Monitoring Greenland and Antarctic supraglacial hydrology from space

Diarmuid Corr, MSci (Hons)

Lancaster Environment Centre

Lancaster University

A thesis submitted for the degree of

Doctor of Philosophy

January, 2024

This thesis is dedicated to my friends, family, and anyone who has made me a cup of tea. I extend my heartfelt thanks for your support, which has helped me reach this point in my academic journey.

Declaration

I declare that the work presented in this thesis is, to the best of my knowledge and belief, original and my own work. The material has not been submitted, either in whole or in part, for a degree at this university or any other. This thesis does not exceed the maximum permitted word length of 80,000 words, including appendices and footnotes but excluding the bibliography. An approximate estimate of the word count is: 46140.

Diarmuid Corr

Monitoring Greenland and Antarctic supraglacial hydrology from space

Diarmuid Corr, MSci (Hons).

Lancaster Environment Centre, Lancaster University

A thesis submitted for the degree of *Doctor of Philosophy*. January, 2024

Abstract

Ice sheets, large masses of glacial ice covering polar regions, influence global sea level and ocean currents. The study of surface water on these ice sheets, supraglacial hydrology, is essential to understand the effects of climate change on ice sheet stability, sea-level rise, and climate systems. This thesis examines supraglacial hydrological systems in Antarctica and Greenland by developing novel methods to classify them using optical satellite imagery (Sentinel-2 and Landsat-8).

Chapter 2 reveals the presence of supraglacial hydrology features, such as lakes and channels, on the West Antarctic Ice Sheet through a novel dual-NDWI and k-means clustering approach. A total of 10,478 features covering 119.4 km² were identified, broadening our knowledge of Antarctica's supraglacial hydrology.

Chapter 3 uses random forest and radiative transfer models to analyse the extent and volume of surface meltwater on the Greenland Ice Sheet from 2014 to 2022. This study assesses supraglacial hydrological features Greenland wide, on a decadal scale, for the first time. The results imply that reductions in firn air content and increases in ice slab content are drivers of increasing meltwater in various drainage basins, particularly in the north, east, and south.

Chapter 4 presents an innovative algorithm that quantifies uncertainty in the prediction of supraglacial hydrology using Bayesian inference with spatial statistics. This probabilistic approach provides predictions for the presence of water

at the pixel level with associated standard deviations, which signify uncertainty. By quantifying uncertainty, this approach is important for understanding the quantity and trends of meltwater flowing into the ocean.

This research advances our understanding of the distribution and dynamics of supraglacial hydrology on ice sheets, providing data and tools for the wider scientific community. These findings contribute to our understanding of the impacts of climate change on polar regions and support machine learning models to map surface water.

Publications

The candidate confirms that the work submitted is their own, except where work that has been included as part of jointly authored publications. The contribution of the candidate and the other authors to this work has been explicitly indicated. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The research presented in Chapter 2 is published in Earth Systems Science Data:

D. Corr et al. (2022). "An inventory of supraglacial lakes and channels across the West Antarctic Ice Sheet". In: Earth Systems Science Data.

The research presented in Chapter 3 and Chapter 4 is prepared as manuscripts to be submitted in the coming weeks as:

D. Corr et al. (In Prep.). "Supraglacial hydrology coverage has increased on the Greenland ice sheet over the last decade". In: Nature Climate Change (Intended).

D. Corr et al. (In Prep.). "A new method for probabilistic prediction of supraglacial lakes on the southwest Greenland Ice Sheet". In: Environmetrics (Intended).

I am listed as a named author in the following publications, this work does not appear in the thesis, but was fundamental in guiding the thesis into what it has become:

T. J. Barnes et al. (In Prep.). "Changes in Supraglacial Lakes on George VI Ice Shelf, Antarctic Peninsula: 1973-2020". In: The Cryosphere Discussions.

