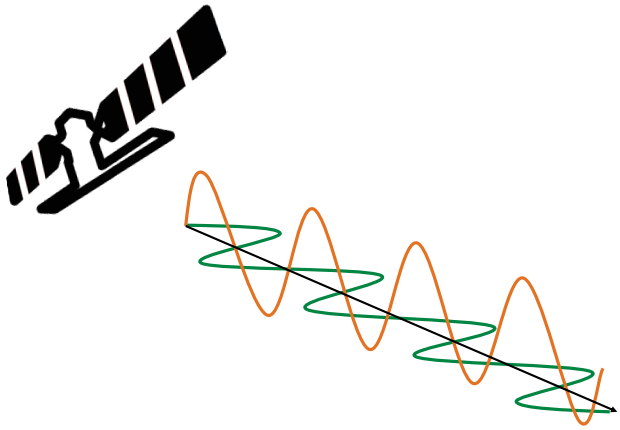


Figure 2.2: Polarisation angles along the satellite line of sight horizontal (H) in green, vertical (V) denoted in orange (redrawn from *Polarisation et contre-polarisation*)

A digital SAR image is a 2-D array formed by rows and columns, called mosaic. Each pixel in the mosaic contains a complex number that carries amplitude and phase information about the backscattered field from all scatterers within the resolution cell.

In general, HV polarisation has been found to help determine the volume of the snowpack whereas HH is more sensitive to surface properties (König et al., 2002). It has been found that dominant co-polarisation is a good indicator for ice lenses or ice layers, whereas dominant cross-polarisation (i.e., volume scattering) characterises homogeneous ice (Kääb et al., 2014). SAR imagery has also been used to detect subglacial lakes (Palmer et al., 2013) and supraglacial ponds on debris-covered glaciers (Watson et al., 2018).

The detected SAR image contains amplitude information of the backscatter signal, whereas the phase SAR image contains information about the phase change, ϕ, between the transmission and reception of the signal, due to scatterers being at different distances from the radar.

2.2.1 Amplitude-based observations

Backscatter coefficient, σ^0

The backscatter coefficient, σ^0, describes the normalised measure of the strength of the radar return signal from a (distributed) target. It is defined per unit area on the ground (*Sentinel-1 SAR Technical Guide: Glossary*).

The energy returning from a glaciated target is dependant on

- the surface roughness,
- the humidity/wetness of snow and ice, and
- the in-depth structure of snow and ice.

The returned energy provides a key insight into the nature of the observed snowpack: wheareas on dry snow, the radar microwaves penetrate into the snowpack and are reflected by many fine ice crystals and ice lenses, they are absorbed by wet snow and ice and reflected away from the sensor (Kääb et al., 2014).

These effects are also dependant on different observation parameters, such as

- the wavelength of the SAR signal,
- the incidence angle of the SAR signal,
- polarisation, and
- the topography and smoothness of the detected surface.

Backscatter time series can detect changes in snow and ice conditions which are related to the amount and variation of ice, air and water in the measured target. Average summer SAR images were found to have potential in assisting the process of mapping glacier outlines, especially in the presence of snow (Kääb et al., 2014).

The amplitudes of the returned backscatter can also be used for *Offset Tracking*, a technique to derive glacier velocities using image-matching algorithms (e.g., see Nascetti et al., 2016, Strozzi et al., 2002).
SAR altimetry observations over ice sheets, e.g., with Cryosat (see McMillan et al., 2018 and Gourmelen et al., 2018) can also monitor the thickness of the glaciers.

Using SAR imagery to detect glacier zones is an application of the imagery that has successfully been performed on glaciers in Norway, Svalbard, on the Tibetan

plateau, as well as in Antarctica (e.g., Huang et al., 2011, Zhou and Zheng, 2017, Zhang et al., 2019, Winsvold et al., 2018 and Barzycka et al., 2019).

Huang et al., 2011, Kundu and Chakraborty, 2015 and Winsvold et al., 2018 also used the amplitude data to track transient snow lines.

Yang et al., 2016 compared various approaches to extract glacier outlines and water-eroded stripes using TerraSAR (X-band, high-resolution SAR) imagery of the Dongkemadi glacier area on the Qinghai–Tibet Plateau. They tested – among other methods – a decision tree for glacier detection and a Canny Operator for detection of glacier edges. In the decision tree approach, a backscatter coefficient threshold on account of the difference between the backscatter coefficient of glaciers and their surroundings was used in the first step. The extracted potentially glacial area was then compared to threshold elevations to separate water-eroded stripes from the glaciated area.

2.2.2 Phase-based observations

Interferometric SAR

Interferometry calculates the phase difference between two images acquired from different look angles or at different times. By observing the same area from slightly different angles in both cases, interferograms can be used both for terrain altitude measurement (i.e., InSAR) and to monitor surface displacement and deformation (i.e., Differential InSAR) (Fletcher, 2007).

In contrast to the amplitude-derived backscatter coefficient, InSAR systems use the phase information received by the system. The SAR interferogram is generated by pixel-wise cross-multiplication of the first SAR image with the complex conjugate of the second. The interferogram amplitude is the pixel-wise multiplied amplitude of the two images, whereas the interferometric phase describes the phase differences between the two images.

The correlation (similarity) between the two images as a function of displacement of one relative to the other is described by the coherence coefficient, γ. A γ close to zero indicates total decorrelation, whereas a γ close to one means full correlation.

The coherence coefficient can itself be broken down as

$$\gamma = \gamma_{temp} \cdot \gamma_{geom} \cdot \gamma_{proc} \cdot \gamma_{thermal} \tag{2.1}$$

where γ_{temp} represents temporal changes in surface scattering properties (physical changes caused by, e.g., wind and precipitation), γ_{geom} represents geometric vari-

ations in the look angle and increases with the deformation gradient, γ_{proc} represents processing-induced errors (e.g., coregistration), and $\gamma_{thermal}$ is a noise component (e.g., speckle).

Repeat-pass interferometry can be used to derive surface deformation from SAR images obtained at different times on repeated orbits. There are several satellite systems allowing the derivation of these parameters; among them ERS-1, ERS-2, Envisat and Sentinel-1 (Fletcher, 2007).

As InSAR can retrieve sub-centimetre surface displacement measurements along the satellite line of sight, it is considered to be suitable to monitor slow-moving glaciers. This technique requires the difference between the images to be very little to retain the capability to reduce the noise, a small baseline (perpendicular line between the two satellites) and short revisit times (Michel and Rignot, 1999).

Analysing time series of interferometric coherence maps, Shi et al., 2019 managed to extract glacier boundaries through threshold-based segmentation and edge detection. They also found that there was a low level of coherence in wet snow and ice and that coherence maps during the ablation period are most suitable for identifying snow- and ice-covered areas.

Fang et al., 2013 use a supervised classification based on the glacial velocities derived from TerraSAR imagery. The approach relies mainly on distinguishing the moving glacial areas from their non-moving surrounding environment. For that, a coherence map and texture information is generated from the TerraSAR data, from which features then are detected, ranked and selected based on a random forest. Using only the velocity fields to identify glaciers, a 75 percent accuracy can be obtained. Zhou et al., 2010 are also using an approach based on the coherence map in glaciated areas and obtain 89 percent accuracy compared to the results derived from the Landsat Thematic Mapper.

2.2.3 Challenges identified in SAR processing

In contrast to multi-spectral imagery, SAR imagery is not as intuitive to interpret for humans. Another challenge in SAR processing is the presence of *speckle*, a granular noise caused by interference of waves reflected from many elementary scatterers. Speckle filtering has been proven to be more accurate with nighttime imagery and at times with low precipitation (Yang et al., 2016).

2.3 Further data sources and combined approaches

Thermal imaging can be used as a supplementary data source and has proven to be especially useful correcting for glacier delineation in debris-covered glaciers (Smith et al., 2014), being used in combination with a digital elevation model that identifies potential debris-covered areas. SAR has also been proven to improve performance mapping debris-covered glaciers (Robson et al., 2015). Smith et al., 2014 managed to design an algorithm to delineate both clean glacial ice and glacial debris tongues as well as lakes on a mountain range in Central Asia using Landsat, ASTER, MODIS (multi-spectral), and a DEM derived from radar imagery.

Chapter 3

Method and Scenario

As earlier mentioned, the goal of this study is to examine the opportunity of using SAR-Data for glacier detection. The overall workflow and main input data of this study are described in Figure 3.1.

In a first step, the radiometrically calibrated backscatter imagery of both the cross- and co-polarisation is examined for

- all surfaces,
- glaciated surfaces that are assumed to be "clean ice" (little or no debris cover), and
- all glaciated surfaces.

These characteristics are then brought in context with weather patterns, such as melt and snow depth. The development of the backscatter throughout the seasons is analysed for significant patterns that can serve the purpose of glacier delineation. Histograms are used to identify the main backscatter patterns of the glaciated areas. Finally, two learning methods, namely k-means clustering and a decision tree approach, are tested for their ability to identify surface types, and especially the glaciated surfaces, from SAR imagery.

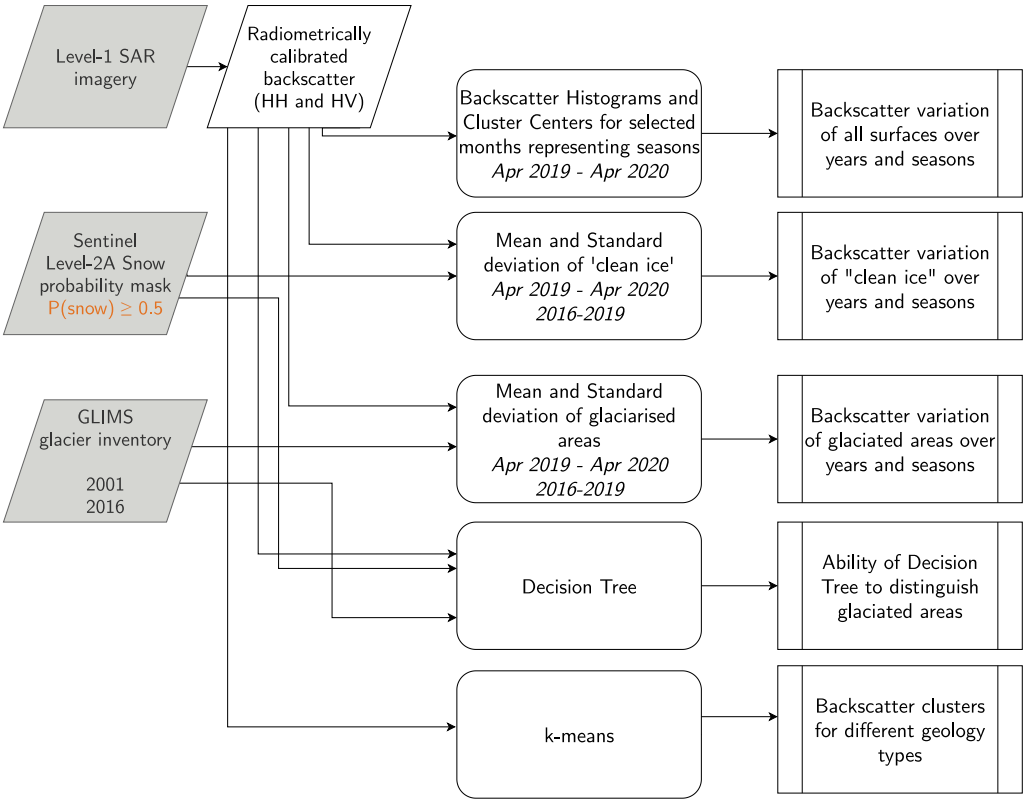


Figure 3.1: Study workflow input data, methods and work packages

3.1 Study area

Disko Island (70°N, 54°W) is the largest island in Greenland, located in the Central West (see Figure 3.2), with an area of 8578 km^2. The island offers a large variety of periglacial phenomena, i.e., ice- and sand wedges, stone polygons, mud circles, as well as palsa and pingo formations, in addition to a varied vegetation (*Arctic Station*). With the presence of rock glaciers and surge-type glaciers, this region is of a particular interest to monitor continuously for the evolution of its surface and landforms, also in context with livelihood in this region.

3.1.1 Geology

The distinct surface types identified on Disko Island and their coverage are presented in Table 3.1 and Figure 3.3. The island is part of the Tertiary volcanic province of West Greenland and is thus mainly made up by volcanic rocks. The ice coverage is about 22 %, mainly consisting of land-terminating glaciers that largely vary in size.

In total, Disko Island contains about 1000 glaciers and ice fields including Ser-

Geology Group	Area in km^2	Area coverage in %
Volcanic rocks	3420	39.73
Ice	1870	21.61
Glaciofluvial and marine deposits	278	3.22
Rock slides	131	1.51
Marine sandstone	16.5	0.19
Non-marine sandstone	12.7	1.48
Lakes	10.7	0.12
Undifferentiated deposits	2770	32.12

Table 3.1: Surface types on Disko Island

mersuaq Ice Cap (918 km^2) and Bræpasset Ice Cap (254 km^2) (Yde, 2011). Apart from these two big ice caps, the remaining glaciers are mostly smaller than 20 km^2 (*GLIMS Viewer*). Yde and Knudsen (Yde and Knudsen, 2007) estimate that 75 glaciers are surge glaciers, i.e., glaciers that at certain times advance up to 100 times faster than usual. One of these surge-type glaciers is Kuannersuit glacier: At a big surge event between 1995 and 1998, when it surged for 11 km and a large amount of sediment from the overriden area was moved (Roberts et al., 2009). Additionally, there is a large presence of rock glaciers, i.e., glaciers consisting more of rocks than snow and ice often due to slow-moving glacial ice being covered by debris (*Glacier Types: Rock*).

3.1.2 Weather during the observation period

Figures 3.4 and 3.5 show the air temperatures at Qeqertarsuaq Heliport (69°15′4″N 53°32′17″W), a location close to sea level in the South-West of Disko Island. For the long-term observations depicted in Figure 3.4, one can see that the temperatures are usually below 0 °C until at least April, and higher than 0 °C until September/October. In the summer, maximum temperatures until 20 °C can be measured. In the winter period of 2016 and 2018, maximum temperatures around 5 °C have been measured. The daily temperatures in the years 2019 and 2020 (see Figure 3.5) are below 0 °C until May, then constantly larger than 0 °C until the end of October 2019. From November 2019 to March 2020, the temperatures are between −20 °C and −12 °C. In beginning of April 2020, there are a few days when temperatures are higher than 0 °C which already may introduce melt.

Precipitation and snow depth are described in Figures 3.6 and 3.7.

The long-term precipitation and snow depth values (see Figure 3.6) show a high snow depth for the winter of 2017 with recorded snow depths up to 0.4 m, whereas the maximum recorded snow depths in the following winters are around 0.25 m. In general, there are days with high precipitation usually during autumn.

Looking at the snow depth values recorded for the period 2019–2020 (see Figure 3.7), one can see that lower altitudes were basically snow-free from May 2019 to October 2020; the largest snow cover being measured towards the end of April 2019. Following some minor precipitation events in the beginning of October, there are a few dozen centimetres of snow. More frequent precipitation events in the beginning of December 2019 lead to an increasing snow depth in Mid-December which, however, melts after a short period.

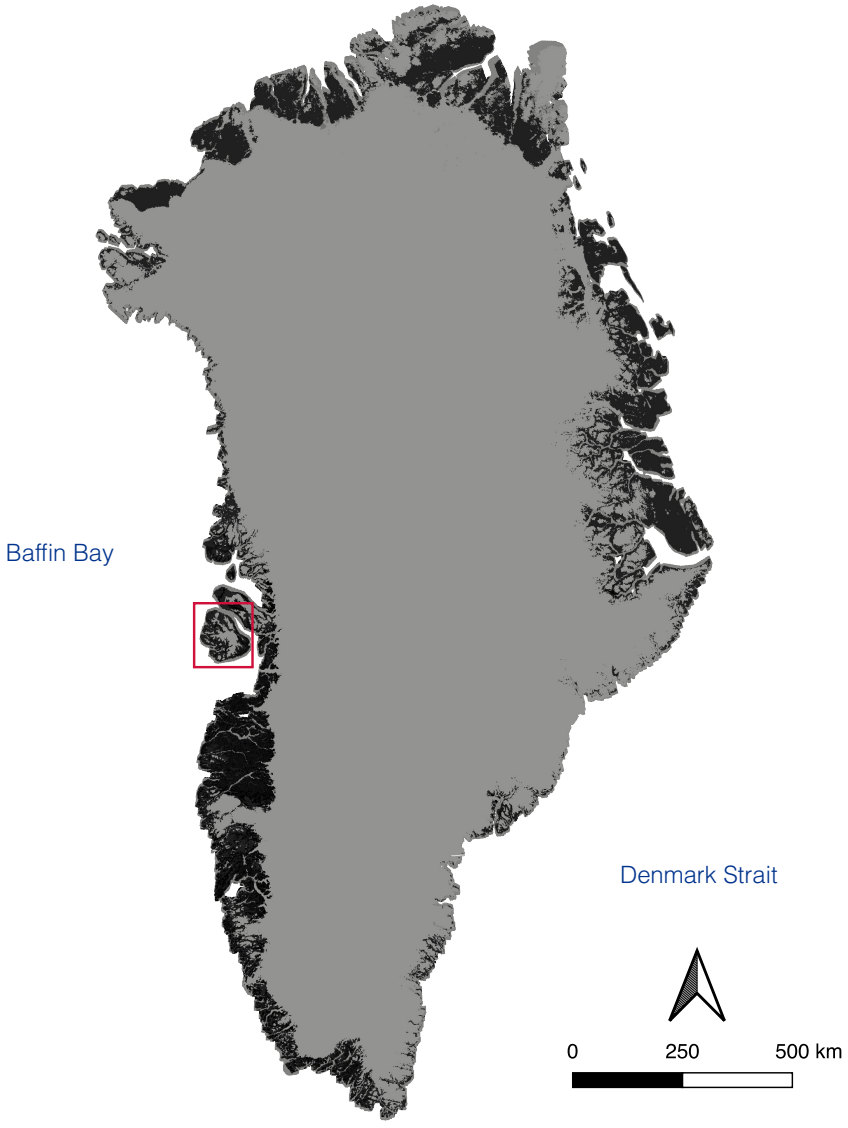


Figure 3.2: Location of Disko Island (red square)

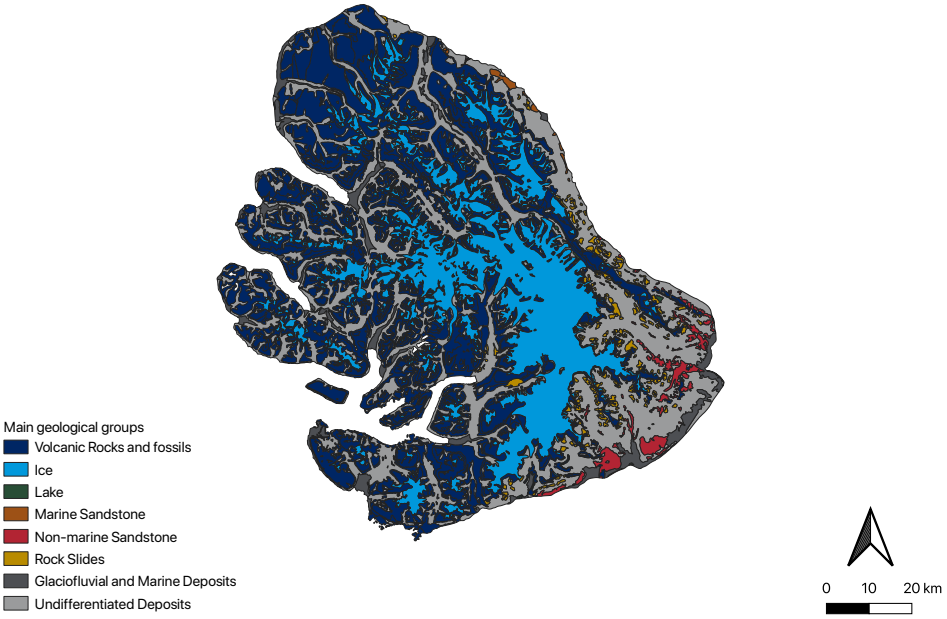


Figure 3.3: Geological map of Disko Island (Geoogical Survey of Denmark and Greenland, 2014)

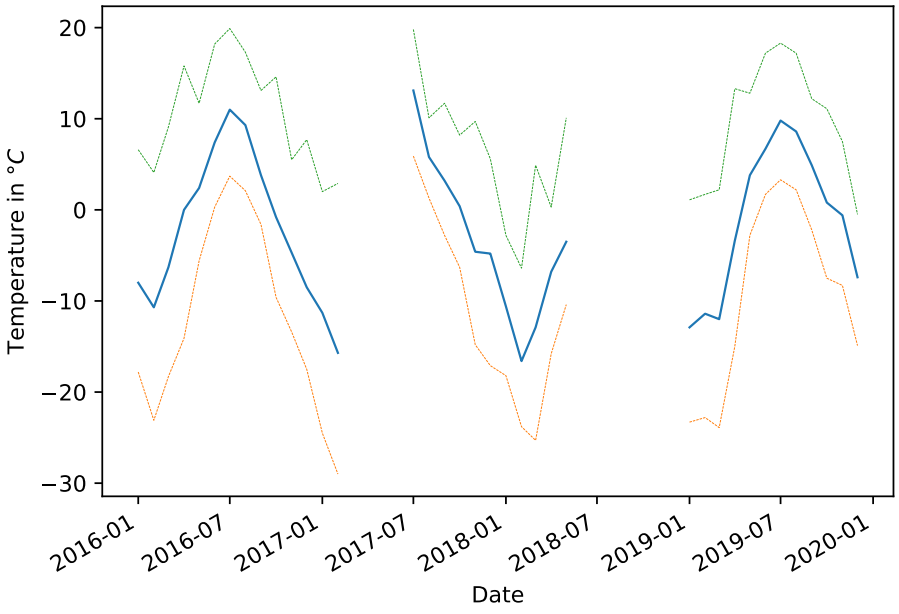


Figure 3.4: Monthly temperatures at Qeqerstarsuaq Heliport in the period January 2016 to December 2019; average (blue), minimum (orange) and maximum (green) monthly temperature (*DMI Vejrarkiv*)