E. Glen et al. (In Prep.). "A comparison of supraglacial meltwater features throughout an extreme high, and an extreme low, melt season: Southwest Greenland". In: The Cryosphere (Intended).

Acknowledgements

I would like to express my gratitude to the individuals and institutions who have supported me throughout this academic journey. I am immensely thankful to my supervisors Amber Leeson and Malcolm McMillan. Without their unwavering encouragement, guidance, and contributions, this thesis would not have been possible. To Israel Martinez Hernandez and Ce Zhang, your mentorship and expertise have guided me through the complex world of statistics and machine learning. To Jennifer Maddalena for all your hard work and for lending me an ear when I needed it most, I am eternally grateful. Thanks to educators, teachers, lecturers, and mentors, you provide the base for me and countless others to excel in their lives. Thanks to my viva examiners, James Lea and Alex Bush, for taking the time to critique this body of work and for an insightful conversation on the topic.

This research was supported by the Engineering and Physical Sciences Research Council and the European Space Agency and was conducted at Lancaster University. Thanks to the UK Centre for Polar Observation and Modelling for providing a community in which polar scientists can flourish. I express my gratitude for all the financial support and opportunities provided to me during my PhD. Thanks to my co-authors and other collaborators, particularly those at Earthwave and S[&]T.

Thanks to the Lancaster Glaciology Group for introducing me to the icy world and welcoming me back to Lancaster each time I visited. To Michel Tsamados and those at CPOM-UCL for accepting me into your research group and providing me with an academic home in London, I thank you very much. Thanks to those in Lancaster and London for making my PhD life so enjoyable. In no particular order, thanks to Jade, Vincent, Emily, Joe, Laura, Robert, Thomas, Dom, Ali, Connor, Eve, Adam, Nick, Will, Jon, Harry, Carmen, Blaise, and Robbie for being around for all

the chats and beers. To Tom, Elias, and Isaac, thanks for those endless discussions of Premier League greats over cups of coffee or work-dodging pints, they got me through this process.

To those friends outside academia who gave me a reality check when I needed one, I thank you. To flatmates, Bobs, and lads, thanks for sticking around. To Lorcan, Rónán, Niall, Mark, Sean, Caomhlán, James, Conor, Shaun, Anna, Juliette, Ciaran, Ffion, Alex, Lea, and Jack, thanks for everything. A special word of thanks must go to those who I met during my PhD and who became some of my closest friends. To Matt, Jez, and Tom, thanks for introducing me to the Lakes, challenging me, and providing well-needed advice and encouragement.

To the Cassidys, thank you for accepting me into your family; you have always made me feel welcome, and I appreciate that deeply. To my siblings, Katie, Eunan, Aoife, Cliodhna, Dara, Calum, Malena, and Claire, thank you for everything. Thanks for the discussions and arguments, parties and laughs, and most importantly, for never disagreeing when I say I am right. To my Nancy, I hope you grow up with natural curiosity and a flair for science. To my parents, Michael and Dolores, you have made me the person I am today. Thank you for the support, care, shelter, cooking, cleaning, dog-walking, and everything else over the last four years and my entire life. I can assure you that the eternal student's studying is now complete.

Finally, to Ellen. I cannot put into words how much I admire you. Thank you for your unwavering support when we lived together, but even more so when we lived apart. Your kindness and love have inspired me to complete this thesis. You have been with me through every high and low, and I thank you greatly.

Abbreviations and acronyms

ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AVHRR	Advanced Very High Resolution Radiometer
A:P	Area-Perimeter Ratio
AIS	Antarctic Ice Sheet
AP	Antarctic Peninsula
Auto	Automated
CNN	Convolutional Neural Network
EAIS	East Antarctic Ice Sheet
EROS	Earth Resources Observation Science
ESA	European Space Agency
DEM	Digital Elevation Model
DN	Digital Numbers
ELA	Equilibrium-Line Altitude
GVI	George VI
GI	Grounded Ice
GIS	Geographic Information Systems
GL	Grounding Line
GrIS	Greenland Ice Sheet
IS	Ice Shelf
INLA-SPDE	Integrated Nested Laplace Approximation with Stochastic Partial Differential Equation

IPQ	Iso-Primetric Quotient
L8	Landsat-8
LIDAR	Laser Imaging, Detection, and Ranging
LWIR	Long-Wave Infra-Red
Man	Manual
MBC	Minimum Bounding Circle
MBR	Minimum Bounding Rectangle
ML	Machine Learning
MODIS	Moderate Resolution Imaging Spectroradiometer
MSI	Multi-Spectral Instrument
NASA	National Aeronautics and Space Administration
NDSI	Normalised Difference Snow Index
NDWI	Normalised Difference Water Index
NEGIS	North East Greenland Ice Stream
NIR	Near Infra-Red
OLI	Operational Land Imager
QA	Quality Assurance
RACMO2	Regional Atmospheric Climate Model
RADAR	Radio Detection and Ranging
RES	Radio-Echo Sounding
RF	Random Forest
RGB	Red, Blue, Green

RMSE	Root Mean Square Error
RTM	Radiative Transfer Model
S2	Sentinel-2
SAR	Synthetic Aperture Radar
SGC	Surpraglacial Channels
SGH	Surpraglacial Hydrology
SGHF	Surpraglacial Hydrology Features
SGL	Surpraglacial Lakes
SMB	Surface Mass Balance
SWIR	Short-Wave Infra-Red
TIRS	Thermal Infra-Red Sensor
TOA	Top-of-Atmosphere
USGS	United States Geological Survey
WAIS	West Antarctic Ice Sheet

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this chapter, I provide an overview of Earth's two major ice sheets, Antarctica and Greenland, and the hydrological system that links their supra-, en- and subglacial environments. I examine the significance and known networks of supraglacial hydrology and its impact on ice sheets. Additionally, I explain the techniques used to map supraglacial features and outline the available remote sensing options. I identify knowledge gaps in the literature and discuss the difficulties we face as a community. Lastly, I outline the thesis aims and objectives, before introducing the structure of the rest of the thesis.

1.1 Ice sheets

Ice sheets are a significant part of the Earth system, influencing global sea level, ocean currents, and biogeochemical processes. An ice sheet is a large mass of glacial ice that covers an area of more than 50,000 km^2. Currently, there are only two ice sheets on Earth, the Antarctic Ice Sheet (AIS) and the Greenland Ice Sheet (GrIS). The melting of these two ice sheets is expected to be the main contributor to the rise in sea levels in the near future (Fox-Kemper et al., 2021). I begin by providing an

overview of the geography and climate of each ice shelf. Following this, I discuss the processes by which the ice sheet interacts with its environment and the observations of mass balance change and sea level rise. Finally, I examine the factors that are causing the current imbalance of the ice sheet.

1.1.1 Physical geography and climate

The AIS covers an area of approximately 14 million km^2, which is almost the whole continent. The average thickness of the ice is more than 2 km, and it holds 61% of the world's fresh water, or the equivalent of 58 m of potential sea level rise (Bamber et al., 2018; Fretwell et al., 2013; D. Vaughan et al., 2013). The AIS is usually divided into three geographic regions: the East Antarctic Ice Sheet (EAIS), the West Antarctic Ice Sheet (WAIS), and the Antarctic Peninsula (AP) (Figure 1.1). Ice that is on top of bedrock and therefore has land at its base is known as grounded ice. Around the grounding line, where the grounded ice meets the ocean, there is floating ice. Floating ice around the coast of Antarctica forms features known as ice shelves. The largest of these, Ross and Filchner-Ronne, are more than 400,000 km^2. Ice streams, which are corridors of fast-flowing ice, transport ice from the interior to the coastline, flowing through glaciers to the ice shelves or directly into the sea.