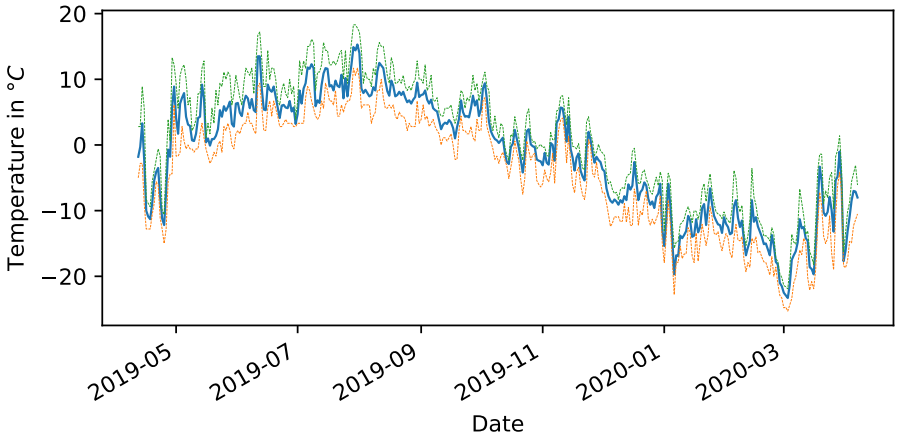


Figure 3.5: Daily temperatures at Qeqerstarsuaq Heliport in the period April 12, 2019 to April 18, 2020; average (blue), minimum (orange) and maximum (green) daily temperatures (*DMI Vejrarkiv*)

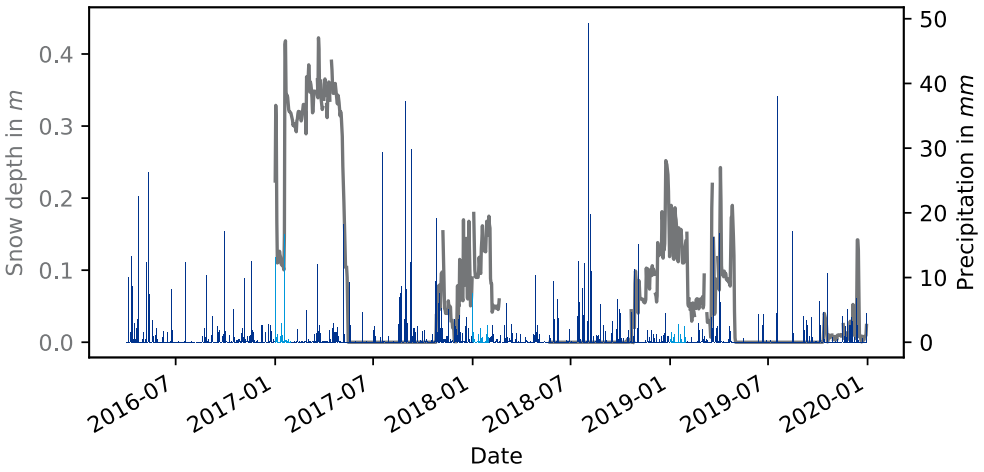


Figure 3.6: Daily precipitation and snow depth at Teleø station, June 2016 to December 2019; snow (azure bar), rain (dark blue bar); snow depth (black line) recorded from January 2017 (source: *ClimateBasis Data*)

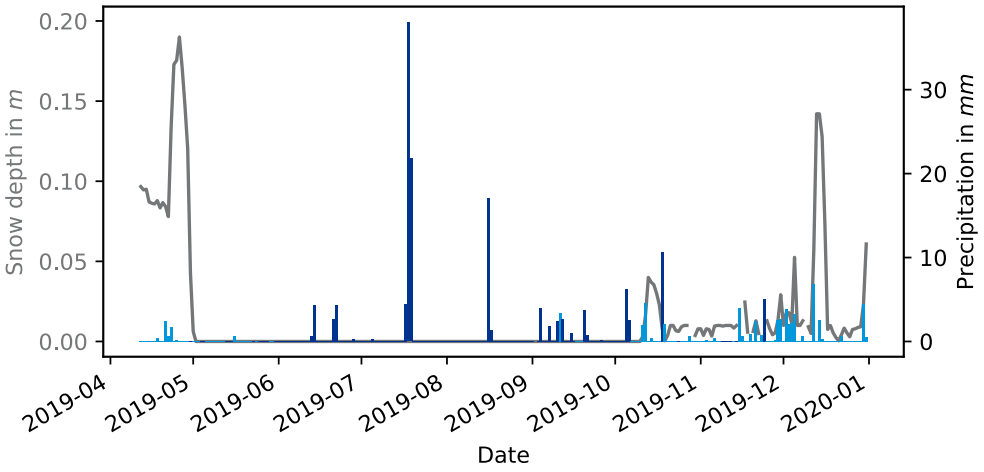


Figure 3.7: Daily precipitation and snow depth at Teleø station, April 2019 to December 2019; snow (azure bar), rain (dark blue bar), snow depth (black line) (source: *ClimateBasis Data*)

3.2 Data Sources

The data products used in this study are

- level-1 (decompressed and focused) radar imagery, being processed and analysed from the Sentinel-1 SAR as the dataset to be analysed,
- a level-2 snow mask derived from Sentinel-2 imagery, serving as both validation dataset for glaciated areas and as a primer for the contextual classification
- the Digital Elevation Model (DEM) ArcticDEM, and
- the GLIMS database containing ground-validated glacier outlines.

Level-1 data from Sentinel-1 is processed and analysed for its backscatter properties. Using the snow mask and the GLIMS database, the potential of Sentinel-1 imagery for reliable glacier delineation throughout the seasons is examined.

3.2.1 Sentinel-1

The main objective of the Sentinel-1 mission is the monitoring of the land and the ocean, supporting the different thematic streams of the Copernicus programme.

It is imaging all global landmasses, coastal zones, and shipping routes. In addition, it covers the global ocean with a dedicated acquisition mode, Wave Mode, to facilitate the mapping of wave currents and sea ice.

The mission operates two polar-orbital satellites, operating on the same orbital plane at 693 km which allows for 'reliable, repeated, wide area monitoring' (*Sentinel-1 SAR Technical Guide*). Sentinel-1A and -1B were launched in April 2014 and April 2016, respectively. With the launch of the second satellite, data can be acquired at a six-day frequency over the same area. Two other spacecraft (Sentinel-1C and Sentinel-1D) are planned to replace the first two satellites at the end of their operational lifespan.
The main payload of the Sentinel-1 satellites is a C-band (centre frequency 5.405 GHz) SAR system which supports operation in dual polarisation (HH+HV, VV+VH) implemented through one transmit chain (switchable to H or V) and two parallel receive chains for H and V polarisation in all acquisition modes, except for WV.

Sentinel-1 imagery is acquired in four acquisition modes with varying swath size and ground resolution. In this study, imagery acquired in Interferometric Wide Swath Mode (IW) is used. The imagery consists of 20 km swaths, with 5 x 20 m ground resolution. Sentinel-1 is using the Terrain Observation with progressive Scans SAR (TOPSAR) mode for data acquisition, scanning the terrain in bursts which are synchronized (< 5 ms) between repeat-pass data takes (see Figure 3.8). IW imagery

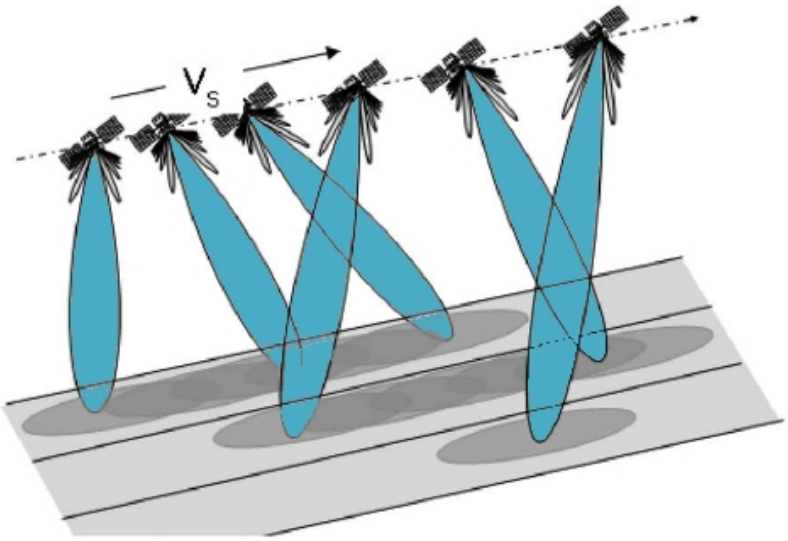


Figure 3.8: TOPSAR Acquisition mode with Subswaths, v_s representing the movement of the satellite (Image courtesy of *Sentinel-1 SAR Technical Guide*)

consists of the three sub-swaths, providing continuous ground coverage. Steering the beam both in range and azimuth direction (i.e., switching between sub-swaths and scanning back and forth within sub-swaths), a nearly uniform Signal-to-Noise ratio and Distributed Target ambiguity ratio can be reached, also reducing speckle (De Zan and Guarnieri, 2006).

Level-1: Ground Range Detected and Single Look Complex

Level-1 products are focused baseline products that have been preprocessed with a Doppler centroid estimator and SLC focusing. Noise and internal calibration errors have been removed, and the products contain echo source packets, as well as attitude and orbit information.

Ground Range Detected (GRD) The GRD imagery contains information on the amplitude of the radar backscatter and is used to derive the backscatter coefficient, σ^0. The initial resolution of the IW GRD imagery is 10 m × 10 m in range and azimuth direction. GRD backscatter intensity is already multi-looked and the slant range coordinates have been projected to the ground range coordinate system, WGS84. Due to ground range detection, phase information is lost.

Single Look Complex (SLC) The SLC imagery carries both amplitude and phase information of the radar backscatter. It is provided in slant range geometry (along the line of sight of the radar system) and single look in each dimension using full available signal bandwidth. The ground resolution of the SLC imagery is 5 m × 20 m in range and azimuth direction. The SLC IW data product consists of one image per subswath and per polarisation channel. Each subswath again consists of a combination of bursts, where each burst has been processed as a separate SLC image (individually focused). The bursts are separated with a black fill demarcation.

The imagery used in this study are Level-1 GRD products observed in IW. Over Disko Island, signals are obtained in HH, i.e., or horizontal transmit and horizontal receive, and HV (horizontal transmit and vertical receive) polarisation.

3.2.2 Sentinel-2

The main goal of the Sentinel-2 mission (*Sentinel-2 User Handbook* 2013) is the monitoring of land surface changes.

Sentinel-2A and -2B were launched in June 2015 and March 2017, respectively. The two satellites, operating in a sun-synchronous polar orbit, phased at 180° are providing systematic coverage of

- all continental land surfaces (including inland waters) between latitudes 56° South and 84° North,
- all coastal waters up to 20 km from the shore,
- all islands greater than 100 km,
- all EU islands,
- the Mediterranean Sea, and
- all closed seas (e.g. Caspian Sea).

In addition, the Sentinel-2 observation scenario includes observations following member states or Copernicus Services requests (e.g. Antarctica, Baffin Bay). Most areas are covered with a temporal resolution of five days (10 days for Antarctica and other, smaller islands).

The main payload is a passive Multi-Spectral Instrument (MSI) obtaining (sunlight) radiation in 13 bands with different ground resolutions as depicted in Figure 3.9 - in the Visible Infrared (VIR) (bandwidth: 350 nm to 780 nm) and Near Infrared (NIR) (bandwidth: 780 nm to 2500 nm) and Short Wave Infrared (SWIR) frequency range (bandwidth: 30–185 nm) (*Sentinel-2 User Handbook* 2013).

In this study, a Level-2A snow/ice mask product, obtained from imagery acquired

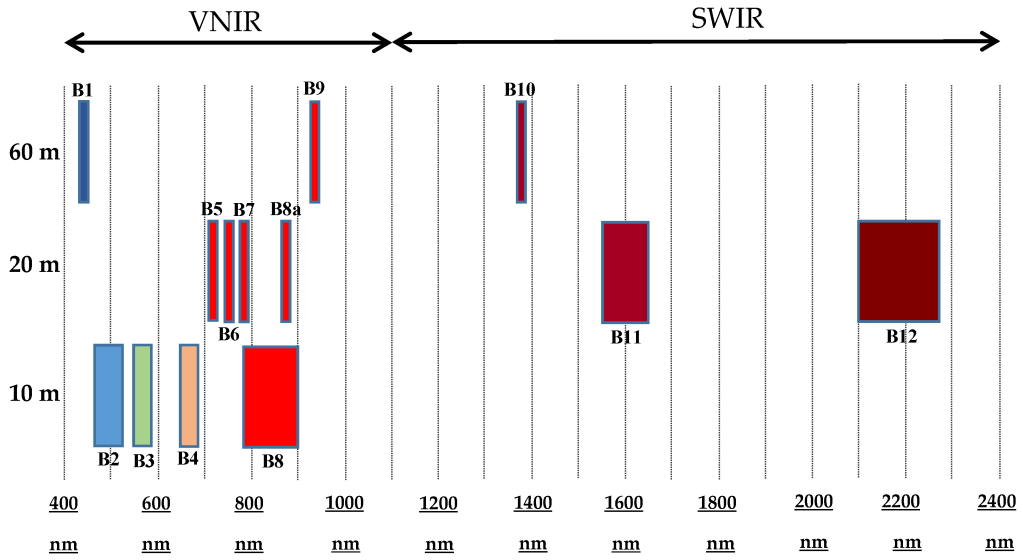


Figure 3.9: Sentinel-2 multi-spectral instrument bands; wavelengths and resolution (Gascon et al., 2017)

in August/ September of each year, serves as validation data for the clustering results and as as a primer for the contextual classification.

Level-2A: Snow/Ice mask

Level-2 data derivation from Sentinel-2 is performed by the Sentinel-2 ground segment. The scene classification operation classifies image data into the following classes:

- clouds (four subclasses),
- shadows,
- cloud shadows,
- vegetation
- soil/deserts,
- water, and
- snow.

One by-product generated throughout this scene classification is a snow/ice probability mask. In this study, this mask derived from imagery in September in the years 2019 is going to be used as a validation dataset for the k-means classification results and as a primer for the contextual classification.

The snow/ice probability mask has been generated from orthoimages projected in UTM/WGS84 from which defective, non-existent and saturated pixels have been

excluded, and resampled with a constant ground sampling distance of the ground resolution of the respective band), that have been resampled to a 60 metre resolution, as follows (see Figure 3.10, and *Level-2A Algorithm Overview*):

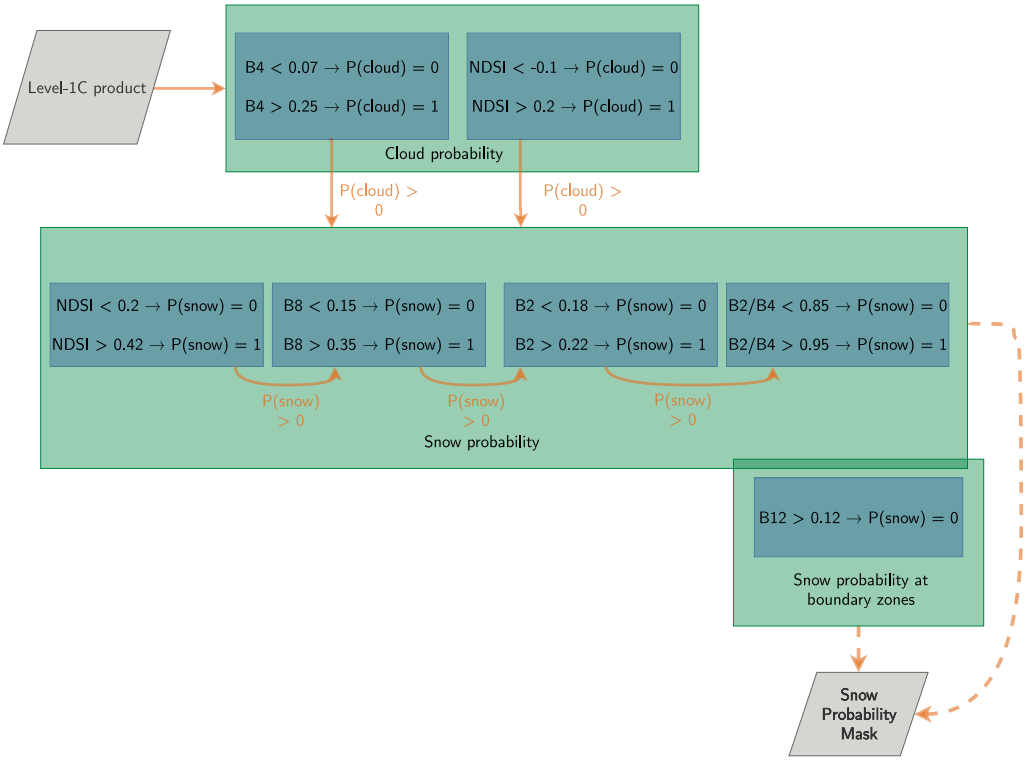


Figure 3.10: Generation of level-2A snow probability mask (*Level-2A Algorithm Overview*)

In a first step, a cloud probability ($P(cloud)$) is assigned to each pixel depending on their reflection from bands 4 (Red), 3 (VIR) and 11 (SWIR)). Threshold values are applied to assign reflectance values a $P(cloud)$ of 0 (lower threshold) and 1. Any reflection values in between these thresholds are assigned a $P(cloud)$ following a linear distribution.

At first, band 4 values are processed. The threshold for a $P(cloud)$ larger than 0 is 0.07. Every band 4 reflected pixel with a reflection intensity larger than 0.25 is assigned $P(cloud)$= 1.0.

Bands 3 and 11 are used to calculate the Normalised Differential Snow Index. This ratio uses the reflection properties of snow and clouds in the VIR and SWIR band. Whereas both snow and clouds reflect VIR frequencies, SWIR frequencies are ab-

sorbed by snow but reflected by clouds. The NDSI is calculated as follows:

$$NDSI = \frac{B3 - B11}{B3 + B11} \tag{3.1}$$

The thresholds used here are -0.1 for a $P(cloud)$ of 0.0, and 0.2 for a $P(cloud)$ of 1.0.

All values with a $P(cloud)$ larger than zero (from either of the two thresholds examined) are then passed on to the snow probability calculation step.

The second step is the calculation of the snow probability ($P(snow)$) on pixels that have been passed on from step 1 and that are located at locations where snow has been detected within the past ten years. The same approach as in step 1 is applied here, assigning a $P(snow)$ according to threshold values looking at the different bands and band ratios. The bands examined in this step are 2 (blue), 3 (green), 4 (red), 8 (NIR) and 11 (SWIR). However, the threshold steps here apply as filters, passing only the pixels to the next filter which have been assigned a $P(snow)$ larger than zero:

- The first filter looks again at the NDSI. A lower threshold of an NDSI is set to 0.2 and the upper threshold ($P(snow)$= 1) is set to 0.42.
- The second filter examines the NIR band 8 with a lower threshold of 0.15 and 0.35.
- The third filter looks at the blue band 2: $P(snow)$ 0 is set for all pixels with a B2 reflectance lower than 0.18 and a $P(snow)$ of 1 is assigned to all values larger than 0.22.
- The fourth filter looks at the ratio of the reflectance between band 2 and band 4. A low B2/B4 ratio indicates the presence of water bodies. Any ratio below 0.85 is thus assigned a $P(snow)$ of zero, whereas ratios larger than 0.95 are assigned a $P(snow)$ of 1.

Finally, extended boundary zones of the areas identified as snow-covered areas throughout the filters in step 2 are then examined to detect pixels that may have been detected as clouds in step 1 as they contained "mixed" pixel values, due to the presence of snow and ground in the same pixel. For this step, band 12 values of these extended boundary zones are analysed and all pixels that contain a band 12 reflection larger than 0.12, are assigned a $P(snow)$ of zero.

The final output is a snow confidence mask, containing a ‘combination of the snow probabilities in each pixel’.

3.2.3 GLIMS Glacier Database

The Global Land Ice Measurements from Space (GLIMS) initiative's 'objectives are to establish a global inventory of land ice, including surface topography, to measure the changes in extent of glaciers and, where possible, their surface velocities. [...] [By] long-term monitor[ing] of the world's glaciers, build[ing] a base of historical data, [to] detect climate changes early, [to] predict and avoid hazards to human communities living in the proximity of glaciers' (Raup et al., 2007). Glaciers are primarily monitored using the ASTER instrument on the Terra spacecraft, complemented by in-situ mass balance measurements and other air- and spaceborne imaging systems.

3.2.4 ArcticDEM

The Digital Elevation Model (DEM) *ArcticDEM* has been used in various processing steps for terrain correction purposes.
It is described as 'an NGA-NSF public-private initiative to automatically produce a high-resolution, high quality, digital surface model (DSM) of the Arctic using optical stereo imagery, high-performance computing, and open source photogrammetry software'. According to the project description, this DEM is meant to be a collection of time-dependant elevation model, also providing a frame for the infrastructure to process more and newer high-quality imagery over time (*ArcticDEM*).

The ArcticDEM contains elevation data of all land area north of 60°N, including Greenland, the State of Alaska in entirety, and the Kamchatka Peninsula of the Russian Federation.

The model has been generated from the high-resolution optical imagery obtained by the WorldView-1, WorldView-2, and WorldView-3 satellites by applying stereo auto-correlation techniques to overlapping image pairs.

For this study, version 3.0 (release 7) of the the mosaic files - compilations of multiple co-registered, blended and feathered strips - have been downloaded at a 10 metre ground resolution; using the *ArcticDEM Explorer*. It is projected to *EPSG:3413*, NSIDC Sea Ice Polar Stereographic North, and referenced to the *WGS 84* horizontal datum.

3.3 Data Processing

3.3.1 Input data

The parameters derived for this project work is the radiometrically calibrated backscatter, σ^0. Processing has been carried out in Sentinel Application Platform (SNAP), mainly via the Graph Processing Framework (GPF) which can process data through terminal commands and via XML processing chains created by the user.

The products used for the analysis are described in Table 3.2 and were downloaded from *ASF Data Search Vertex* and *Copernicus Open Access Hub*. Their footprint is outlined in Figure 3.11.