The GrIS covers an area of approximately 1.7 million km^2, which is equivalent to around 80% of the surface of Greenland. The average thickness of the ice is 1.5 km, and it holds 7% of the world's fresh water, which is the equivalent of 7 m of potential sea level rise (Bamber et al., 2018; D. Vaughan et al., 2013). In the interior, the ice can be up to 3.2 km thick (Figure 1.2), while coastal mountains encircle most of the ice sheet's edge (Morlighem et al., 2017). There are no ice shelves, as the ice sheet reaches the sea and drains the interior through outlet glaciers and ice streams. These two types of glaciers are determined by the settings in which they end, either marine or

Figure 1.1: **Labelled map of Antarctica. Image credit: Landsat Image Mosaic of the Antarctica team** `https://lima.usgs.gov/documents/LIMA_overview_map.pdf`

land-terminating. Large marine-terminating glaciers, such as the Petermann glacier in northwest Greenland, form floating ice tongues before calving into the sea. Ice streams, such as the 600 km long North East Greenland Ice Stream (NEGIS), also drain a considerable amount of the ice sheet.

The atmospheric conditions of Antarctica and Greenland are shaped by their polar locations and expansive ice sheets, yet they have distinct characteristics due to their different geographies and surrounding oceans. Antarctica, the

Figure 1.2: **Map of Greenland with ice sheet thicknesses. Image credit: Eric Gaba – Wikimedia Commons user: Sting**

southernmost continent, has an extreme polar climate with prolonged periods of polar day and night, and temperatures that can drop to -80°C during winter (*Antarctic weather* 2019). Coastal areas tend to be warmer and receive more

precipitation than in the interior (Lenaerts et al., 2019), with an average annual temperature ranging from around -10 °C near the coast to -60 °C on the highest peaks (*Antarctic weather* 2019). The Southern Ocean is known for its extreme conditions and is encircled by the Antarctic Circumpolar Current, which plays a major role in redistributing heat and nutrients throughout the world's oceans and regulating global climate (Marshall and Speer, 2012; Rintoul, 2018). Its frigid waters are kept close to freezing point due to the influence of the Antarctic ice sheet, which facilitates the formation of sea ice during winter (Comiso and Nishio, 2008; Gloersen, 1992; Massom and Stammerjohn, 2010). Katabatic winds, which originate in the interior of the ice sheet, can reach high speeds and affect local weather conditions (Barral et al., 2014; Grazioli et al., 2017; M. R. van den Broeke, 1997). Although coastal areas receive some precipitation, the interior is exceptionally dry, making Antarctica the coldest and driest continent on Earth (*How cold is the Antarctic?* 2017). Greenland, on the other hand, is classified as tundra, which means it has at least one month with average temperatures high enough to melt snow, but no months when the average temperature exceeds 10°C, around the ice-free coast (Kottek et al., 2006). It has an Arctic climate with milder temperatures than Antarctica and also experiences polar day and night, with winter temperatures dropping to -50 ° C or below (Hanna et al., 2021), while the average summer temperature ranges from +8 °C near the coast to -16 °C at the highest interior peaks (Hanna et al., 2021). The region is influenced by the Arctic Ocean, which moderates coastal temperatures (Buch, 2002). It is bordered by the Atlantic Ocean to the south and the Arctic Ocean to the north. The boundary currents of the Arctic Ocean, the East Greenland Current and the West Greenland Current, transport cold, nutrient rich waters from the Arctic Ocean southward until they mix with warmer Atlantic waters, creating a dynamic transition zone with varying temperatures and salinities (Buch, 2002; Gou, Pennelly, and Myers, 2022; Håvik et al., 2017; Münchow, Falkner,

and H. Melling, 2015). Sea ice also forms seasonally during Arctic winter (Comiso and Nishio, 2008; Gloersen, 1992).