Acquisition Time	Satellite	Pass Direction	Relative Orbit No.	Polarisation
March, June, September, December representing seasons from 2016-2019	Sentinel-1A	Ascending	90	HH HV (HV from 2017)
12-day repeat imagery from April 12, 2019 to April 18, 2020	Sentinel-1B	Descending	171	HH HV

Table 3.2: Details of Sentinel-1 imagery used in this study

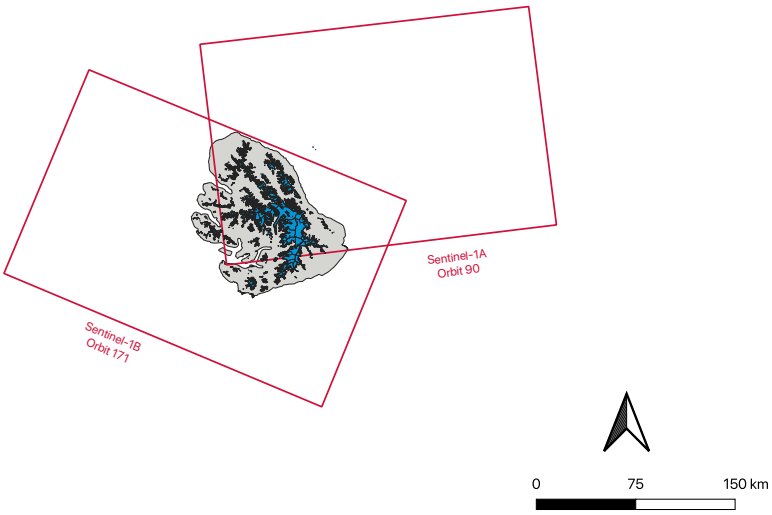


Figure 3.11: Footprint of Sentinel-1 imagery used in this study

Radiometrically calibrated backscatter coefficient

The radar backscatter coefficient, σ^0, for both the cross- and co-polarised (HV and HH) data was derived from Sentinel-1 Level-1 GRD imagery following the approach of Winsvold et al., 2018 (presented in Figure 3.12), using the graph processing tool, *gpt*, with the Sentinel-1 Toolbox in the *Sentinel Application Platform (SNAP)*.

The steps are described in Figure 3.12:

1. A *precise orbit file* is applied to refine the initial orbit state vectors with more precise, post-processed orbit files (available a few weeks after the product has been generated), which contain accurate satellite position and velocity information.
2. A noise lookup table (which is provided for Level-1 products) removes the *thermal noise*, i.e., the distortion of the radar reflectivity caused by the instrument itself.
3. A *radiometric calibration* step removes significant radiometric bias, such that the backscatter value in the pixel truly represents the radar backscatter of the reflecting surface.
4. The radiometrically calibrated backscatter image was *subset* given a predefined

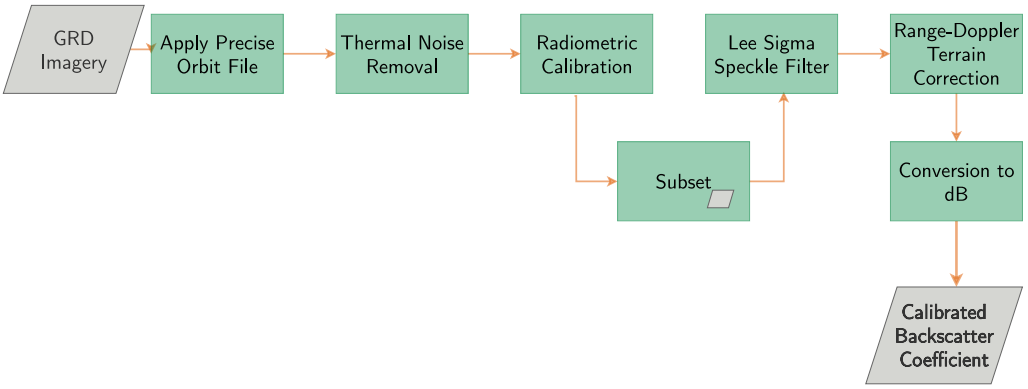


Figure 3.12: Derivation of radiometrically calibrated backscatter

polygon in Well-Known Text (WKT) format (in WGS84).

5. A *Lee Sigma Speckle filter* was applied to remove speckle, the granular noise 'due to interference of radar signals reflected from many elementary scatterers in resolution cell' (Lee et al., 1994), averaging the variations of grey level in adjacent resolution cells (assuming a Gaussian distribution).
6. *Range-Doppler terrain correction* compensates for geometric distortion caused by topography, i.e., foreshortening and shadow. It also orthorectifies the image and geocodes the SAR scenes from images in radar geometry using metadata information (i.e., orbit state vector information, radar timing annotations, slant to ground range conversion parameters), as well as the input DEM, in this case the *ArcticDEM*. As the ArcticDEM also has a ground resolution of 10 m × 10 m, the resolution remains as the original imagery.
7. The linear backscatter values were converted to decibels, $\sigma^0_d B$.

Due to limited working memory capacity, the processing steps were split in two parts, such that steps 1 to 4 were carried out in the first processing part and steps 5 to 7 were applied to the subset image.

3.3.2 Validation Data

The validation datasets used for this study are a glacier dataset from the GLIMS inventory and a Snow/ Ice Mask derived from Sentinel-2 imagery, downloaded as a level-2A product. Their footprint is outlined in Figure 3.13.

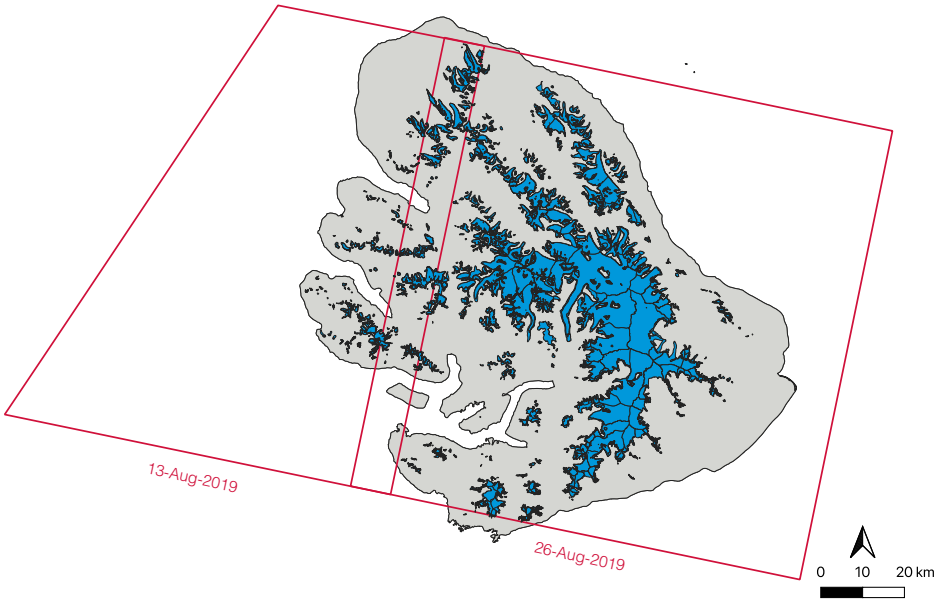


Figure 3.13: Disko Island with validation data: GLIMS inventory (azure) and the extent of the two snow probability masks (red)

GLIMS inventory for Disko Island

The glacier outlines obtained by the GLIMS initiative (GLIMS and NSIDC, 2005 and Cogley et al., 2015) were downloaded using the *GLIMS Viewer*. It contains glacier outlines from two years, namely 2001 and 2016. In total, there are 1094 glaciers of various sizes (0.020 km^2 to 135.317 km^2) recorded in the database, most of them being between 5 and 10 km^2.

Snow/Ice probability mask

The snow confidence mask used in this study has been downloaded for imagery from two Sentinel-2B images. The main criteria for the selection of the ice mask data was a low cloud cover percentage, i.e., a cloud cover percentage lower than 15 %. As it can be seen in Table 3.3, the total snow ice percentage is lower than the 19 % at Disko Island. This is mainly due to the image boundaries being larger than the island extent, but can also be due to the fact that some of the glaciers on Disko Island are debris-covered and thus hard to recognise on multi-spectral imagery (see Figure 3.13).

Acquisition Time	Satellite	Pass Direction	Relative Orbit No.	Cloud Cover [%]	Snow Ice [%]
2019-08-26 15:38:19Z	Sentinel-2B	Descending	11	10.89	10.46
2019-08-13 15:28:19.024Z	Sentinel-2B	Descending	111	1.12	0.5

Table 3.3: Details of Sentinel-2 level-2A data containing ice mask

Before it could be used as a validation dataset, the snow/ice probability mask was preprocessed in *QGIS Geographic Information System. Open Source Geospatial Foundation Project.* In a first step, the two images were merged. Then, a so-called "Majority Filter" from the *System for Automated Geoscientific Analyses (SAGA)* toolbox was applied to create more homogeneity in the image, thus, filling smaller "holes" in the classified areas. Here, the square search mode and a 10% threshold were applied.

3.3.3 k-means Clustering

Clustering partitions data into groups, called clusters. The *k-means* clustering algorithm is a statistical method to partition a number of observations into k clusters, maximising the distance between the values in between the clusters while minimising the variance within the cluster. Each observation belongs to the cluster with the closest mean.

3.3.4 Decision Tree Classification

Decision trees are a common supervised learning method used for classification. The purpose of the decision tree is to create a model to predict the value of a target variable based on sequential decisions inferred from a number of data features. Probabilities are assigned to the events, and values are determined for each outcome. A major goal of the analysis is to determine the best decisions.

The decision tree consists of nodes forming a rooted tree with the end notes – called leaves – being assigned to one of the target classes. The depth of the decision tree increases its complexity and decreases the user's ability to interpret the decision paths. It is applied both for binary classification and multi-output classification. This method is intuitive as its decision rules are easy to visualise and thus interpret. At the same time, it does not require data preparation. Designing a decision tree, the phenomenon of overfitting – creating complex trees that capture the training data

well, but do not generalise the task enough – should be avoided, in the easiest way by limiting the depth of the tree and the number of input features to which the decision rules are fitted.

The criteria to establish decision rules for the decision tree model in this study is the *Gini Index* impurity measure, which describes the probability of a specific feature to be classified incorrectly when picked randomly. A Gini Index of zero means that all elements in this node belong to a certain class whereas a Gini Index of one indicates a random distribution of the elements across various classes. As quality measure, the decision trees were cross-validated by splitting the input data in a training and test set with the test set being a third of the size of the input. This way, the tree's performance can be measured both against its input data and overfitting can be detected before applying the derived decision rules to possibly more different datasets.

As the surface type distributions of Disko Island are highly imbalanced, i.e. some surface types covering a much larger area and thus providing more pixels as input samples for the decision tree classification, the algorithm was trained with the option to assign weights to the different classes inversely proportional to their frequency in the input data, (Lior and Oded, 2008).

Chapter 4

Results

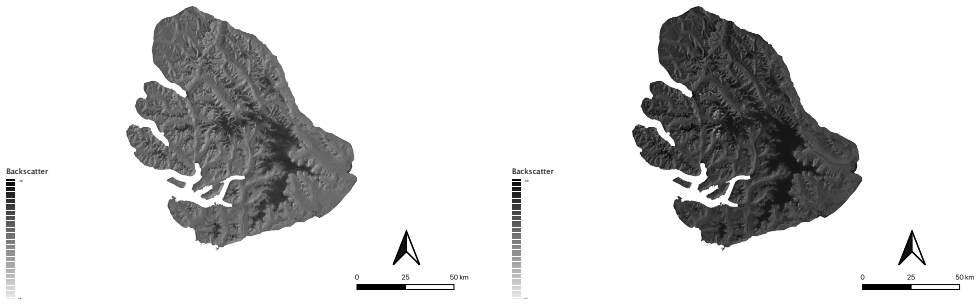
Following the processing steps described in Section 3.3, the intermediate and final products are presented here.

First, the radiometrically calibrated backscatter imagery (output from the process described in Figure 3.12) is presented. Then, this radiometrically calibrated backscatter is analysed quantitively first for all surfaces, then for clean ice identified by the ice mask and finally all glaciated areas. Based on the previous analysis, one unsupervised and one supervised classification method are tested towards their performance to recognise ice and other geology types.

4.1 Radiometrically calibrated backscatter

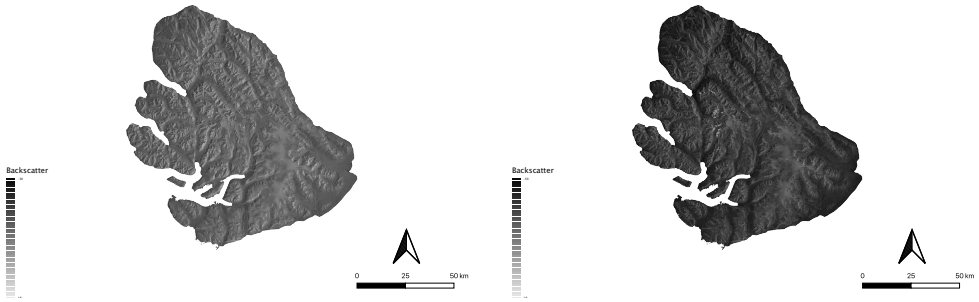
The radiometrically calibrated backscatter is displayed in Figures 4.1 and 4.2 for one day in June and December, respectively.

In general, the backscatter intensity is more distinct in June than in December, especially looking at the HH polarisation (Figure 4.1a and 4.2a). There is also a clear change in backscatter patterns from summer to winter for most of the surface. In the June imagery, the Sermerssuaq ice field in the West, spanning over the middle to the South of the island, is clearly distinguishable in both HH (Fig. 4.1a) and HV (Fig. 4.1b) polarisation.



(a) HH polarisation **(b)** HV polarisation

Figure 4.1: Radiometrically calibrated backscatter image on June 11, 2019



(a) HH polarisation **(b)** HV polarisation

Figure 4.2: Radiometrically calibrated backscatter image on December 20, 2019

4.2 Seasonal variations of backscatter for all surface types

To determine how the backscatter of all surfaces is varying over the seasons, the backscatter was analysed quantitavely using histograms, looking at HH and HV as well as the ratio between the two polarisations. Imagery from March, June, September and December – representing the seasons – were examined in this context. Histograms count the number of pixels that belong to each *bin*, a value range, and give thereby a good indication of the backscatter value distribution.

Figures 4.3 and 4.4 show the distribution of HH and HV backscatter values for selected months representing the seasons (June, September, December, March). It gets obvious that the backscatter changes significantly throughout the seasons for both polarisations.

Looking at the development of the HH backscatter distribution (Fig. 4.3a–4.3d), the backscatter values can be grouped in roughly six groups with different density

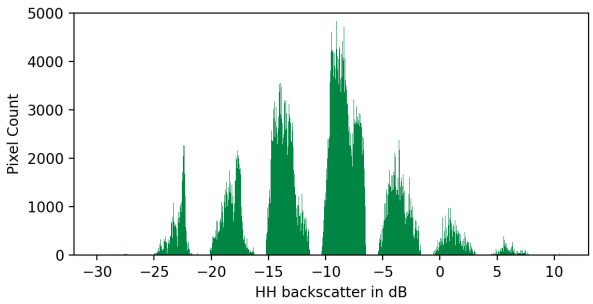
distribution. The spectral range of the backscatter for all surfaces is constant between 8 dB and −25 dB with most of the pixels in the range −5 dB to −15 dB. From June to September (Figure 4.3a and 4.3b), there is a shift towards the two "central" backscatter groups. The main "group" backscatter ranges are constant. There is a similarity between the backscatter recorded during the seasons with snow cover, i.e., December and March. Here, the backscatter groups are however shifted towards slightly higher values in March compared to December.

A variation in the backscatter throughout the seasons is also recognisable in the HV polarisation. Here, the backscatter distribution in June and March, and in September and December are similar to each other. The backscatter spectrum for all surfaces reaches from about −32 dB to about −5 dB and there are two to three ranges that describe the backscatter of most pixels. Whereas the backscatter in June and September can be grouped into 7 to 8 groups, for March and December there are only 6 main backscatter groups. Looking at the shift from June to September, the backscatter values shift again towards the more "dominant" groups, this is especially visible for the lower backscatter range.

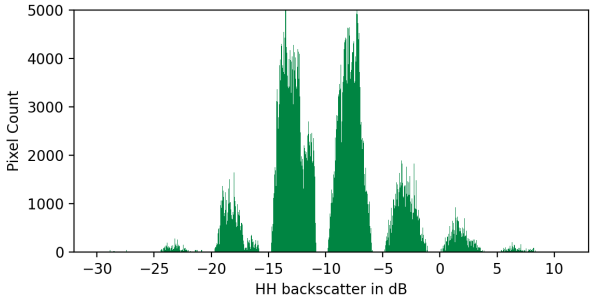
Comparing the backscatter for HH and HV, one can see that the backscatter distribution in June for HH (Figure 4.3a) and HV (Figure 4.4a) are similar, but shifted in the backscatter amplitudes, whereas there are less similarities comparing HH and HV histograms for the other months. Whereas in HH there seems to be one "dominant" backscatter group, the backscatter distribution is more spread in HV in December and March.

Comparing HH and HV backscatter recorded in December – depicted in Figures 4.3c and 4.4c – the pixel distribution is not just shifted between the two polarisations, but varies even within the "groups". This is for example visible in the backscatter group −5 to −9 dB in HH compared to the group −14.5 to −17 dB in HV as well as −10 to −14.5 dB in HH compared to the group −18 to −22 dB in HV.

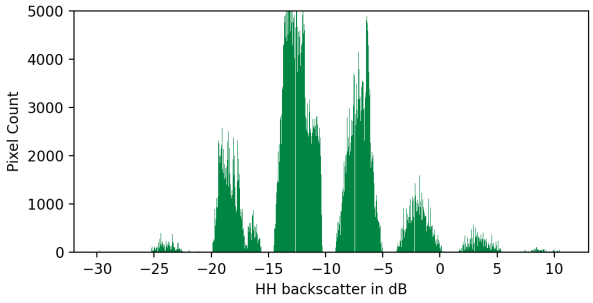
In total, backscatter can be identified as a seasonally varying property for the entire image, i.e., all surfaces, capturing the seasonal variations in ground cover to some extent.



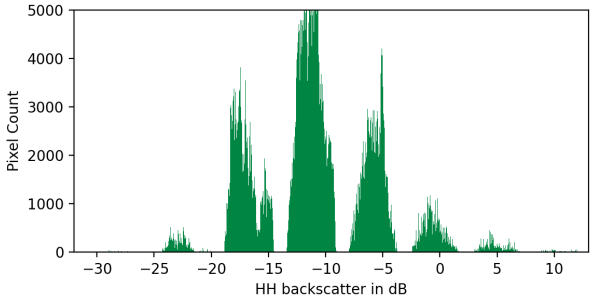
(a) June 11, 2019



(b) September 15, 2019

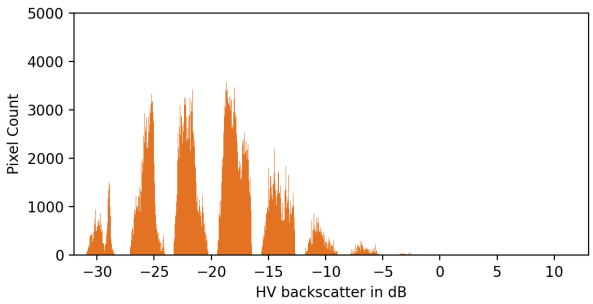


(c) December 20, 2019

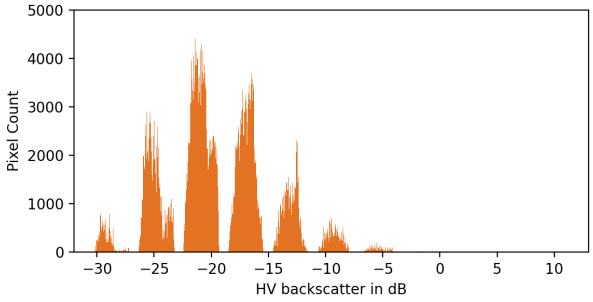


(d) March 13, 2020

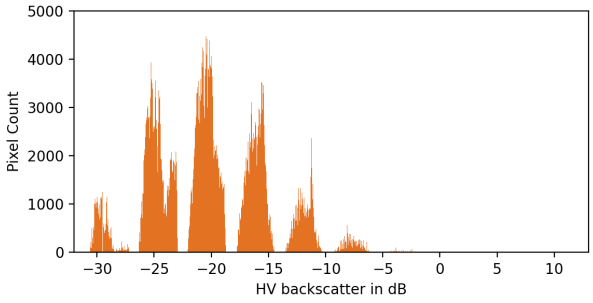
Figure 4.3: Histogram of HH backscatter for selected months in 2019



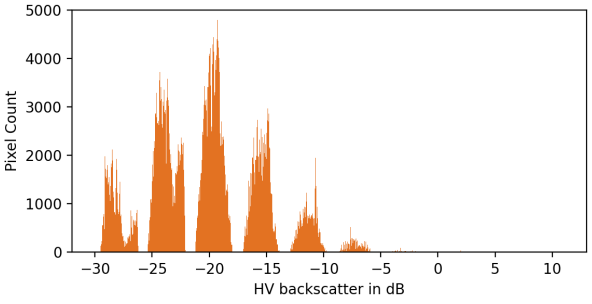
(a) June 11, 2019



(b) September 15, 2019



(c) December 20, 2019

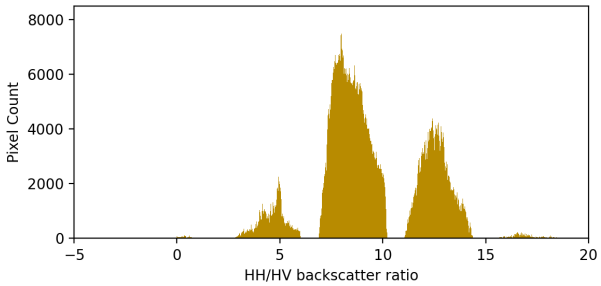


(d) March 13, 2020

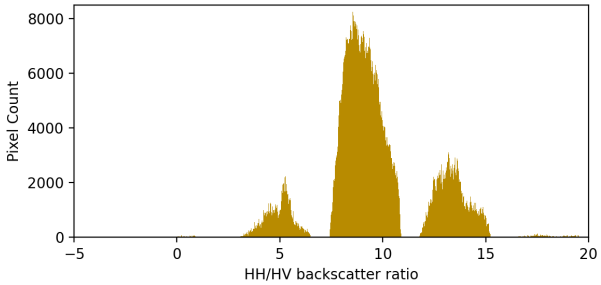
Figure 4.4: Histogram of HV backscatter for selected months in 2019

As the backscatter in both polarisations has been identified as a seasonally varying parameter, also the relation between the two backscatters was also analysed throughout the seasons.

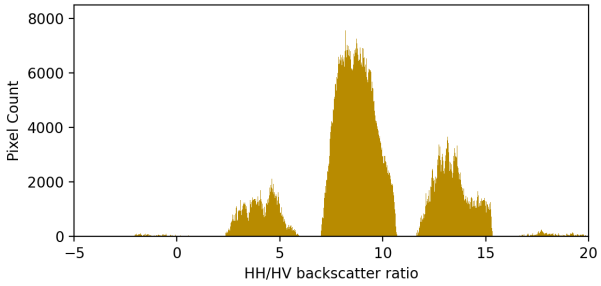
Figure 4.5 shows the pixel-wise difference between the two polarisations, also known as the polarisation ratio. For this ratio, three groups can be identified. The distribution of ratio count varies less than the quantities in the polarisations themselves throughout the seasons.



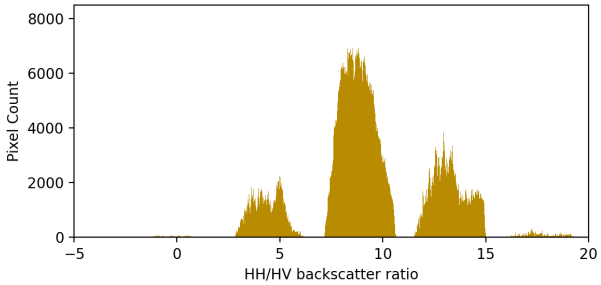
(a) June 11, 2019



(b) September 15, 2019



(c) December 20, 2019



(d) March 13, 2020

Figure 4.5: Histogram of HH-HV backscatter ratio for selected months in 2019

4.3 Seasonal variations of backscatter for glaciarised areas

After examining the backscatter patterns for all surfaces of Disko Island, the focus is now shifted towards the glaciated areas, to see if glaciers belong to one of the specific groups/clusters that have been identified in Section 4.2.

First, the backscatter development of clean ice and firn (identified by the Sentinel-2 snow probability mask) over the seasons is examined before extending the area of examination to all the glaciated areas identified by the GLIMS inventory from 2016. To look at the backscatter patterns, the mean and standard deviation of the backscatter are calculated for both polarisations and the difference between the two polarisations.

The arithmetic mean, $\bar{x}$ is calculated as follows:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{4.1}$$

with x_i being the value in the pixel and n being the total number of pixels.

The standard deviation is the measure of the extent to which values of a variable differ from the arithmetic mean, i.e., describes the spread of the values, and is calculated as

$$\sigma_0 = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \tag{4.2}$$

4.3.1 Seasonal variations of backscatter for clean ice

Looking at the backscatter mean and standard deviation from the 12-day repeat imagery in 2019–2020 (Figure 4.6), one can see that the polarisations follow the same trend, but the HH-HV ratio develops differently especially during the summer months (see Figure 4.7).

The standard deviation is about 2.5 dB to 3 dB in the HH spectrum over the seasons, whereas in HV, the standard deviation is lower, especially during the snow-free months (May to July) with a standard deviation of about 1.5 dB, and a standard deviation of about 6 dB for the winter months. However, HV varies largely in the transition periods, represented by the dates May 6 and 18, 2019 and August 22, 2019.

After the melt period in May, a clear drop in HH and HV reflectance is visible. In both polarisations, an additional drop in backscatter can be identified in mid-June. This can be related to increased precipitaion during this period. As for the summer-autumn transition, HH shows a more gradual increasing backscatter, whereas the

change of HV change is more direct, bearing in mind that HV also shows a larger variance here.

The different summer-autumn transition patterns for HH and HV are also well-reflected in variation of HH-HV ratio variation during this period (Figure 4.7). In the winter period (September to March), the backscatter is constant within the standard deviation range at about 7. For spring, summer and autumn, no clear ratio pattern could be identified with the ratio being both larger and smaller than 7 in this period.

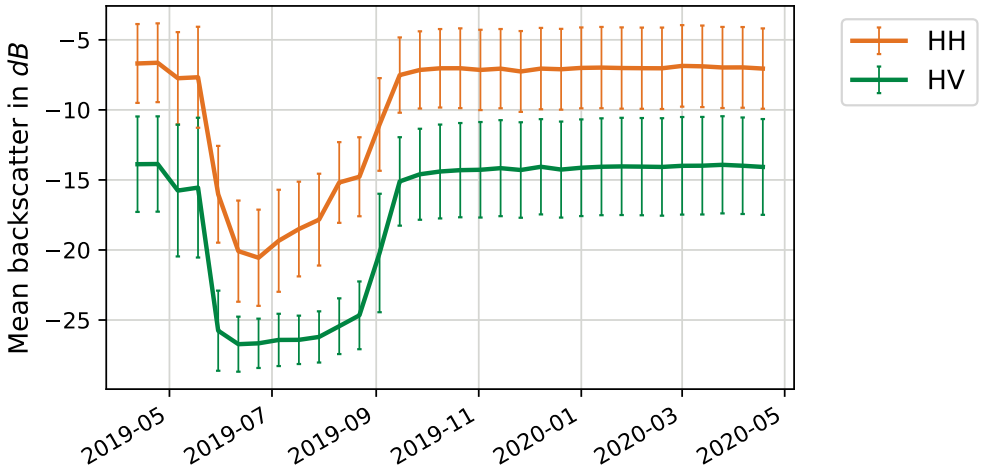


Figure 4.6: Mean and Standard deviation of backscatter for clean ice from 12-day repeat imagery, 2019–2020

Over a longer term (2016–2019, see Figure 4.8), the clean ice backscatter mean and standard deviation follow the same pattern as could be identified in the 12-day repeat imagery. In the summer month of June, there is a dip in the mean backscatter and a smaller standard deviation. Also here, the standard deviation of HH is larger than the one of HV in the summer. The backscatter through December and March is constant.

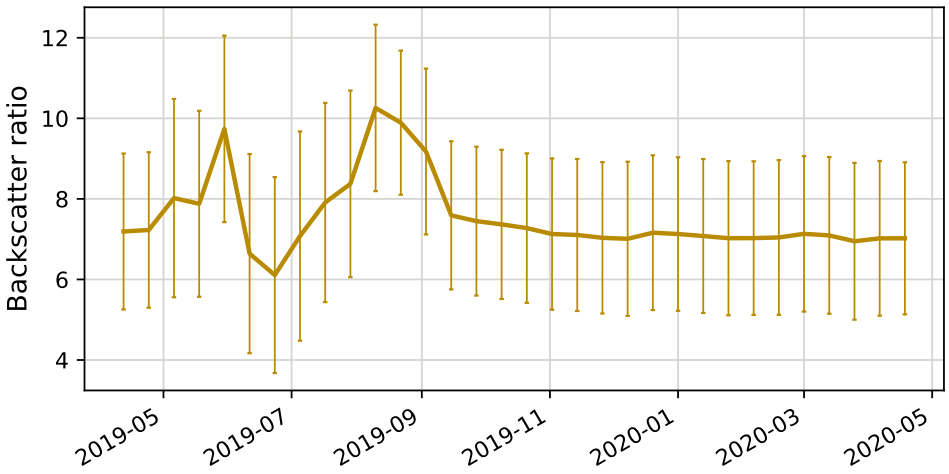


Figure 4.7: Mean backscatter ratio HH-HV for clean ice from 12-day repeat imagery, 2019–2020

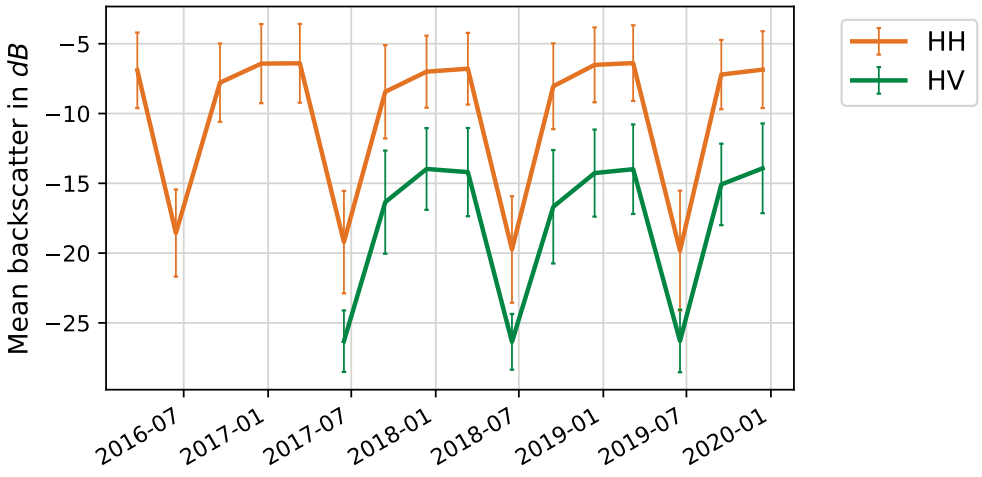


Figure 4.8: Mean and Standard deviation of backscatter for clean ice throughout the seasons, 2016–2019

4.3.2 Seasonal variations of backscatter for all glaciated areas

Looking at the backscatter development over all the glaciated surfaces (see Figure 4.9), the overall trends identified for clean ice can also be identified here: a constant mean backscatter in the winter season and more backscatter variation during melt and summer season, with the mean backscatter dropping from the beginning of the spring season.

The standard deviation for HH is at about 3.5 dB with some large outliers in the summer period (June to beginning of July) which show a large spread of backscatter values in HH polarisation.

As for HV, the standard deviation is more constant at 4 dB with two large outliers in May 2019 that show a standard deviation of more than 5 dB.

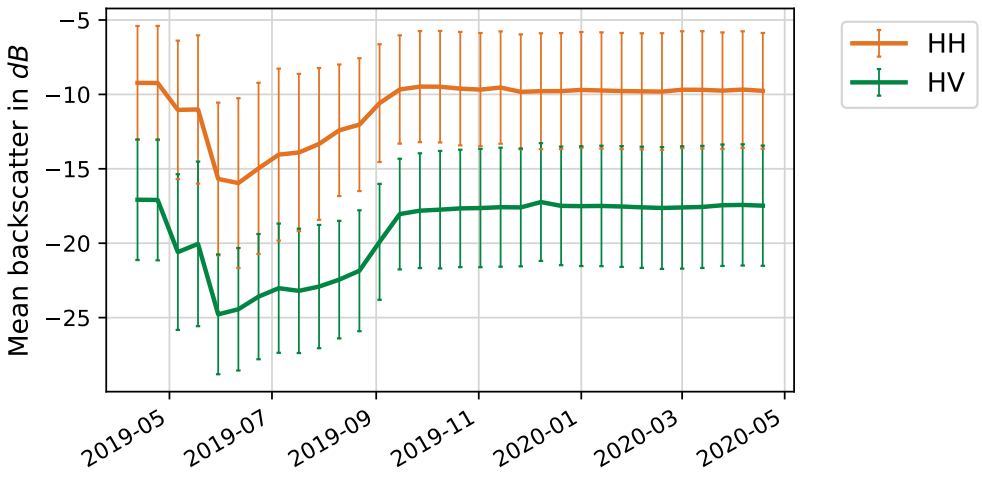


Figure 4.9: Mean and Standard deviation for all glaciarised areas of backscatter from 12-day repeat imagery, 2019–2020

Looking at the HH-HV ratio, the mean ratio is more constant throughout the course of the year than the ratio for just clean ice. However, also the standard deviations are larger. An overall trend of a constant HH-HV ratio over the snowy season is visible.

The long-term mean backscatter (Figure 4.11) follows again the same trends as for clean ice, again with the standard deviations in HH being larger in the summer.

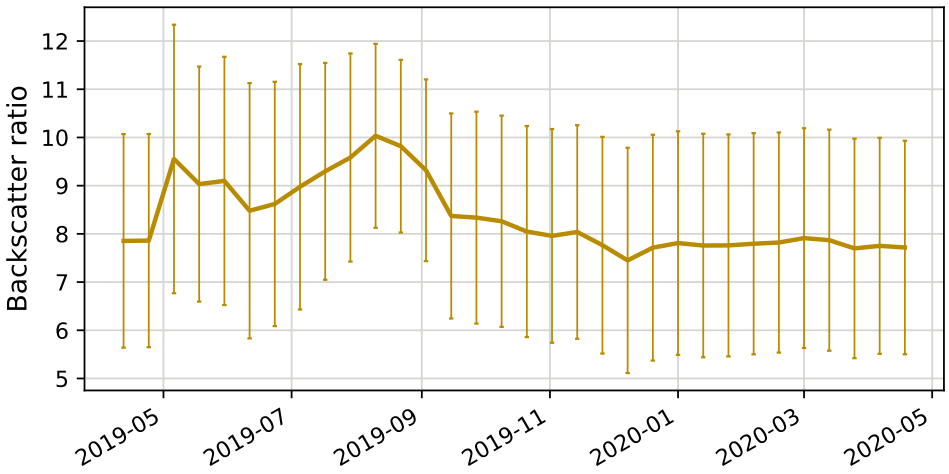


Figure 4.10: Mean backscatter ratio for all glaciarised areas HH-HV an from 12-day repeat imagery, 2019–2020

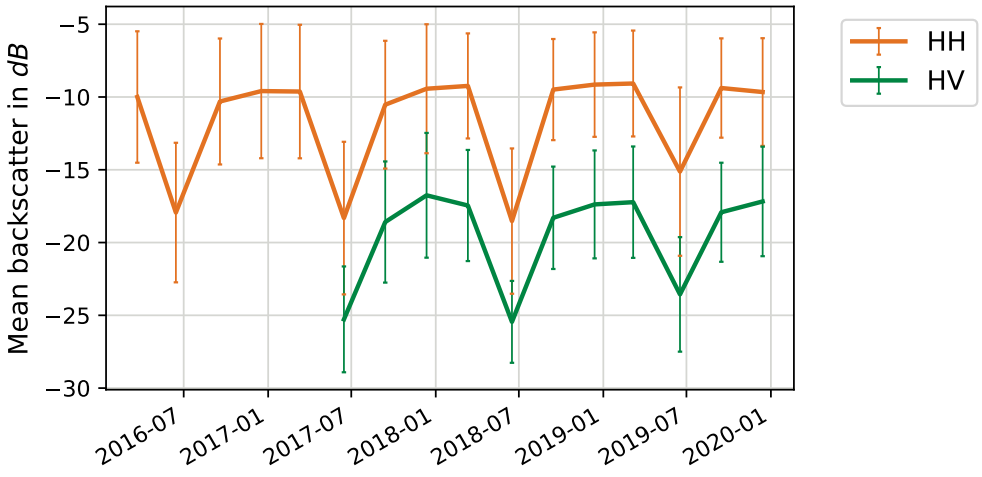


Figure 4.11: Mean and Standard deviation of backscatter throughout the seasons, 2016–2019

4.4 K-means Clustering: Backscatter patterns for surface types

To better understand the seasonal variations of the spectral signatures, they were investigated further by applying a k-means clustering with seven target clusters (six, as this is the most frequent count of bins in the histograms plus one extra cluster accounting for *No Data* values) to imagery from different seasons. k-means clustering is dividing the input data into target groups, i.e., clusters, based on their distance from the mean of its cluster – that should be minimised – and the distance from the mean of the other clusters – that should be maximised.

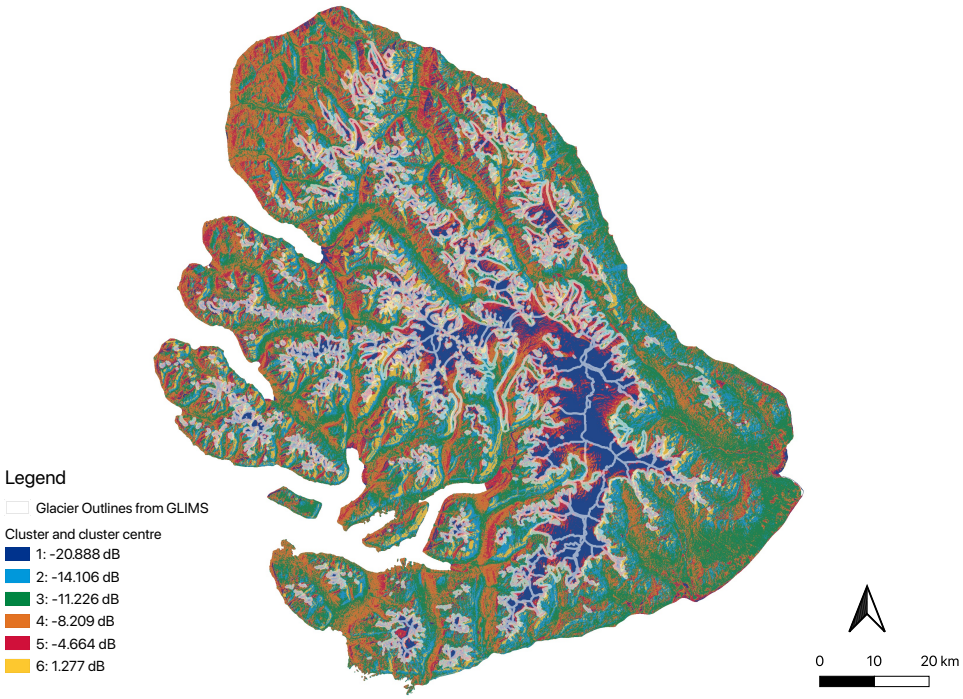


Figure 4.12: Clusters from k-means clustering for June 11, 2019 using HH polarisation

The maps in Figures 4.12 and 4.13 show the output of the k-means clustering for the backscatter, HH and HV, respectively, recorded in June 2019. It is clear that most ice is well-represented by cluster 5 in HH and cluster 3 in HV. However, there are also larger areas that are not assumed to be covered by ice that are located in those clusters.

Tables 4.1 and 4.2 describe how the different geology types are represented across the different clusters, giving more insight in backscatter similarities across the different geology types.

Table 4.1: Cluster percentage per geology type from HH clusters derived for June 11, 2019

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Volcanic rocks	3.29	18.87	26.52	25.53	13.26	12.52
Ice	43.90	4.54	8.58	15.28	26.30	1.40
Lakes	13.80	1.35	9.75	26.84	48.00	0.25
Marine sandstone	0.61	10.13	44.57	35.64	7.45	1.61
Non-marine sandstone	0.48	16.45	42.50	27.88	10.25	2.40
Rock slides	0.25	27.98	36.72	22.58	4.67	7.80
Glaciofluvial and marine deposits	3.99	1.78	38.95	40.62	14.53	0.12
Undifferentiated deposits	0.69	18.31	44.27	27.26	6.15	3.31

As for HH, 70 % of all the ice are captured in two clusters, namely cluster 1 and 5. About 44 % of the ice are captured in cluster 1 with a cluster centre at −20.888 dB, which is a little bit higher than the mean backscatter identified for clean ice at that date (Figure 4.6). However, cluster 1 also contains about 14 % of all the lakes on the island. Apart from that, there are few landscape types that largely have the same HH backscatter as 43 % of the ice in June 2019. Cluster 5 – identified by a mean of −4.66 dB – captures another 26 % of all the ice. This cluster also captures a large fraction of lakes and some fractions of marine deposits and volcanic rocks. Here, probably some of the ice with more debris cover is captured, mostly lying on the outer edges of the glaciers (Fig. 4.12).

Looking at HV (Figure 5.3), cluster 2 detects most of the ice well, but it also contains a few non-ice regions, for example in the North-West of the island, and in regions close to the sea. The cluster centre at −23.342 dB is a bit higher than the mean backscatter for clean ice at this date. 60 % of all the ice , but also (45 %) of the lakes, are captured in cluster 2 (see Table 4.2). There are also larger fractions of other geology groups, especially glaciofluvial and marine deposits and volcanic rocks that show similar backscatter amplitudes.

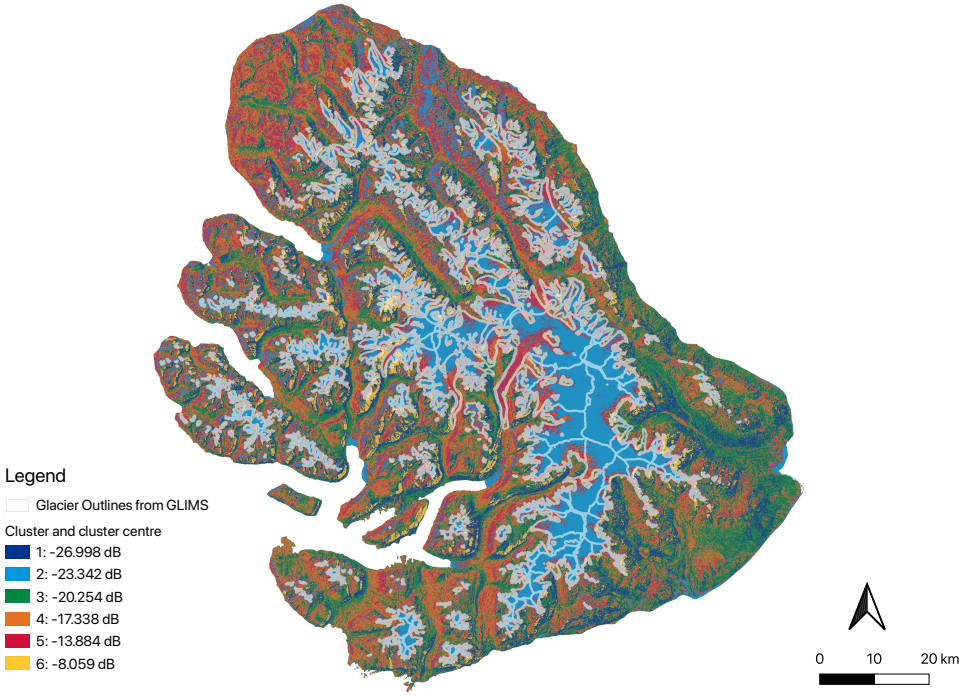


Figure 4.13: Clusters from k-means clustering for June 11, 2019 using HV polarisation with glacier outlines from GLIMS (white)

Table 4.2: Cluster percentage per geology type from HV clusters derived for June 11, 2019

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Volcanic rocks	16.94	8.58	21.91	25.74	17.99	8.86
Ice	2.86	60.67	6.31	10.66	18.56	0.94
Lakes	0.29	45.79	5.32	15.88	32.49	0.23
Marine sandstone	6.06	7.01	25.18	30.95	30.49	0.30
Non-marine sandstone	17.11	3.66	38.63	24.49	15.26	0.85
Rock slides	24.84	1.13	33.18	25.80	9.86	5.19
Glaciofluvial and marine deposits	6.72	16.29	27.31	31.30	18.24	0.14
Undifferentiated deposits	17.30	3.29	36.74	28.66	11.79	2.22

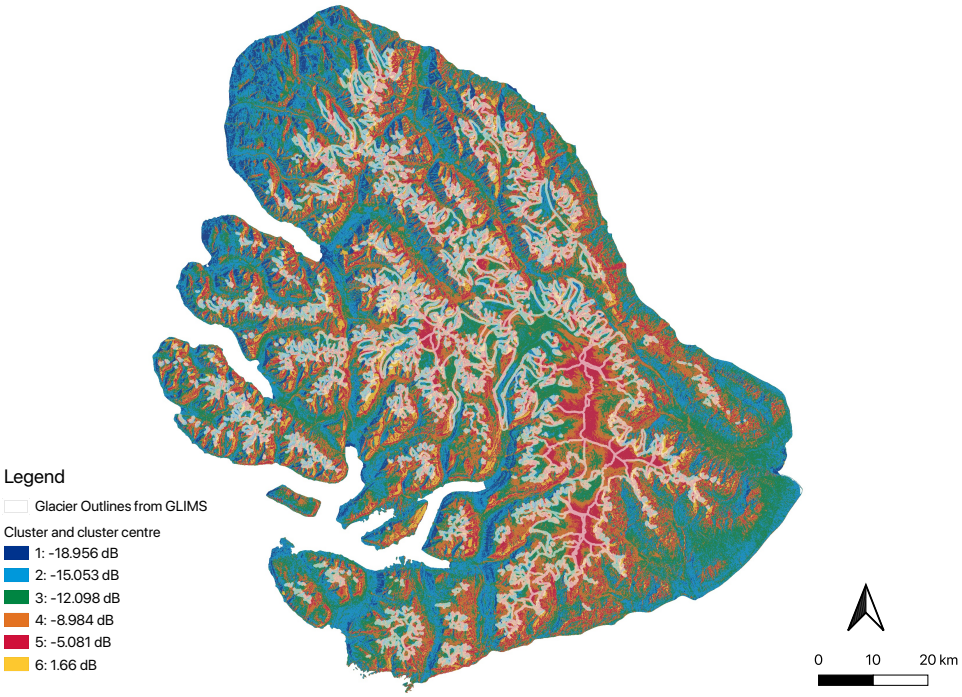


Figure 4.14: Clusters from k-means clustering for December 20, 2019 using HH polarisation with glacier outlines from GLIMS (white)

The maps in Figures 4.14 and 4.15 show the output of the k-means clustering for the backscatter, HH and HV, respectively, recorded in December 2019. In contrast to the clustering results from the June imagery, there is not obvious 'ice' cluster to be identified from the clustering performed in December.