A study of 17 Antarctic weather stations showed that 13 of them had a positive (warming) trend in their annual mean temperature from 1979 to 2018, with the Vernadsky station on the western Antarctic Peninsula warming at a rate of 0.46 $\pm$ 0.15 °C per decade (Turner et al., 2020). Warming was most noticeable in spring (Turner et al., 2020). In Greenland, overall temperature trends for 2001-2019 are generally insignificant due to a cooling period for 2013-2019; however, significant coastal warming $\sim$+1.7 °C in summer and $\sim$+4.4 °C in winter has been observed for the period 1991-2019 (Hanna et al., 2021).

1.1.2 Interactions between ice sheets and the Earth system

The relationship between the ice sheets and their surrounding climate is important. Ice sheets interact directly with four main components: atmosphere, lithosphere, ocean, and sea ice (Table 1.1, Fyke et al., 2018). The interactions that are not relevant to this thesis are not discussed here, and only those that are essential for understanding supraglacial hydrology processes are examined. This interaction is a two-way system; i.e. each component exerts its effect on the ice sheet, while, properties of the ice sheet also exert an effect on the components. These interactions determine how external forces (such as human-induced climate change) affect changes in the ice sheet (Fyke et al., 2018). The complex relationship between ice sheets and climate can be explained by the concept of mass balance. The mass balance is the net difference between the mass gained through accumulation processes (e.g. precipitation) and the mass lost through ablation processes (e.g. melting at the surface and bed, sublimation). The contribution of Earth's ice sheets to the global rise of sea level depends on the ice sheets' mass balances.

Table 1.1: **Interaction between ice sheets and surrounding Earth system components. Source: adapted from Fyke et al., 2018.**

Interface	Direction	Interaction
Ice Sheet - Atmosphere	To Ice Sheet	Surface Energy Fluxes
		Surface Mass Fluxes
	From Ice Sheet	Surface Topography
		Ice Sheet Extent
		Surface Type
Ice Sheet - Lithosphere/Mantle	To Ice Sheet	Geothermal Heat Flux
		Subglacial Water Pressure
		Bed Elevation and Type
	From Ice Sheet	Ice Base Normal Stress-
		-and Mass Loading
		Ice Basal Velocity
Ice Sheet - Ocean	To Ice Sheet	Sub-Ice Shelf Energy Fluxes
		Sub-Ice Shelf Mass fluxes
	From Ice Sheet	Ice Melt Runoff
		Ice Shelf Geometry
		Iceberg Calving
		Gravitational Sea Level Effects
Ice Sheet - Sea Ice	To Ice Sheet	Sea Ice Back Stress
	From Ice Sheet	Ice Shelf Displacement

The thickness and size of the GrIS is determined by current climate conditions, which are responsible for the accumulation of mass from precipitation, the loss of mass from melting and sublimation, and the loss of ice to the oceans by the calving

or melting of the icebergs. The Antarctic Ice Sheet (AIS) is also affected by climate, but there are some important differences due to the higher elevation of the ice sheet, the more uniform atmospheric circulation, and the cold temperatures of the Southern Ocean (Fyke et al., 2018). In recent times, almost all (90%) of the precipitation over the Greenland Ice Sheet (GrIS) has been in the form of snow, corresponding to the accumulation of mass (Ettema et al., 2009). The remaining 10% of precipitation has been in liquid form (rainfall) at lower altitudes of the ice sheet during summer. However, these proportions can vary significantly from year to year due to internal climate variability (B. Noël et al., 2015). At present, the mass gain of the AIS is almost entirely due to snowfall, as the atmosphere is not warm enough for rainfall even in the summer months (van Wessem et al., 2014).

Surface melting is caused by a net positive surface energy balance over snow/ice at 0°C. In their respective summers, the incoming solar energy is at its peak, and the amount of solar energy absorbed by the ice or snow is largely determined by the surface albedo, which is higher in winter due to the presence of highly reflective snow. Once liquid water is produced at the surface, the porosity of the surface and subsurface, its temperature, and its density determine its fate. Ice is virtually impermeable, leading to the formation of supraglacial meltwater lakes (e.g., M. Moussavi et al., 2020; Stokes et al., 2019; Sundal et al., 2009; K. Yang et al., 2021) or englacial firn aquifers (Montgomery et al., 2020) that store water locally and supraglacial channels (or moulins) that transport water across (or through) the ice surface (Hoffman et al., 2018; Leidman et al., 2023; Spergel et al., 2021). Compared to GrIS, Antarctic surface melt extent, and its contribution to the mass balance, is much smaller (Fyke et al., 2018; Picard and Fily, 2006).

Surface Mass Balance (SMB) is a component of the overall mass balance. SMB is the net difference between accumulation and ablation on the surface of the ice sheet (Kittel et al., 2021). The mass loss from basal melting and iceberg calving is known

as ice discharge. The contribution to the rise of global sea level for the respective ice sheets is the net difference between ice discharge and SMB. Systems that have a negative mass balance lose mass; for grounded ice, this mass loss directly contributes to sea-level rise. Mass loss from floating ice shelves and tongues does not directly contribute to sea level rise, as ice has a lower density than water and therefore floats, displacing the water below.

Similarly, sea ice does not directly contribute to sea level rise; however, it has a major effect on the dynamics of Arctic and Antarctic ice sheets. During winter, it serves as a protective cover over the ocean, acting as an insulator to limit the exchange of heat between the warmer ocean and the ice sheets (Cornish et al., 2022; Fuerst et al., 2016; Jenkins et al., 2010; Notz, 2009). This helps to maintain lower temperatures at the base of the ice sheet, reducing the extent of basal melt (Bintanja et al., 2013; Hellmer, 2004; Pritchard et al., 2012) and the thinning of Antarctic ice shelves (Jenkins et al., 2010; Massom et al., 2018; T. A. Scambos et al., 2004; D. J. Wingham, Wallis, and Shepherd, 2009). Sea ice also acts as a physical barrier to restrain the flow of ice from the interior of the ice sheet into the ocean and provides structural support to icebergs and floating ice tongues (Fuerst et al., 2016; Gudmundsson, 2013; Sun, Riel, and Minchew, 2023). Furthermore, sea ice has a high albedo, meaning that it reflects a significant portion of incoming solar radiation away from the ice sheet, cooling it and reducing surface melting (Curry, Schramm, and Ebert, 1995; Kashiwase et al., 2017).

The balance of interactions between ice sheets and climate is significantly different for the two ice sheets: the atmosphere has a greater influence on the GrIS, while the ocean plays a more critical role in the alteration of the AIS (Fyke et al., 2018; Picard and Fily, 2006). Understanding climatic variability is crucial to understanding how ice sheets respond to changing climate conditions.