Looking at Table 4.3, showing how the different geology types are represented in the clusters, there are three clusters covering in total about 80 % of the ice area: cluster 3 (centre: −12.098 dB), cluster 4 (centre: −8.984 dB) and cluster 5 (centre: −5.081 dB). These clusters however don't exclusively cover ice, but are also assigned to other geology types: The cluster 3 backscatter pattern is a value that can be found frequently throughout all surface types, whereas only percentages lower than 20 % of the other geology types are assigned to cluster 5.

For HV – the same results presented in Table 4.4 – a similar pattern is visible: about 80 % of the ice is contained in cluster 1 (centre: −26.843 dB)), cluster 2 (centre: −23.893 dB)) and cluster 4 (centre: −17.965 dB). The same spread of different backscatter throughout the geology types that could be observed in HH is also present in HV.

Comparing the backscatter clusters of both months, one can see that the backs-

Table 4.3: Cluster percentage per geology type from HH clusters derived for December 15, 2019

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Volcanic rocks	11.03	18.94	21.80	21.02	16.13	11.09
Ice	1.77	11.61	30.62	32.99	20.95	2.06
Lakes	9.02	56.32	23.50	10.00	0.90	0.25
Marine sandstone	3.06	21.73	46.79	22.52	4.41	1.49
Non-marine sandstone	4.30	26.54	34.51	21.33	10.20	3.13
Rock slides	3.26	17.04	25.96	27.45	19.00	7.28
Glaciofluvial and marine deposits	11.73	49.95	31.54	6.23	0.46	0.10
Undifferentiated deposits	4.01	24.22	33.08	24.44	11.45	2.80

catter is shifting on most surfaces from summer to winter, but that these shifts are not necessarily related to a certain geology type.

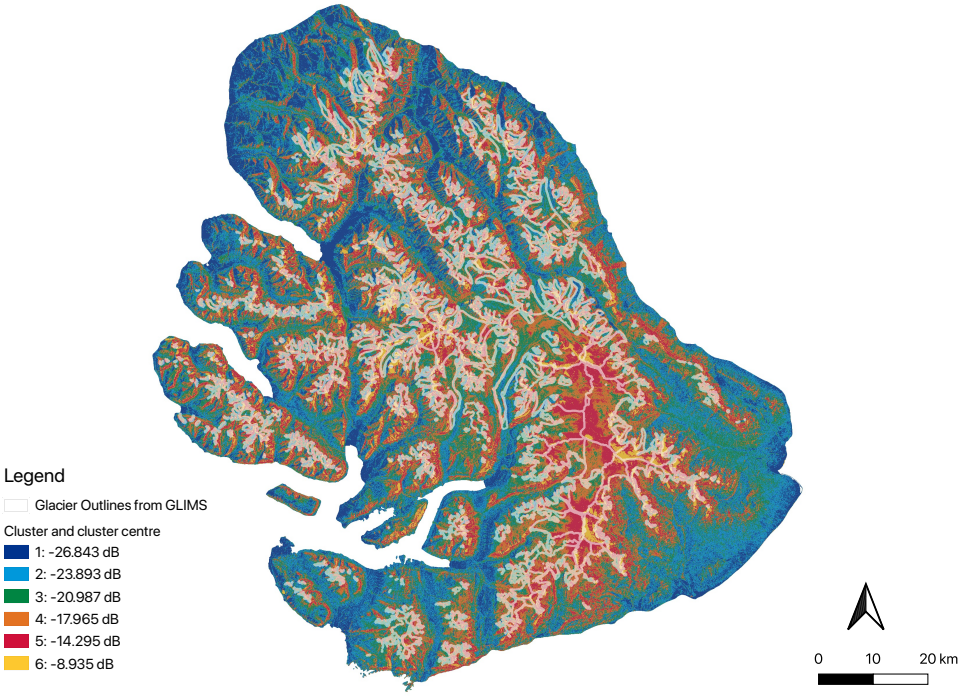


Figure 4.15: Clusters from k-means clustering for December 20, 2019 using HV polarisation with glacier outlines from GLIMS (white)

Table 4.4: Cluster percentage per geology type from HV clusters derived for December 15, 2019

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Volcanic rocks	22.49	20.15	19.12	15.33	15.25	7.66
Ice	23.99	33.43	7.09	26.40	1.25	7.85
Lakes	10.16	5.45	21.80	0.21	62.15	0.23
Marine sandstone	16.39	3.61	42.01	1.23	36.58	0.17
Non-marine sandstone	31.10	10.67	41.48	1.82	14.81	0.11
Rock slides	25.33	27.68	19.38	18.44	5.14	4.03
Glaciofluvial and marine deposits	10.48	2.08	32.28	0.30	54.80	0.07
Undifferentiated deposits	25.21	20.35	28.55	9.63	14.69	1.57

4.5 Decision Tree

Clustering gave some idea of how to identify glacial ice from the backscatter values of the two polarisations. However, one could also see that a more nuanced method is required to improve the detection of glaciers / glacial ice as there was a high presence of other geology types in the dominant clusters for ice.

To find out if there were any additional patterns that have not been discovered so far, also combining HH and HV backscatter for the glacier detection, the decision tree supervised learning method was tested. The *Gini index* was used here as a measure of impurity and the sample data was assigned weights to account for the different quantities of ice-covered and ice-free areas (described in more detail in Section 3.3.4).

In the first step, the tree was trained to match the backscatter values to a binary target mask represented by the GLIMS outlines from 2016 (three target groups: *No Data*, *No Ice*, *Ice*). Here, the backscatter from the different seasons, represented by the imagery from the selected months in the period 2017 - was used separately and the resulting trees were used to detect the glaciated areas in the following years, i.e. 2018 and 2019.

To measure the decision tree's performance, they were cross-validated by splitting the input data in a training and test set with the test set consisting of one third of the total number of input values. This way, the tree's performance can be measured both against its input data and overfitting can be detected before applying the derived decision rules to possibly more different datasets.

The trained decision trees were then used to predict the different target groups (i.e., *Ice*, *No Ice*, *No Data*) for imagery recorded on different dates and in different years and examined towards the types of geology that get identified as false positive for ice.

For readability, the root of the decision tree is moved by one level down as the first branch only identifies *No Data* pixels.

4.5.1 Training with binary dataset

June (representing summer)

Decision Tree Characteristics For the June imagery, the decision tree diverges at depth 3 (see Fig. 4.16), and gives a cross-validation score of 86 %.

In this case, the decisions are solely made on the HV polarisation:
Applying the threshold $-31.405 \leq \text{HV} \leq -23.55$ dB, 80 % of all ice pixels are identi-

fied correctly with large confidence, but also 16 % of all non-ice areas are identified as ice, thus predicting a much larger ice cover than it is the case. 16 % of all ice pixels get identified as false negative already after applying the first threshold ($HV \leq -23.55$). Looking back at the reflection mean and standard deviation for HV (Figures 4.8 and 4.11), this threshold corresponds to the HV backscatter range for clean ice (-24 dB to -29 dB). As the possibly debris-covered areas have a reflection range from -22.5 to -29 dB, this decision tree captures most of the clean ice but identifies "dirty" ice as false negative.

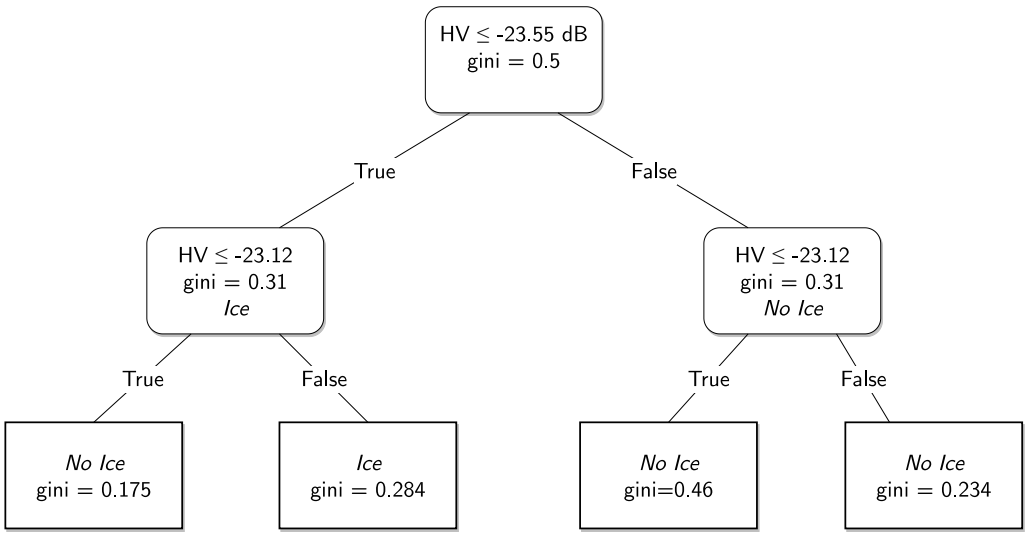


Figure 4.16: Decision tree graph trained to June 15, 2017

Prediction maps Looking at the predictions for the following years in the maps 4.17 through 4.19, there is a large number of false positives in the predictions for June 2017 and 2018, especially in June 2018. The prediction for June 2019 shows fewer false positives and in general a smaller ice cover.

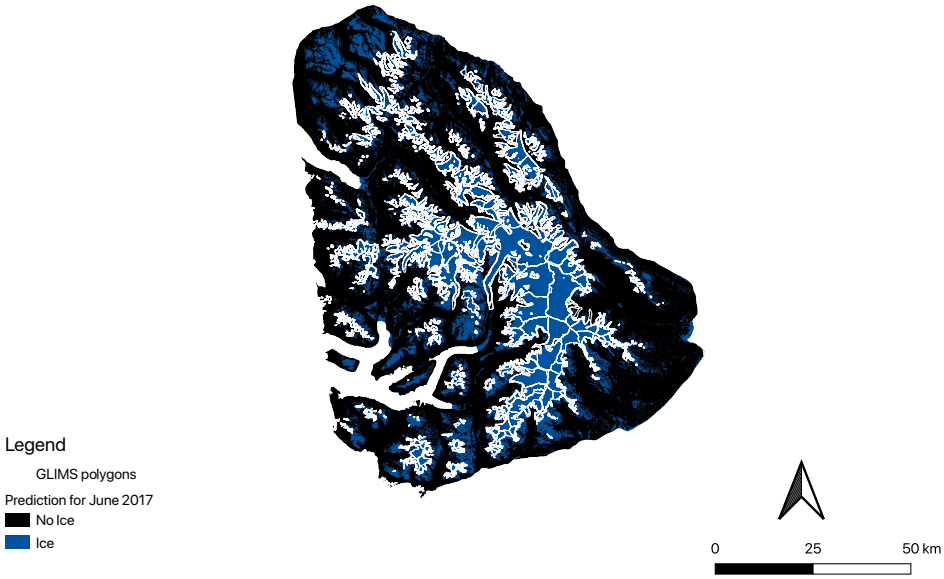


Figure 4.17: Decision tree prediction for June 2017

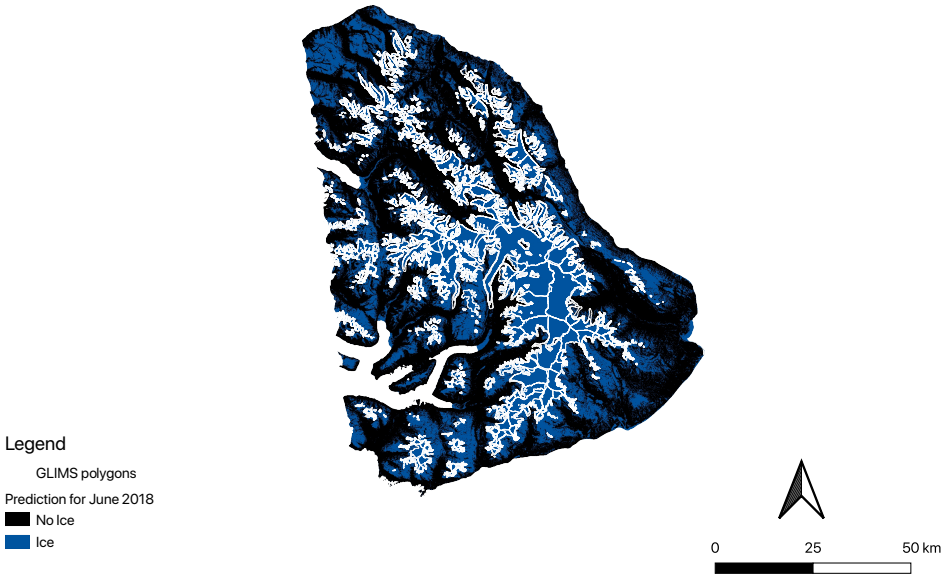


Figure 4.18: Decision tree prediction for June 2018

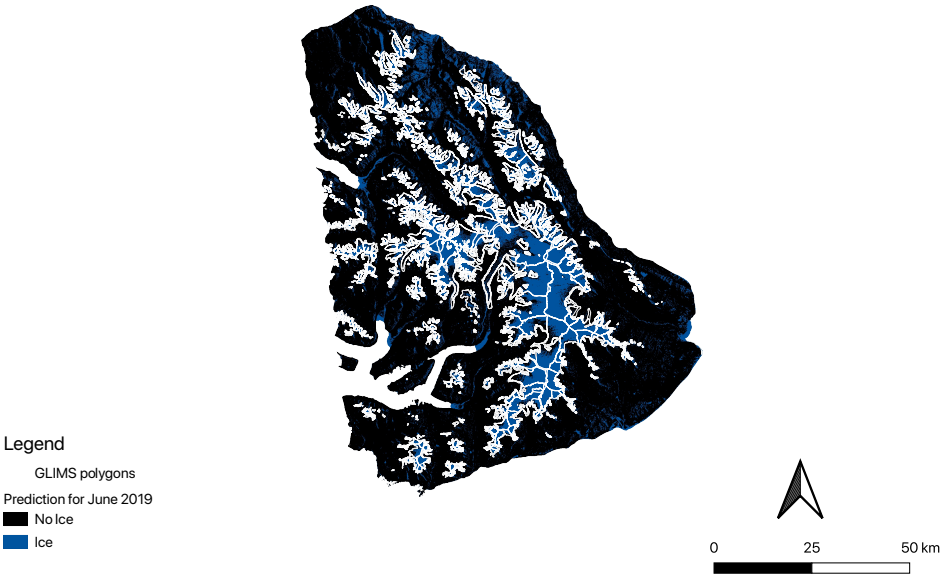


Figure 4.19: Decision tree prediction for June 2019

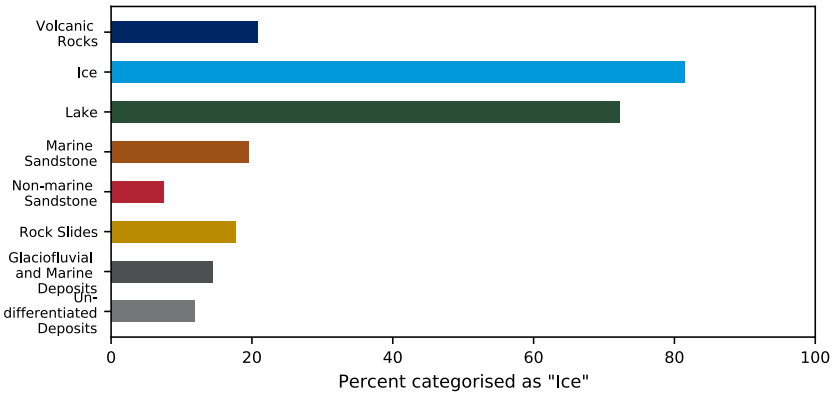
Geology Types Figure 4.20 shows the percentage of each geology type that has been predicted as ice for the years 2017, 2018 and 2019, using the decision tree described above. Figure 4.21 describes how much of the area detected as *Ice* is actually ice and the fractions of false positives per geology type.

In June 2017 (Fig. 4.20a), 80 % of the ice gets detected correctly. However, also 70 % of the lakes get detected as ice. Marine sandstones, volcanic rocks and rockslides show 20 % of false positive predictions. This may have been caused by the location of these surface features being close to glaciers.

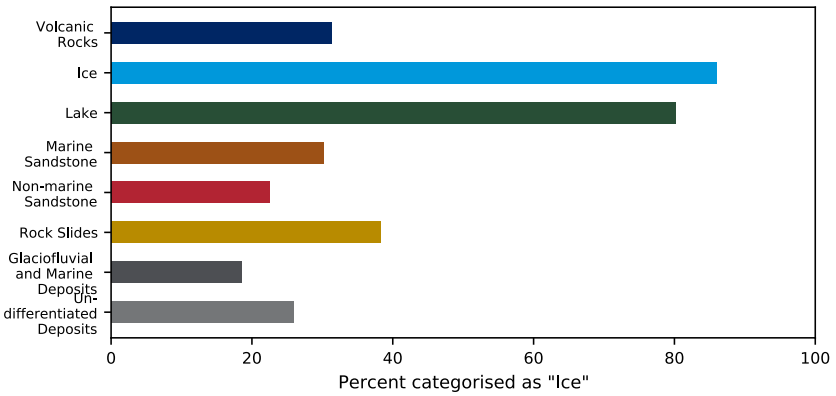
Looking at the same results for June 2018 (Fig. 4.20b), 84 % of the ice gets predicted correctly, again probably covering most of the clean ice. 80 % of the lakes do again get detected as ice. In general, there is a larger number of false positives in the prediction for that year, which was also visible in the map (Fig. 4.18). This can be related to the larger number of precipitation events in June 2018 compared to 2017 and 2019.

For June 2019 (Fig. 4.20c), the ice detection is not as successful as in the previous two years with only about 60 % of the ice being identified. This may be related to melt and actual shrinkage of glaciers as the validation geology map has not been updated for this particular year. Also here, more than 60 % of the lakes are again false positives, as well as marine sandstone (about 40 %). As for other groups, there are much fewer false positives compared to the previous years (below 20 %).

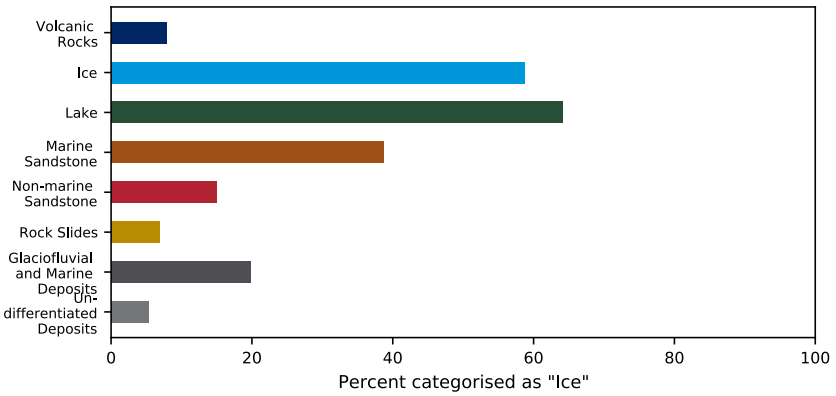
Although lakes were largely detected as *Ice*, they are not the main contributor of false positives for any of the years (see Fig. 4.21). The percentage of correctly detected ice varies from 45 % to 70 %. The three surface types with largest surface cover are the main contributors to false positives. The contribution of volcanic rocks to the false positives varies largely from June 2017 and 2018 to 2019, with in general a smaller fraction of false positives in June 2019.



(a) June 2017

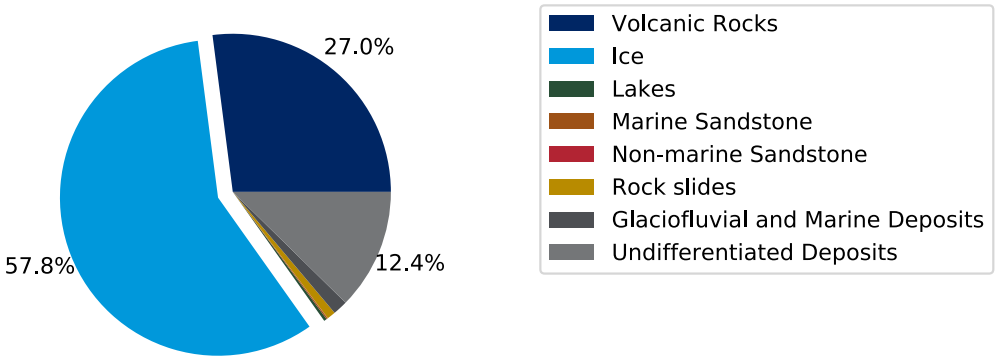


(b) June 2018

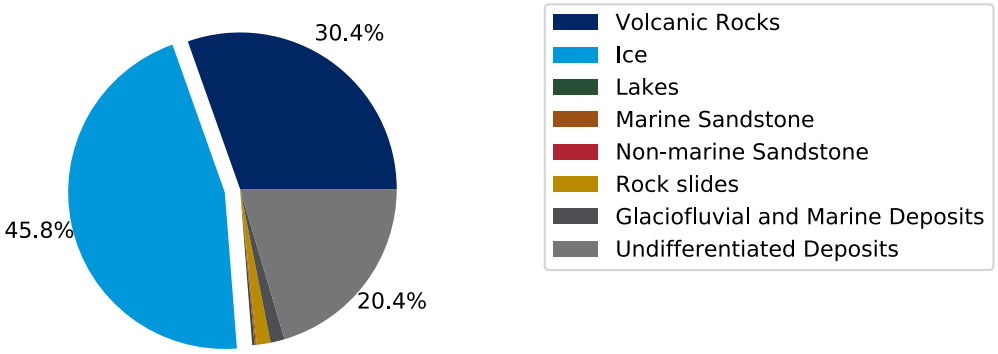


(c) June 2019

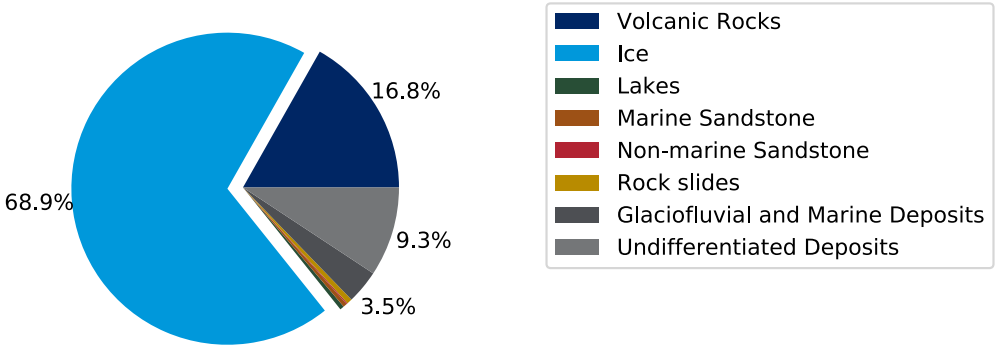
Figure 4.20: Percentage classified as ice per geology type with decision tree from June 2017



(a) June 2017



(b) June 2018



(c) June 2019

Figure 4.21: Geology types classified as ice with decision tree from June 2017

September (representing autumn)

Decision Tree Characteristics As the tree predicted in June, also the September decision tree (4.22) is diverging at depth 3, however showing larger impurity values. The cross-validation score was 83.7%, reflecting a slightly worse performance than the tree fitted and tested in June. In contrast to the previous tree, the decisions are based on both HH and HV polarisation: The first threshold $HH \leq -4.7dB$ is already ruling out a large amount of no ice areas at depth 2. After this point, a lot of pixels would be identified correctly as ice but there is also a large number of false positives. The second threshold applied is $HV > -16.3dB$ which gives good confidence in predicting ice positives and capturing 21 % of all ice pixels correctly. However, this second threshold also removes a large amount (69 %) of ice pixels and classifies them as non-ice pixels.

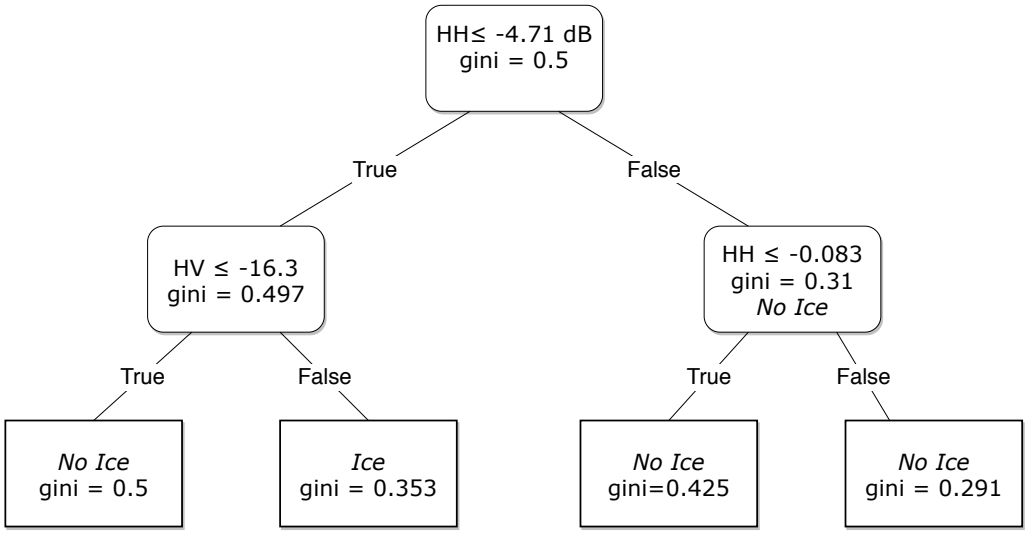


Figure 4.22: Decision tree graph trained to September 13, 2017

Prediction maps The prediction maps for September in the years 2017 to 2019 are very similar. As previously stated, this decision tree seems to apply thresholds that rule out most backscatter and leave just a fraction of the actual ice cover identified as ice. This does however also lead to very few false positives.

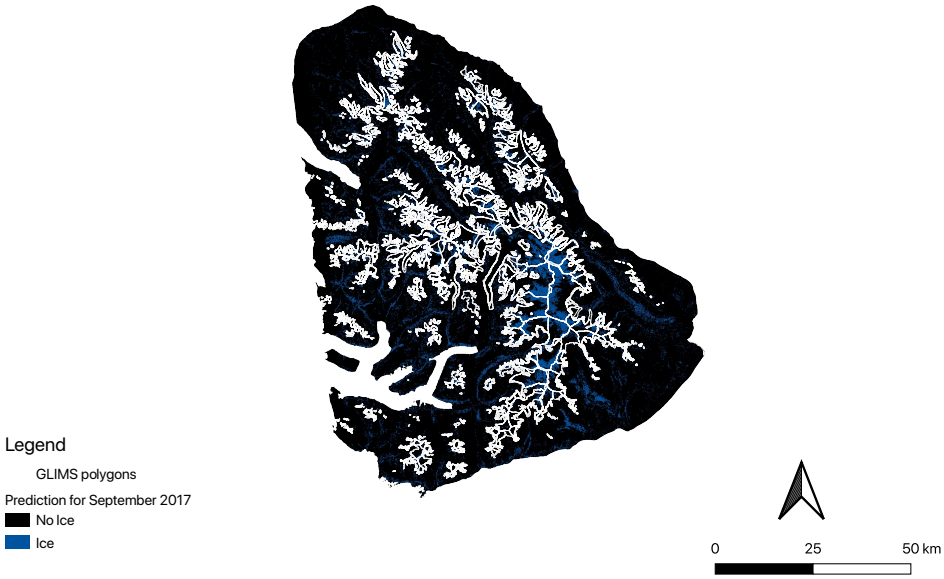


Figure 4.23: Decision tree prediction for September 2017

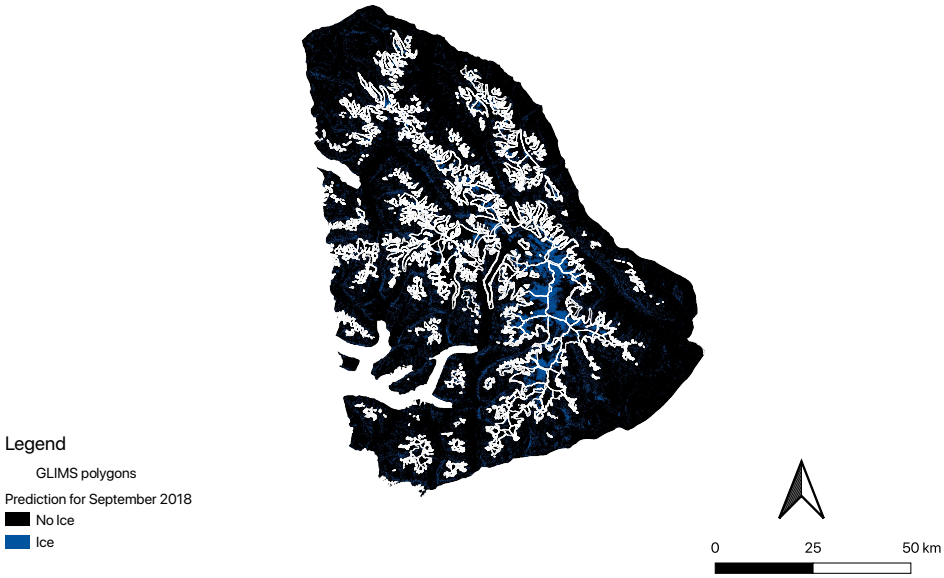


Figure 4.24: Decision tree prediction for September 2018

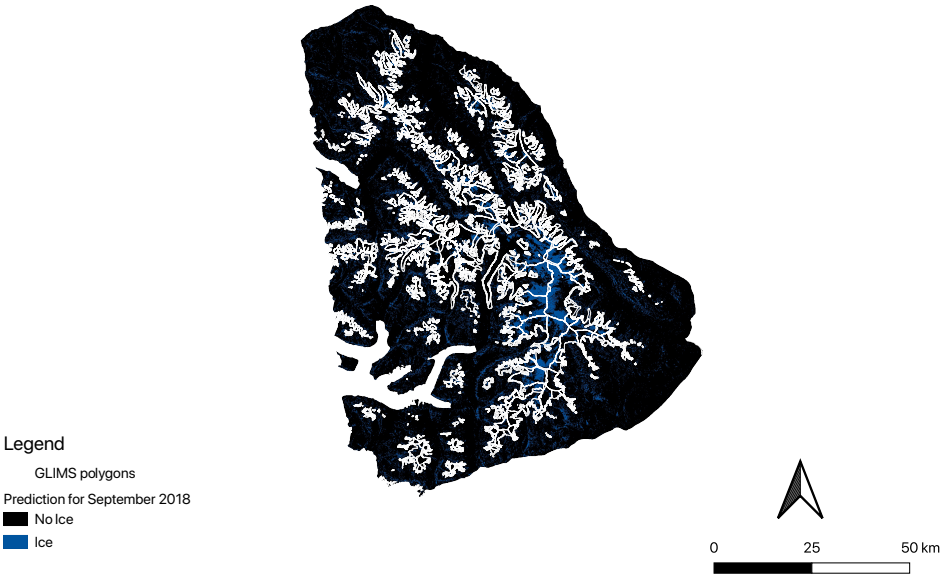
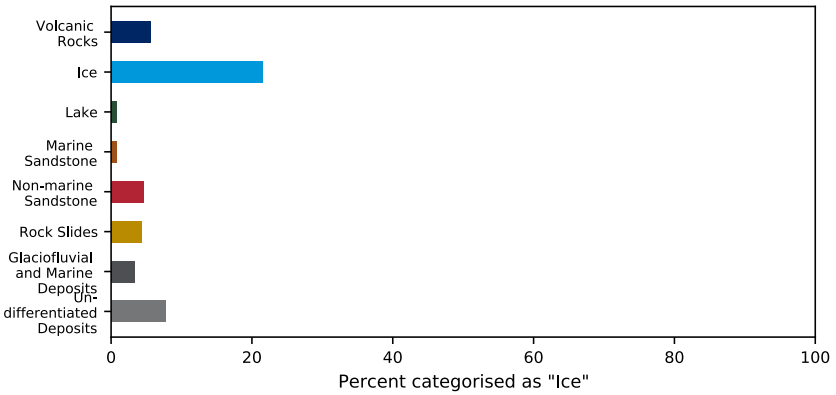


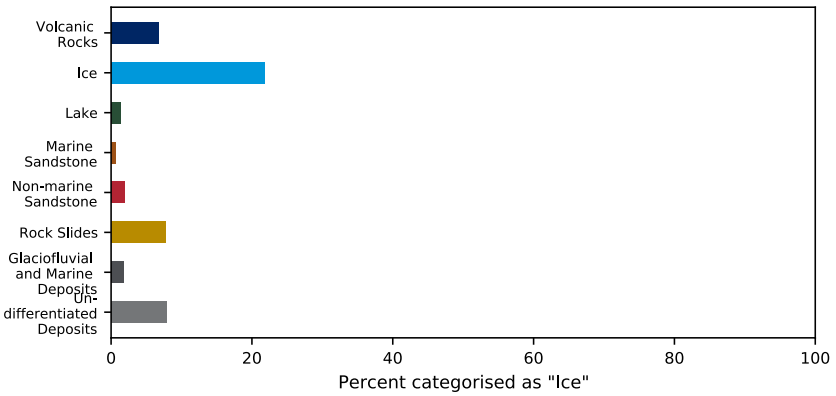
Figure 4.25: Decision tree prediction for September 2019

Geology Types In the prediction for September 2017 about 21 % of the ice is recognised correctly. In contrast to June, the lake signal is not giving false positives. In general, the false positives from the other geology groups are below 10 % This may however be due to the more restrictive rules that this tree contains. This pattern continues throughout the following years; in general, the reflection of the different geology types does not seem to change throughout the years.

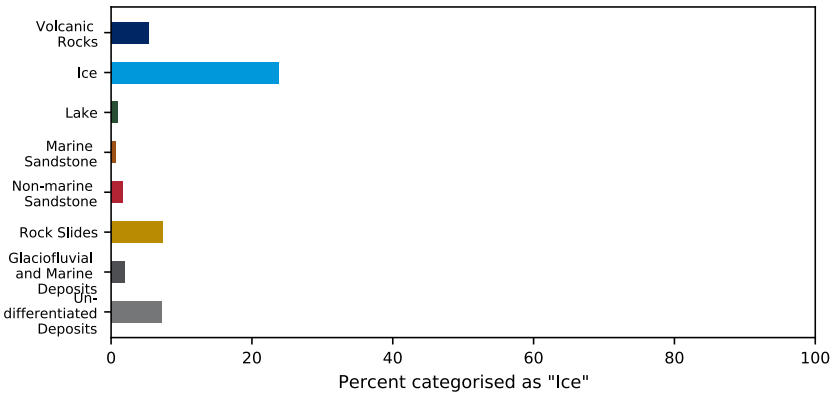
This trend compared to the June imagery is also visible in Figure 4.27: In total, about 50 % of the area detected as *Ice* is actually ice. As for the other surface types, the geology types with the largest surface cover also contribute the most to the false positives.



(a) September 2017

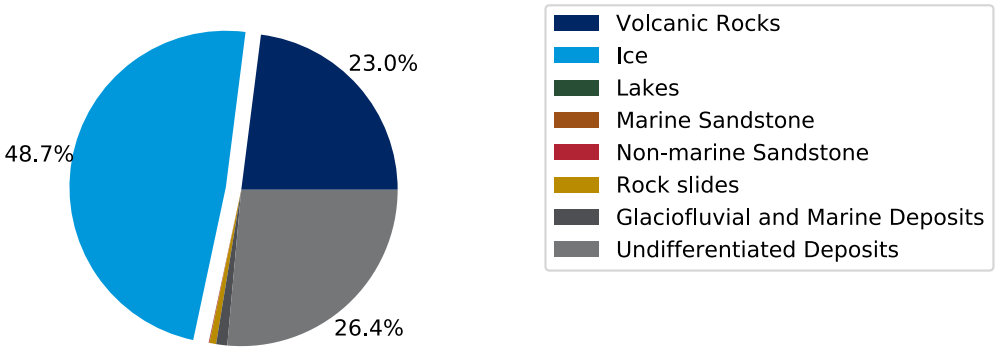


(b) September 2018

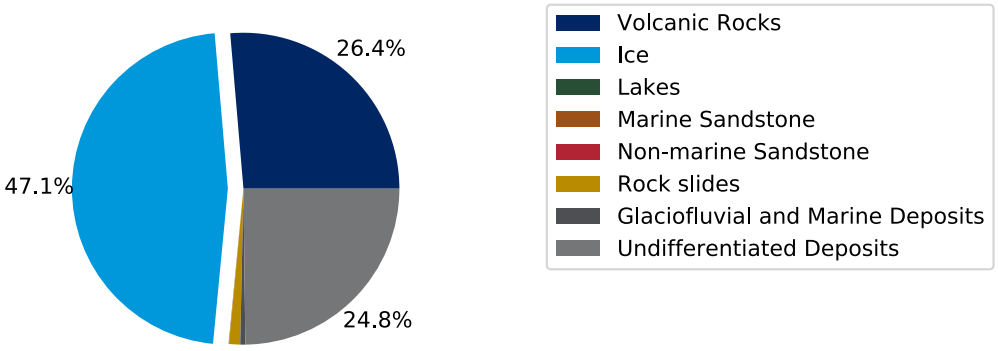


(c) September 2019

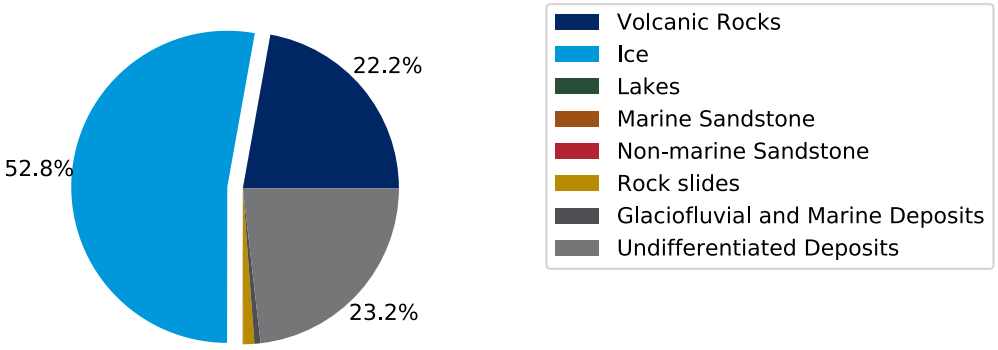
Figure 4.26: Percentage classified as ice per geology type with decision tree from September 2017



(a) September 2017



(b) September 2018



(c) September 2019

Figure 4.27: Geology types classified as ice with decision tree from September 2017

December (representing winter)

Decision tree characteristics A decision tree of depth 3 fitted to imagery from December 12, 2017 (see Fig. 4.28) gives a cross-validation score of 74.5 %. The decisions are based on values from both polarisations. Looking at the impurity values in the leaves, one can see that the tree is not converging and that there is low confidence for all decision paths.

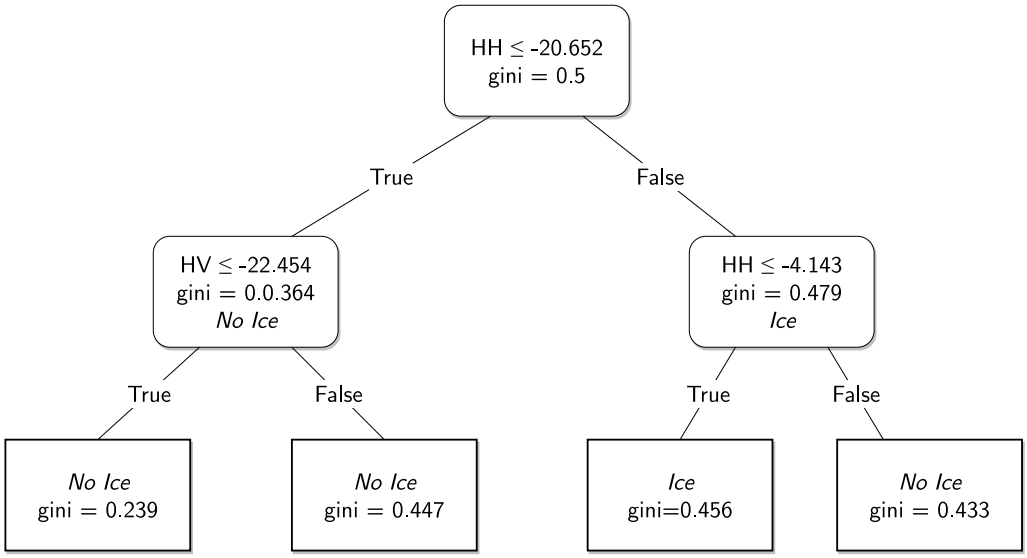


Figure 4.28: Decision tree graph for December

Decision tree including DEM At first sight, including a DEM improves the decision tree's performance by 10 % with a cross-validation score of 89.3 %. However, looking at the paths leading towards the identification of ice, there is a large bias towards high elevation locations. All glaciers below 800 m are not recognised. Above this elevation threshold, the tree is still struggling to confidently identify ice/no ice areas.

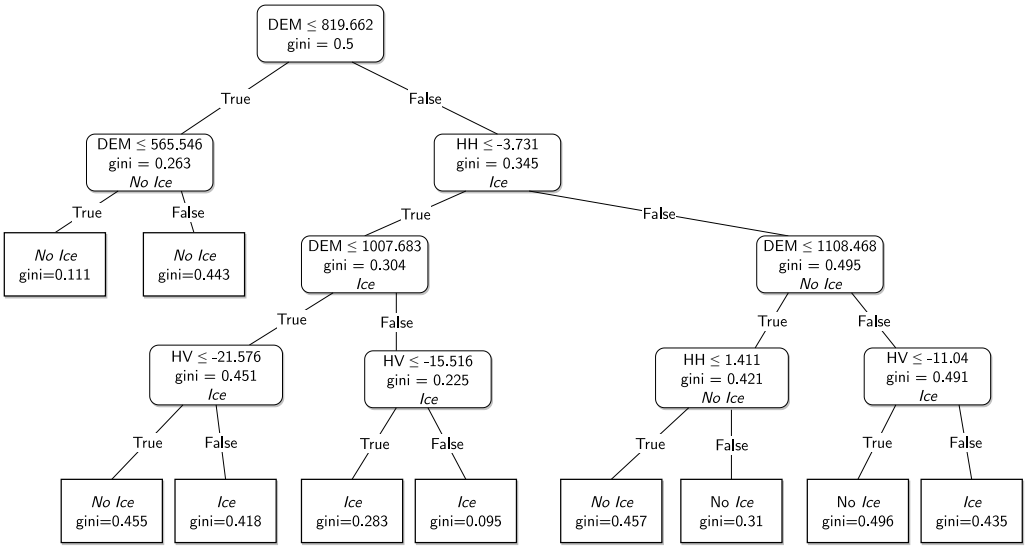


Figure 4.29: Decision tree graph for December including DEM

4.5.2 Training with geology types

In another experiment, the decision tree was trained to detect geology types for imagery from October 9, 2019, assigning specific weights to the training dataset for geology groups.

A decision tree of depth 3 that has been fitted with the *balanced* option can be seen in Figure 4.30. What is remarkable here is the tree's ability to predict the lake class with large confidence already at depth 2 based solely on HH values (following two thresholds: all HH < -20.481 dB with little impurity (0.164), and $-20.481 \leq$ HH -18.194 dB with larger impurity). The ice class cannot be confidently predicted; given an impurity of 0.818, pixels predicted as *Ice* can correspond to any geology class.