1.1.3 Observations of mass balance change & sea level rise

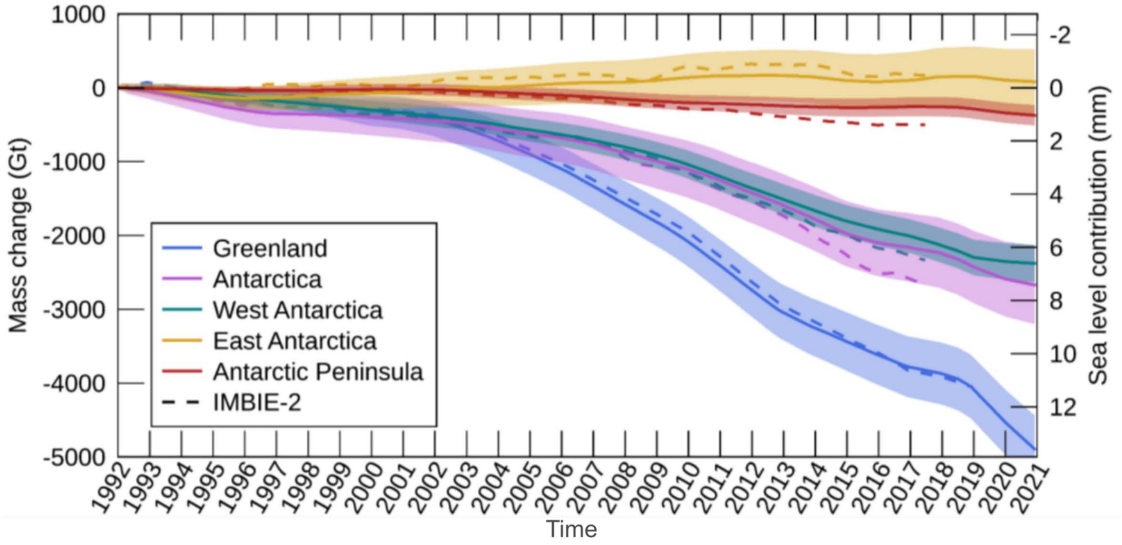


Figure 1.3: **Cumulative ice sheet mass changes. The estimated uncertainty 1σ of the cumulative change is shaded. The dashed lines show the results of previous assessments (IMBIE-2) (Otosaka et al., 2022; Shepherd et al., 2018, 2020).**

The total mass of ice on AIS and GrIS has decreased considerably in the last few decades (Madsen et al., 2022; Moon et al., 2020; Otosaka et al., 2022; Rignot et al., 2019; Shepherd et al., 2018, 2020; Slater et al., 2021). The Input-Outflow method is used to calculate the difference in mass between SMB and ice discharge (Rignot et al., 2008a; Rignot et al., 2008b). Satellite altimetry and gravimetry measurements are used to estimate the mass balance of the ice sheet. From 1992 to 2020, AIS has lost 2671 ± 530 Gt of ice (-92 ± 18 Gt yr^{-1}), and GrIS has lost $4892 \pm 699Gt$ (-169 ± 16 Gt yr^{-1}), resulting in a global sea level rise of 7.4 ± 1.5 mm and 21 ± 2 mm, respectively (Figure 1.4). The greatest mass loss in Antarctica is due to ice discharge from WAIS (Figure 1.4), which peaked at -131 ± 21 Gt yr^{-1} between 2012 and 2016 and slowed to -94 ± 25 Gt yr^{-1} from 2017 to 2020 (Otosaka et al., 2022). The collapse of the Larsen B ice shelf in the 2000s was a major factor in the mass changes on the AP, reaching

its peak at -21 ± 12 Gt yr^{-1} between 2007 and 2011 (Cook and D. G. Vaughan, 2010; Otosaka et al., 2022; Rignot et al., 2004). The mass change for EAIS has remained relatively stable at $+3 \pm 15$ Gt yr^{-1} over the 30-year period from 1992 to 2021 (Otosaka et al., 2022). In comparison, the rate of mass loss in Greenland is much greater and more variable, ranging from -35 ± 29 Gt yr^{-1} in the period from 1992 through 1996 to -280 ± 38 Gt yr^{-1} between 2007 and 2011, while 2019 is the year with the highest mass losses due to increased surface melting (-444 ± 93 Gt yr^{-1}, Otosaka et al., 2022; Tedesco and Fettweis, 2020).