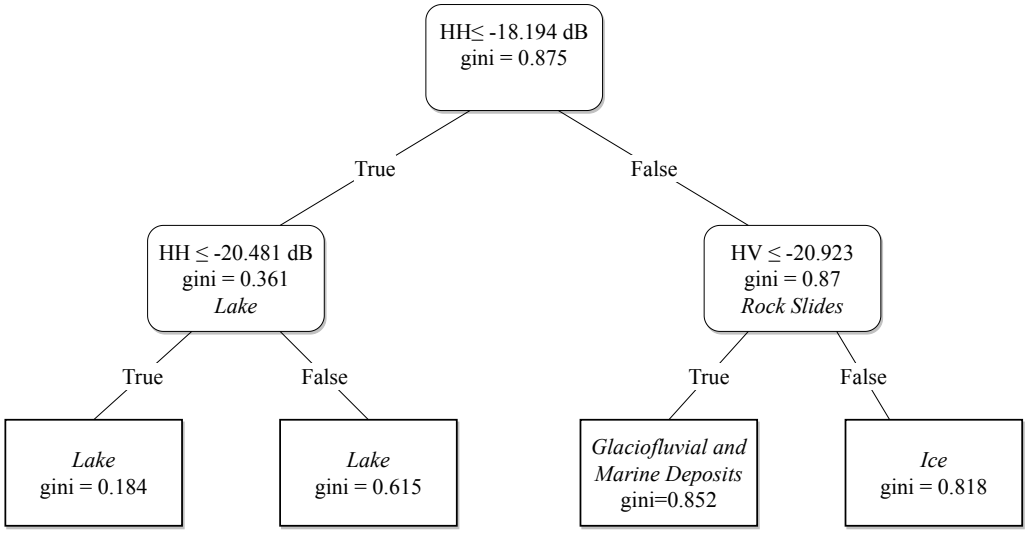


Figure 4.30: Balanced decision tree graph trained for October using *balanced* option

Chapter 5

Discussion

Recalling the underlying aim of this study, backscatter properties from glaciers on Disko Island were examined and the potential use of SAR imagery for glacier detection in this region was evaluated.

Given that SAR is weather- and illumination-independent, a positive result would yield potential for glacier monitoring both in regions with frequent cloud cover, or even allow for glacier monitoring ability throughout all seasons, as the SAR signal is also penetrating through loose snow. A potential use case here could be the constant monitoring of the dynamics of surge-type glaciers.

Having analysed the SAR backscatter distribution and its seasonal variation, one unsupervised (k-means clustering) and one supervised (decision tree classification) learning method have been applied to SAR imagery from different seasons.

The output of these methods, i.e. the identified classes and – for the decision tree – the decision paths, have been examined towards geology types that interfere with the detection of glacial ice, leading to false positives.

The following discussion will put the findings of this study in context with the objectives presented in Chapter 1.4.1 and compare them with work that has been done in other geographical regions. Potential opportunities for improvement of the presented methods will be given, along with an evaluation of the overall objective.

5.1 Detecting glaciated areas from SAR imagery using unsupervised and supervised learning methods

There have been several studies working on the recognition of glacier zones (e.g., Kundu and Chakraborty, 2015, Winsvold et al., 2018 and Barzycka et al., 2019).

This study explored a new approach, looking at the potential of using SAR imagery to detect glaciated areas themselves.

The analysis of the backscatter for the entire region of Disko Island identified clearly distinguishable ranges of backscatter values. Applying a k-means clustering to imagery from different seasons showed that these ranges did not prove to be assigned to specific surface geology types. Further, the spatial distribution of the backscatter amplitude ranges varied throughout the seasons.

For ice, there were two to three ranges of backscatter that captured the majority of glaciated and ice areas. This aligns well with the study of glaciers in Svalbard performed by Winsvold et al., 2018, where it was shown that that Sentinel-1 C-band SAR backscatter showed two distinct backscatter ranges for firn and ice, −6 dB to 0 dB and −18 dB to −13 dB, respectively. This can be explained by the fact that during the warm season with firn transitioning to wet snow and melting surface ice, more energy is absorbed.

However, especially in the imagery recorded during the winter, other geology types had similar backscatter amplitudes to that demonstrated by ice; this was particularly evident for lake areas. In winter, this may be due to the lake being frozen and thus essentially acting as an ice area. In summer, it may be related to the fact that the glacier surfaces in the lower altitudes have melted and thus absorbed more of the incoming signal, as it was also shown in Winsvold et al., 2018 study. More thorough analysis, including ground validation, would be needed to understand the reason for the backscatter similarities.

The results from the k-means clustering showed that the task of glacier detection from SAR imagery was not going to be solved based on one single threshold. Testing a decision tree approach offered the chance to include more input data in the classification algorithm and capture groups in different backscatter ranges in an attempt to detect ice reliably and confidently. Consequently, a decision tree method was applied to the same imagery (including both polarisations) as a means to establish decision paths for ice detection.

First, a decision tree was trained to distinguish between *Ice* and *No Ice* areas. For the June imagery, the suggested decision path was based solely on the co-polarised HV backscatter, correctly predicting up to 60 % to 80 % of all the ice throughout the years. However, 16 % of non-ice areas was also included in this group, a relatively large number of false positives. As already seen in the k-means results, a large fraction of the lakes showed a similar backscatter pattern during this season. Apart from the lakes, marine sandstone was often false-positive classified (20 % to 40 %). It remains

to be checked if this path was chosen due to debris cover on some of the glaciers and what type of debris this would be. For September, decision rules that would correctly capture more than 20 % of the ice could not be established. This is due to the temperatures being just around 0 °C, causing melt and refreeze during this period.

Comparing the different months, the most stable time to establish a glacier detection algorithm was June. Here, the backscatter patterns were most constant for most geology types and thus there was a similar number of false positives year-over-year. Finding different ways to detect these geology types and eliminate the stable false positives can improve the presented backscatter-based decision tree.

A balanced decision tree towards geological groups reaffirmed what was observed in the k-means clustering: there are no clear backscatter patterns that can be assigned to a certain geology group. The only geology type that was properly recognised was the “lake” group as it was represented strongly in exactly one cluster. It therefore became evident that training the algorithm towards the binary GLIMS mask was more successful than capturing all geological groups.

In summary, the two threshold-based learning methods presented and tested in this study do not yet offer a reliable glacier detection method. Clean and dry ice can easily be detected, but debris cover and surface melt impose difficulties and can lead to a larger number of false positives or stricter rules for the decision tree, limiting its ability to detect certain types of ice.

5.2 Surface types interfering with the recognition of glacial ice from SAR imagery

In June 2017, 70 % of the lakes were false-positive interpreted as ice. This may be related to melt on top of the glaciers causing the presence of supraglacial lakes that confuse the decision tree. This theory can be backed by the multispectral-derived snow probabilities for the false-positive detected pixels from the decision trees. These lakes would likely not be predicted as ice areas from the multi-spectral imagery. As lakes don’t cover a large area of Disko Island, using the multi-spectral imagery to rule these out would only give a small improvement of the total area estimation. In general, the decision tree in June seems to recognise clean ice well, also over a longer term (2 years).

K-means clustering revealed that the shifts happening in the backscatter are not necessarily linked to a certain geology group or ground type; the change in backs-

catter values is in fact more varied than that. This also applies to the glaciated areas, corresponding to the earlier-identified shift through the seasons; whereas most of the ice is identified by the lowest HH backscatter cluster in June, the same areas are covered by cluster 4 and 5 in December, corresponding to a shift in backscatter by about 10 dB. It should also be noted that for the non-ice areas, the backscatter groups in the South did not change as much as they did in the North.

In total, clean ice was well-recognised in the June imagery (particularly using HV polarisation) though the k-means clustering indicated that a large fraction of the lake areas seemed to have the same backscatter pattern as ice. As for the December imagery, there were some backscatter groups that were dominant in the glacier areas, but the values/backscatter groups were much more scattered than in June.

The backscatter variations over the entire region are also a result of backscatter from vegetation and other features. Even though there are some dominant groups that likely represent the "ice"/glacier group, there are also non-ice areas that get captured as members of those clusters, as well as some ice areas being captured in different clusters. The spread of the different backscatter clusters over the entire area proves that the simple k-means clustering is not suitable to reach the goal of detecting glaciers from radar data. However, given that there are two to three backscatter groups that are mostly found in glaciated areas, backscatter could still be proven to help distinguish different types of ice in glaciated areas.

5.3 SAR reflection patterns for clean ice and debris-covered glaciers

A pattern for the backscatter of glacier areas throughout the seasons was identified with a standard deviation of about 3 dB for clean ice and about 4.5 dB for all glacier areas in both polarisations.

As Kundu and Chakraborty, 2015 also observed on glaciers in the Central Himalaya region, the seasonal backscatter variation is linked to changes in snow grain size, water content in the firn column, and the depth in both the accumulation and ablation area. Valenti et al., 2008 identified the same trend looking at co-polarised (VV) SAR imagery of wet snow from ENVISAT/ASAR during different seasons: melt causes a drop in backscatter whereas temperatures closer to the freezing point lead to a rise of backscatter over the glaciated areas.

Figure 5.1 shows this trend: as soon as the temperatures rise above 0 °C, the backscatter value lowers for both polarisations as the snow is getting wet and the

grains are larger. Approaching 0 °C again in October, the backscatter values return to their earlier levels, due to both refreezing of the wet snow cover in the uppermost layer (creating ice bodies) and accumulation of fresh snow.

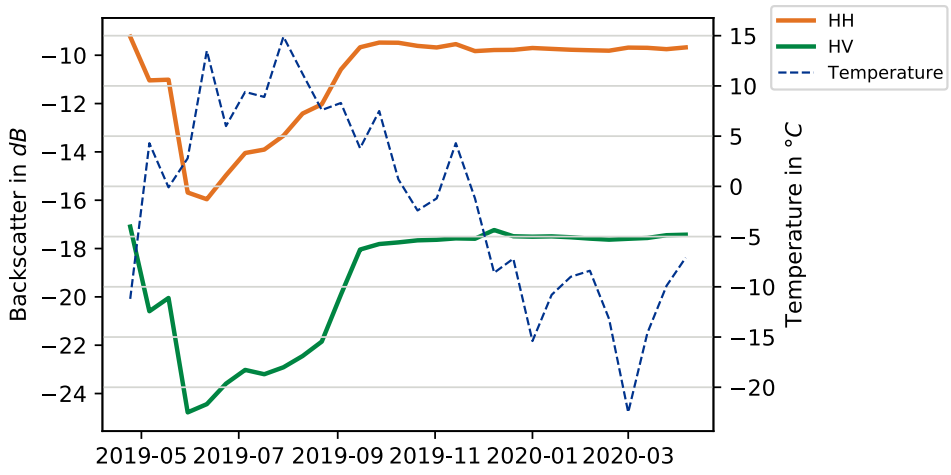


Figure 5.1: Mean backscatter variation for HH and HV polarisations of all glaciated areas (left scale) and mean temperature (right scale)

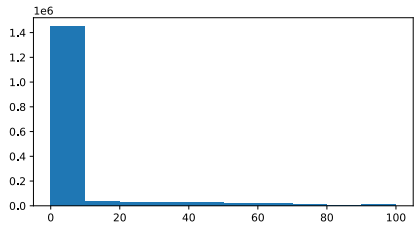
By analysing backscatter recorded with Sentinel-1 for different glacier zones, Winsvold et al., 2018 recorded glacier backscatter values on Svalbard between −18 dB to 0 dB in the period October (2015) to April (2016); the lower range −18 dB to −13 dB described the ice zone while backscatter values closer to zero described firn. In summer, the firn layer – partly melted into wet snow – and the uppermost layer of some of the ice melted, lowering the backscatter to −20 dB to −18 dB. This change in backscatter is thus less extreme than what was observed in this study.

Long and Drinkwater, 1994 examined the backscatter of the Greenland Ice Sheet using SEASAT imagery at different locations. In a region of the Greenland Ice Sheet at a similar latitude to Disko Island – i.e., 70°N – the backscatter range observed in July was −12 dB to −8 dB for VV co-polarisation. In general, they found a seasonal backscatter variation of up to 10 dB, larger than the mean variation that was observed in this study for the glaciers of Disko Island.

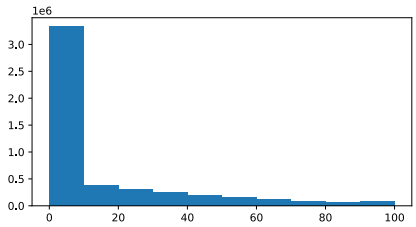
5.4 How well are the methods performing compared to glacier detection from multi-spectral imagery?

k-means To compare the k-means clusters to the snow mask from Sentinel-2, the results of the k-means algorithm for glaciers derived from GLIMS and the respective snow probability were connected for the month of June which consistently gave the best results in this study. Here, the clusters that captured most of the "ice" area were examined.

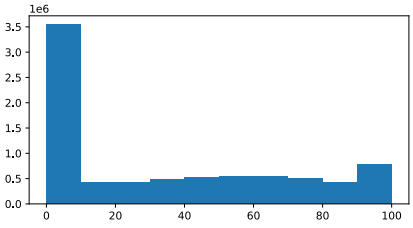
For HH polarisation, clusters 1, 4 and 5 captured 85 % of the ice. The snow probabilities corresponding to the pixels in these clusters are depicted in Figure 5.2. What is remarkable here is that in cluster 1 (44 % of the detected ice) very few pixels are recognised to have a snow probability larger than zero. In effect, these would not have been detected as ice using the snow probability product derived from the multi-spectral image. Cluster 5 (26 % of the ice) contains a much larger amount of pixels that could confidently be predicted as ice.



(a) Cluster 1



(b) Cluster 4

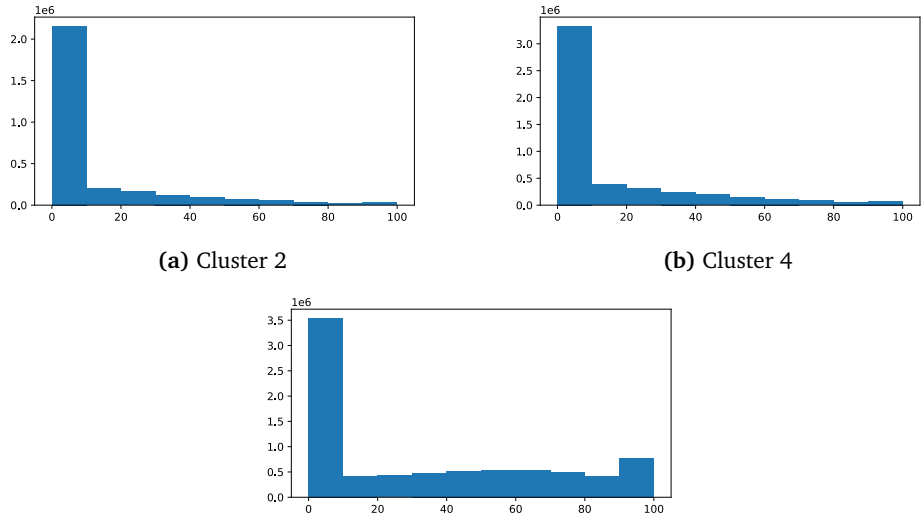


(c) Cluster 5

Figure 5.2: Sentinel-2 snow probability for glaciated areas in selected clusters from HH polarisation

Looking at clusters 2, 4 and 5 in the HV polarisation cluster image for the same date (see Figure 5.3), a similar pattern can be observed, with cluster 5 containing the most pixels with a larger snow probability and cluster 2 – the cluster containing 60 % of the ice – not containing many pixels with a snow probability larger than 50 %. Compared to cluster 1 in HH, however, cluster 2 in HV shows a larger amount of pixels

with higher snow probabilities, again emphasizing the result that HV polarisation is more suitable for predicting ice from SAR imagery.



(a) Cluster 2

(b) Cluster 4

(c) Cluster 5

Figure 5.3: Sentinel-2 snow probability for glaciated areas in selected clusters from HV polarisation

Figure 5.4 shows the clustering result overlaid on an RGB-composite image from Sentinel-2. Clean ice and snow/firn is detected well, including smaller ice areas among other geology types. For some regions, both on ice and non-ice areas, the backscatter captures more variation than the optical Sentinel-2 image, whereas other clean-looking areas on Sentinel-2 are not recognised as ice in the radar imagery.

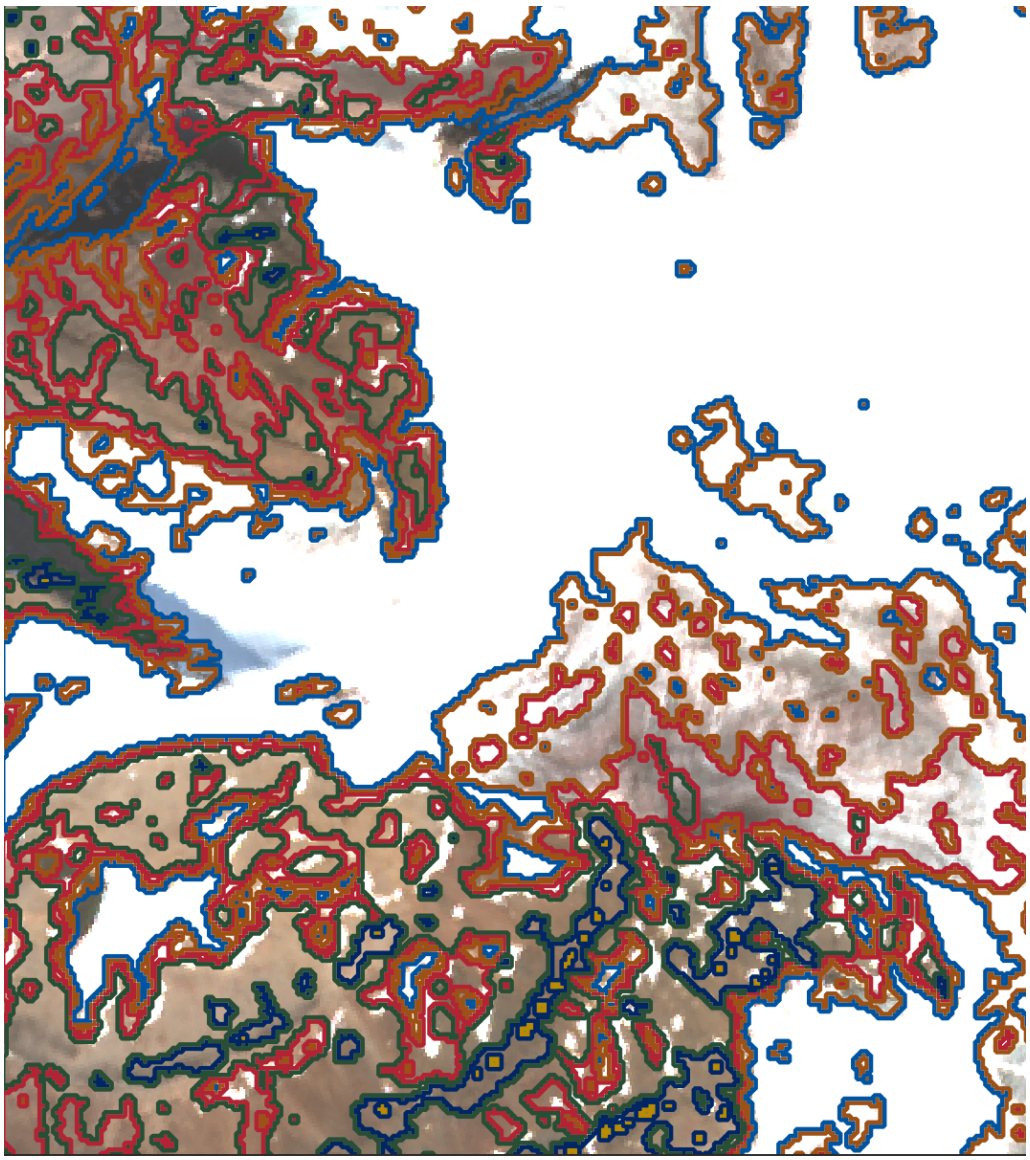
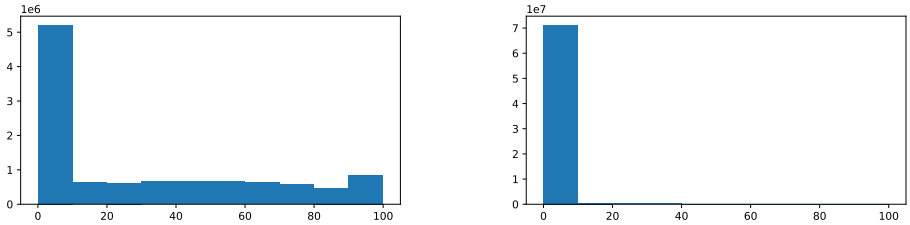


Figure 5.4: Cluster outlines derived from Sentinel-1 HV imagery on June 11, 2019 with RGB composite image from Sentinel-2 captured on June 12, 2019 (*Copernicus Open Access Hub*)

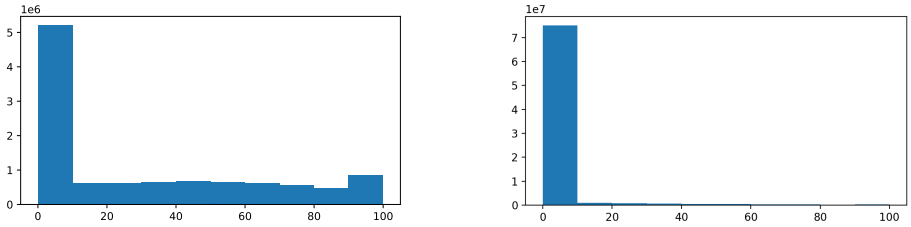
Decision Tree Comparing the decision tree output to the snow probability from Sentinel-2, one can see for the results in June 2019 (Figure 5.5) that the multi-spectral image and the decision tree strongly agree on a large number of pixels they false-positively detect as *No Ice*. As for the *Ice* group, the result is more nuanced with snow probabilities across the entire range. As in the k-means analysis, there is a large number of *Ice* pixels that would not have been detected as ice from the multispectral-derived snow mask.

Figure 5.6 for September shows a similar pattern as the results from June: Areas that were detected as *No Ice* (i.e., false negatives) in the SAR imagery were also not detected as snow or ice in the multi-spectral imagery.



(a) Ice **(b)** No Ice

Figure 5.5: Snow probability for glaciated areas in ice and non-ice class from decision tree for June 2019



(a) Ice **(b)** No Ice

Figure 5.6: Sentinel-2 snow probability for glaciated areas in ice and non-ice class from decision tree for September 2019

Figure 5.7, an overlay of decision tree results on the Sentinel-2 RGB composite image, shows a similar pattern to the earlier overlay of k-means clustering. A smaller area of ice is classified as ice than what is visible in the Sentinel-2 imagery. Smaller areas of snow and firn are also captured here, although the performance seems to be worse than the one shown by k-means. Examining both overlay images, it is evident that minor areas of false positive ice detection could be ruled out by using a simple filter assigning a minimum size to a potential glaciated area.

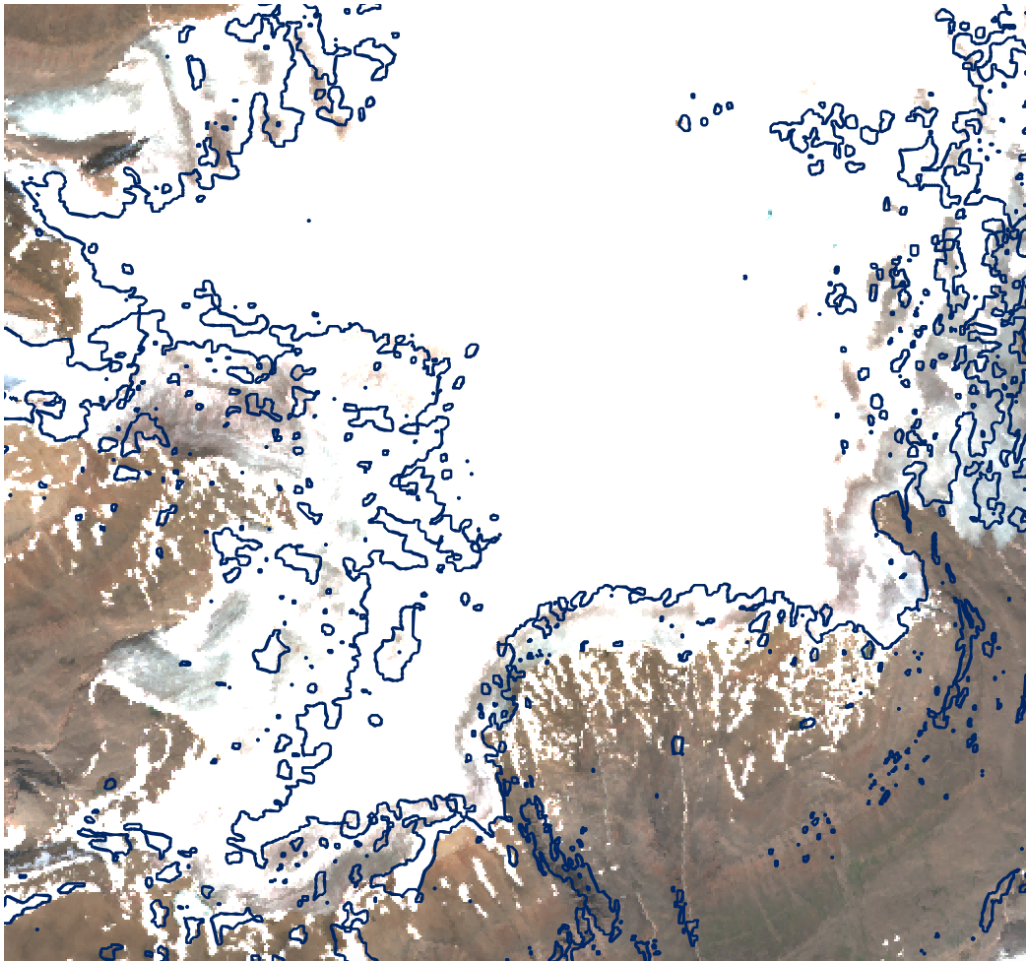


Figure 5.7: Glacier outlines derived from Sentinel-1 imagery by decision tree on June 11, 2019 compared to RGB composite image from Sentinel-2 captured on June 12, 2019 (*Copernicus Open Access Hub*)

5.5 Is SAR imagery effective for recognition of glaciers among different surface types?

The methods tested in this study are not nuanced enough to confidently predict all ice areas. Some different ice areas can however be detected. One main challenge identified is the ruling out of false positives. To serve this purpose, more knowledge of the specific glaciers is needed to identify whether certain debris cover can confuse an unsupervised learning algorithm.

5.5.1 Possible improvements

Multi-spectral imagery provides an opportunity to rule out some of the areas of SAR imagery generating false positives. For example, the already-presented snow mask could serve as a primer with a reasonable buffer to also detect non-clean ice, such as kilometre-long debris-covered glacier tongues. Previously detected glacier outlines could also serve as a primer for the present detection of glaciers; however, care must be taken such that a bias masking potentially critical changes is not introduced. Finally, the recorded backscatter can be combined with elevation data to assess in a more nuanced way which subregions are identified.

As the decision tree already gives promising results, a more complex supervised classification method could improve the outcome of the task of glacier detection from SAR backscatter. For example, *Random Forest* clustering, potentially including additional data sources such as data from several months and seasons, interferometry, etc., can provide an opportunity to better capture backscatter characteristics. Indeed, interferometry has been used in similar analyses (see Section 2.2.2) and has proven useful to assist glacier delineation and track glacier movement. It would be most suitable to apply such a complex method to a smaller region first to obtain results that can then be easily normalised and interpreted.

Additionally – as mentioned in Section 2.2 – different machine learning algorithms incorporating more information have been successfully applied to map glacier zones. It would be interesting to find out if they could also serve to detect glaciers themselves.

Chapter 6

Conclusion

Reliably detecting glacial ice automatically and throughout the seasons remains an interesting and complex task, both in the presented region of Disko Island and in other regions with a large number of debris-covered glaciers. Past studies have shown SAR imagery to be a remote sensing data source with the ability to capture debris-covered glaciers to a certain extent and to detect other glacier-related properties.

This study aimed to assess the potential of SAR imagery for the recognition of glaciers among different surface types by looking at one unsupervised and one supervised learning method to detect glaciated areas. The clusters resulting from k-means clustering, the unsupervised method, have been analysed for surface types that could be assigned to a given cluster. Certain backscatter ranges (i.e., clusters) were identified as characteristic for glaciers. However, other surface types show backscatter reflection patterns similar to those characteristic for glaciers; this collision combined with a seasonal variation of backscatter reflection patterns to reduce the effectiveness of the unsupervised method.

The decision tree method, which included SAR imagery with both HH and HV polarisation, enabled a more nuanced means of glacier detection than the k-means clustering based on a single polarisation. SAR imagery from the summer months, in this case June, has proven to be best suited for this supervised method; the decision tree captures more than 70 % of the ice while keeping the fraction of false positives mostly below 50 %.

Remaining to be explored are possible combinations of imagery from different seasons to rule out certain areas that are identified as false positive by the currently examined methods. The integration of the polarisation ratio as an input to the decision tree or clustering algorithm can also be considered.

Bibliography

Alaska Satellite Facility. *ASF Data Search Vertex*. URL: https://search.asf.alaska.edu/.

ArcticDEM Explorer. URL: https://livingatlas2.arcgis.com/arcticdemexplorer/.

Asiaq Greenland Survey Nuuk, Greenland. *ClimateBasis Data*. Greenland Ecosystem Monitoring programme. URL: gem.dk.

Bamber, Jonathan L. and Andres Rivera (2007). 'A review of remote sensing methods for glacier mass balance determination'. In: *Global and Planetary Change*. ISSN: 09218181. DOI: 10.1016/j.gloplacha.2006.11.031.

Barzycka, Barbara, Małgorzata Błaszczyk, Mariusz Grabiec and Jacek Jania (2019). 'Glacier facies of Vestfonna (Svalbard) based on SAR images and GPR measurements'. In: *Remote Sensing of Environment* 221.November 2018, pp. 373–385. ISSN: 00344257. DOI: 10.1016/j.rse.2018.11.020.

Cogley, Graham et al. (2015). *GLIMS Glacier Database*. National Snow and Ice Data Center. URL: http://dx.doi.org/10.7265/N5V98602.

Copernicus Research User Support. URL: http://rus-copernicus.eu.

Danish Meteorological Institute (DMI). *DMI Vejrarkiv*. retrieved June 12, 2020. URL: https://www.dmi.dk/vejrarkiv/.

De Zan, Francesco and Andrea Monti Guarnieri (2006). 'TOPSAR: Terrain observation by progressive scans'. In: *IEEE Transactions on Geoscience and Remote Sensing*. ISSN: 01962892. DOI: 10.1109/TGRS.2006.873853.

European Commission (2020). *About Copernicus*. [retrieved May 18, 2020]. URL: https://www.copernicus.eu/en/about-copernicus.

European Space Agency. *Level-2A Algorithm Overview*. URL: https://earth.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-2a/algorithm.

European Space Agency. *Copernicus Open Access Hub*. URL: https://scihub.copernicus.eu/.

European Space Agency. *Sentinel-1 SAR Technical Guide*. URL: https://sentinel.esa.int/web/sentinel/technical-guides.

European Space Agency. *Sentinel-1 SAR Technical Guide: Glossary*. URL: https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-1-sar/appendices/glossary.

Fahnestock, Mark et al. (2016). 'Rapid large-area mapping of ice flow using Landsat 8'. In: *Remote Sensing of Environment* 185, pp. 84–94. ISSN: 00344257. DOI: 10.1016/j.rse.2015.11.023. URL: http://dx.doi.org/10.1016/j.rse.2015.11.023.

Fang, L., O. Maksymiuk, M. Schmitt and U. Stilla (2013). 'Determination of glacier surface area using spaceborne SAR imagery'. In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-1/W1, pp. 105–110. ISSN: 1682-1777. DOI: 10.5194/isprsarchives-XL-1-W1-105-2013. URL: http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XL-1-W1/105/2013/.

Fletcher, Karen (2007). *InSAR Principles: Guidelines for SAR Interferometry Processing and Interpretation*. Tech. rep.

Gascon, Ferran et al. (2017). 'Copernicus Sentinel-2A calibration and products validation status'. In: *Remote Sensing*. ISSN: 20724292. DOI: 10.3390/rs9060584.

Geoogical Survey of Denmark and Greenland (2014). *Geological Map of Greenland, 1:500 000*.

GLIMS and NSIDC (2005). *Global Land Ice Measurements from Space glacier database*. updated 2018; Compiled and made available by the international GLIMS community and the National Snow and Ice Data Center, Boulder CO, U.S.A. DOI: 10.7265/N5V98602.

GLIMS Viewer. [retrieved, March 3, 2020]. URL: http://www.glims.org/maps/glims.

Global Terrestrial Network for Glaciers (2018). *About GTN-G*. retrieved June 1, 2020. URL: https://www.gtn-g.ch/about_gtng/.

Gourmelen, N. et al. (2018). 'CryoSat-2 swath interferometric altimetry for mapping ice elevation and elevation change'. In: *Advances in Space Research* 62.6, pp. 1226–1242. ISSN: 18791948. DOI: 10.1016/j.asr.2017.11.014.

GTN-G Viewer. retrieved June 5, 2020. URL: http://www.glims.org/maps/gtng.

Guenin, G. *Polarisation et contre-polarisation*. [retrieved July 18, 2020]. URL: http://g.guenin.chez-alice.fr/theo/polarisa.htm.

Huang, Lei et al. (2011). 'Classification and snow line detection for glacial areas using the polarimetric SAR image'. In: *Remote Sensing of Environment*. ISSN: 00344257. DOI: 10.1016/j.rse.2011.03.004.

InterACT. *Arctic Station*. [retrieved June 11, 2020]. URL: https://eu-interact.org/field-sites/arctic-station.

International Association of Oil and Gas Producers (IOGP). *EPSG:3413*. URL: http://epsg.io/3413.

Kääb, Andreas et al. (2014). 'Glacier mapping and monitoring based on multispectral data'. In: *Global Land Ice Measurements from Space* 3, pp. 75–112. DOI: 10.1007/978-3-540-79818-7. URL: http://folk.uio.no/kaeaeb/publications/Chapter_4_kaab_v3.pdf.

Kargel, Jeffrey S, Gregory J Leonard and Michael P Bishop (2014). *Global Land Ice Measurements from Space*. ISBN: 9783540798170. DOI: 10.1007/978-3-540-79818-7.

Kundu, Sanchayita and Manab Chakraborty (2015). 'Delineation of glacial zones of Gangotri and other glaciers of Central Himalaya using RISAT-1 C-band dual-pol SAR'. In: *International Journal of Remote Sensing* 36.6, pp. 1529–1550. ISSN: 13665901. DOI: 10.1080/01431161.2015.1014972.

König, Max et al. (2002). 'Detection of superimposed ice on the glaciers Kongsvegen and midre Lovénbreen, Svalbard, using SAR satellite imagery'. In: *Annals of Glaciology* 34, 335–342. DOI: 10.3189/172756402781817617.

Lee, J. S. et al. (1994). 'Speckle filtering of synthetic aperture radar images: A review'. In: *Remote Sensing Reviews* 8.4, pp. 313–340. DOI: 10.1080/02757259409532206. URL: https://doi.org/10.1080/02757259409532206.

Lior, Rokach and Maimon Oded (2008). *Data Mining With Decision Trees: Theory And Applications.* Series in Machine Perception and Artificial Intelligence v. 69. World Scientific. ISBN: 9789812771711. URL: http://search.ebscohost.com/login.aspx?direct=true&db=e000xww&AN=236037&site=ehost-live.

Long, David G. and Mark R. Drinkwater (1994). 'Greenland ice-sheet surface properties observed by the Seasat-A scatterometer at enhanced resolution'. In: *Journal of Glaciology* 40.135, 213–230. DOI: 10.3189/S0022143000007310.

McMillan, Malcolm et al. (2018). 'Assessment of CryoSat-2 interferometric and non-interferometric SAR altimetry over ice sheets'. In: *Advances in Space Research*. ISSN: 18791948. DOI: 10.1016/j.asr.2017.11.036.

Meredith, M. et al. (2019). 'Polar Regions'. In: *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*.

Michel, Remi and Eric Rignot (1999). 'Flow of Glaciar Moreno, Argentina, from repeat-pass Shuttle Imaging Radar images: comparison of the phase correlation

method with radar interferometry'. In: *Journal of Glaciology* 45.149, 93–100. DOI: 10.3189/S0022143000003075.

NASA Earth Observatory. *NASA Earth Observatory Glossary*. retrieved August 8, 2020. URL: https://earthobservatory.nasa.gov/glossary/f/h.

Nascetti, A. et al. (2016). 'Exploiting Sentinel-1 amplitude data for glacier surface velocity field measurements: Feasibility demonstration on baltoro glacier'. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 41.July, pp. 783–788. ISSN: 16821750. DOI: 10.5194/isprsarchives-XLI-B7-783-2016.

Palmer, Steven J. et al. (2013). 'Greenland subglacial lakes detected by radar'. In: *Geophysical Research Letters* 40.23, pp. 6154–6159. ISSN: 00948276. DOI: 10.1002/2013GL058383.

Paul, F. et al. (2013). 'On the accuracy of glacier outlines derived from remote-sensing data'. In: *Annals of Glaciology* 54.63, 171–182. DOI: 10.3189/2013AoG63A296.

Paul, Frank et al. (2015). 'The glaciers climate change initiative: Methods for creating glacier area, elevation change and velocity products'. In: *Remote Sensing of Environment* 162, pp. 408–426. ISSN: 00344257. DOI: 10.1016/j.rse.2013.07.043.

Pfeffer, W. Tad et al. (2014). 'The Randolph Glacier Inventory: a globally complete inventory of glaciers'. In: *Journal of Glaciology* 60.221, 537–552. DOI: 10.3189/2014JoG13J176.

Polar Geospatial Center. *ArcticDEM*. University of Minnesota. URL: https://www.pgc.umn.edu/data/arcticdem/.

Pörtner, H.-O. et al. (2019). 'IPCC, 2019: Summary for Policymakers'. In: *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*.

QGIS.org. *QGIS Geographic Information System. Open Source Geospatial Foundation Project.* URL: http://qgis.org.

Radić, Valentina and Regine Hock (Nov. 2013). 'Glaciers in the Earth's Hydrological Cycle: Assessments of Glacier Mass and Runoff Changes on Global and Regional Scales'. In: *Surveys in Geophysics* 35. DOI: 10.1007/s10712-013-9262-y.

Raup, B.H. et al. (2007). *The GLIMS Geospatial Glacier Database: A New Tool for Studying Glacier Change*. DOI: 10.1016/j.gloplacha.2006.07.018.

Roberts, David H. et al. (2009). 'Ice marginal dynamics during surge activity, Kuannersuit Glacier, Disko Island, West Greenland'. In: *Quaternary Science Reviews* 28.3, pp. 209 –222. DOI: https://doi.org/10.1016/j.quascirev.2008.10.022. URL: http://www.sciencedirect.com/science/article/pii/S0277379108003004.

Robson, Benjamin Aubrey et al. (2015). 'Automated classification of debris-covered glaciers combining optical, SAR and topographic data in an object-based environment'. In: *Remote Sensing of Environment*. ISSN: 00344257. DOI: 10.1016/j.rse.2015.10.001.

Sentinel-2 User Handbook (2013). retrieved June 9, 2020. European Space Agency. URL: https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2_User_Handbook.

Sentinel Application Platform (SNAP). URL: http://step.esa.int/main/toolboxes/snap/.

Shi, Yueling et al. (Feb. 2019). 'Assessing the Glacier Boundaries in the Qinghai-Tibetan Plateau of China by Multi-Temporal Coherence Estimation with Sentinel-1A InSAR'. In: *Remote Sensing* 11.4, p. 392. ISSN: 2072-4292. DOI: 10.3390/rs11040392. URL: http://www.mdpi.com/2072-4292/11/4/392.

Smith, Taylor, Bodo Bookhagen and F. Cannon (Oct. 2014). 'Improving semi-automated glacial mapping with a multi-method approach: areal changes in Central Asia'. In: *The Cryosphere Discussions* 8, pp. 5433–5483. DOI: 10.5194/tcd-8-5433-2014.

Snow, National and Ice Data Center. *Glacier Types: Rock*. [retrieved June 27,2020]. URL: https://nsidc.org/cryosphere/glaciers/gallery/rock.html.

Strozzi, Tazio et al. (2002). 'Glacier motion estimation using SAR offset-tracking procedures'. In: *IEEE Transactions on Geoscience and Remote Sensing* 40.11, pp. 2384–2391. ISSN: 01962892. DOI: 10.1109/TGRS.2002.805079.

System for Automated Geoscientific Analyses (SAGA). URL: http://www.saga-gis.org/en/index.html.

Valenti, Linda, David Small and Erich Meier (Feb. 2008). 'Snow cover monitoring using multitemporal ENVISAT/ASAR data'. In: pp. 11–13. DOI: 10.5167/uzh-6433.

Watson, C. Scott, Owen King, Evan S. Miles and Duncan J. Quincey (Nov. 2018). 'Optimising NDWI supraglacial pond classification on Himalayan debris-covered glaciers'. In: *Remote Sensing of Environment* 217, pp. 414–425. ISSN: 00344257. DOI: 10.1016/j.rse.2018.08.020. URL: https://linkinghub.elsevier.com/retrieve/pii/S0034425718303961.

Winsvold, Solveig H. et al. (2018). 'Using SAR satellite data time series for regional glacier mapping'. In: *Cryosphere* 12.3, pp. 867–890. ISSN: 19940424. DOI: 10.5194/tc-12-867-2018.

Yang, Yan et al. (2016). 'Extraction of glacier outlines and water-eroded stripes using high-resolution SAR imagery'. In: *International Journal of Remote Sensing* 37.5,

pp. 1016–1034. ISSN: 13665901. DOI: 10.1080/01431161.2016.1145365. URL: http://dx.doi.org/10.1080/01431161.2016.1145365.

Yde, Jacob C. (2011). 'Greenland Glaciers Outside the Ice Sheet'. In: *Encyclopedia of Snow, Ice and Glaciers*. Ed. by Vijay P. Singh, Pratap Singh and Umesh K. Haritashya. Dordrecht: Springer Netherlands, pp. 478–484. ISBN: 978-90-481-2642-2. DOI: 10.1007/978-90-481-2642-2_643. URL: https://doi.org/10.1007/978-90-481-2642-2_643.

Yde, Jacob C. and N. Knudsen (Oct. 2007). '20th-century glacier fluctuations on Disko Island (Qeqertarsuaq), Greenland'. In: *Annals of Glaciology* 46, pp. 209–214. DOI: 10.3189/172756407782871558.

Zhang, Jingxiao, Li Jia, Massimo Menenti and Guangcheng Hu (Feb. 2019). 'Glacier Facies Mapping Using a Machine-Learning Algorithm: The Parlung Zangbo Basin Case Study'. In: *Remote Sensing* 11, p. 452. DOI: 10.3390/rs11040452.

Zhou, Chunxia and Lei Zheng (Nov. 2017). 'Mapping Radar Glacier Zones and Dry Snow Line in the Antarctic Peninsula Using Sentinel-1 Images'. In: *Remote Sensing* 9, p. 1171. DOI: 10.3390/rs9111171.

Zhou, Jianmin, Zhen Li and Qiang Xing (2010). 'Deriving glacier border information based on analysis of decorrelation in SAR interferometry'. In: 2.3, pp. 212–217.

Appendix A

Tools

RUS Copernicus Service This work/project has been carried out using the resources made available by the *Copernicus Research User Support* (RUS) service. RUS is a free service which aims at promoting the uptake of Copernicus data by providing cloud-based working environments customised to users's needs, together with the support of technical and thematic experts as well as dedicated training resources.

Sentinel Application Platform (SNAP) The Sentinel Application Platform (SNAP) architecture is a framework containing toolboxes for different Earth Observation Data, among them Sentinel-1 and Sentinel-2. In this project, the Graph Processing Framework (GPF) which facilitates the creation of user-defined processing chains (as XML-files) has mainly been used, along with the desktop application.

QuantumGIS (QGIS) QGIS is an Open Source Geographic Information System which supports viewing, editing, and analysis of geospatial data *QGIS Geographic Information System. Open Source Geospatial Foundation Project.*

Python Python 3.7 was used to read and analyse the raster data covering satellite imagery, the DEM and other binary masks. For the supervised and unsupervised learning algorithms, the *scikit-learn* package was used.

GDAL The Geospatial Data Abstraction Library – GDAL – is an open source library to read, write, translate and process raster and vector formats. Its functions are accessible via QGIS and through calling python scripts from the terminal. In this study, GDAL was used to clip raster data and merge bands from satellite imagery.

NTNU

Norwegian University of
Science and Technology