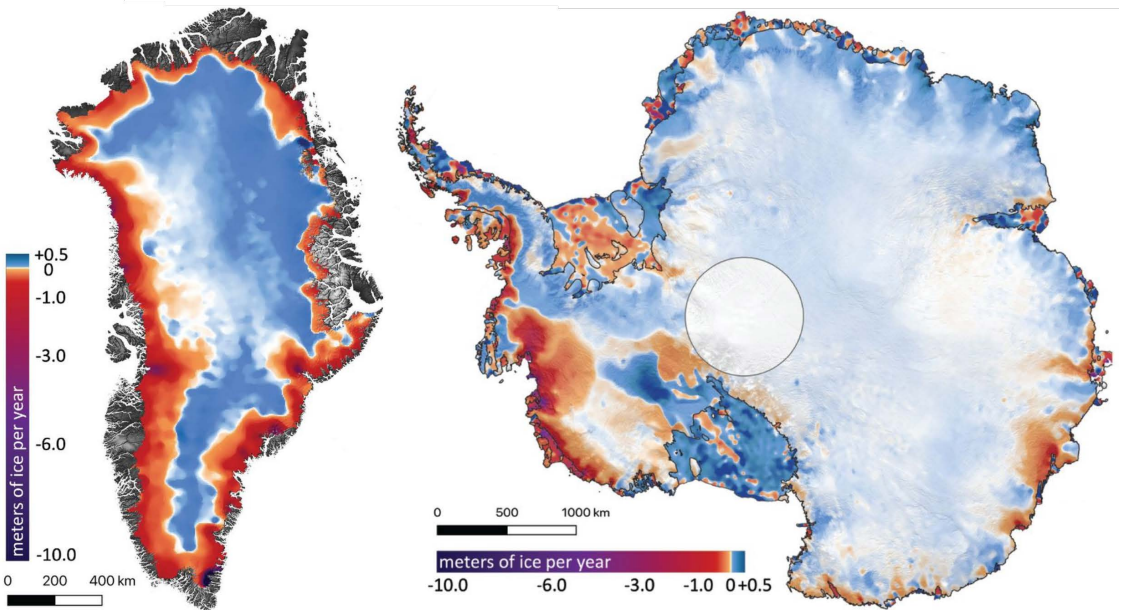


Figure 1.4: **Mass loss from Greenland and Antarctic Ice Sheets in metres of ice equivalent per year from 2003 to 2019. Adapted from B. Smith et al., 2020.**

1.1.4 Drivers of ice sheet imbalance

Ice sheets are sensitive to temperature changes, and the annual cycle of warmer summers has an effect on the amount and intensity of surface melting (Clarkson,

Eastoe, and A. Leeson, 2021; D. Liang et al., 2021; Vermeer and Rahmstorf, 2009). As temperatures rise, the level of surface meltwater runoff, surface melting, basal melting, and precipitation also increases (Lenaerts et al., 2019; Rignot et al., 2019; B. Smith et al., 2020). Research has shown that for every 1 °C of summer warming, the Greenland Ice Sheet (GrIS) loses 91 Gt yr $^{-1}$ of surface mass and experiences a 26 Gt yr $^{-1}$ increase in solid ice discharge (Hanna et al., 2021). Surface melting is much more widespread in Greenland than on the Antarctic Ice Sheet (AIS), where it occurs only around the margins and on ice shelves. Similarly, runoff is only present around the coast of GrIS, while on AIS, most meltwater is thought to refreeze in the firn (Lenaerts et al., 2012, 2017, Figure 1.5). Precipitation is a major factor in the growth or loss of ice sheets. More snowfall can lead to an increase in ice thickness, while less snowfall or more rainfall can cause a decrease (Boening et al., 2012; McIlhattan et al., 2020; Medley and E. R. Thomas, 2019; B. Noël et al., 2015). Generally, precipitation, which is mostly snow in areas higher than 500 metres above sea level, increases from the centre of the ice sheet to the edges (Figure 1.5). Values of less than 20 mm of water per year are seen in the interior of the East Antarctic Ice Sheet, while areas such as the Antarctic Peninsula and southeast Greenland have more than 2000 mm of water per year (Koenig et al., 2016; Miège et al., 2013).

Wind patterns can influence snow redistribution and snowdrift formation, which affects the surface characteristics of ice sheets. Strong winds can lead to increased sublimation, where the ice is transformed directly from the solid state to water vapour (Bintanja et al., 2013; Box and Steffen, 2001; Cullen et al., 2014). Sublimation occurs in coastal regions and is negligible or negative in areas of the interior (Figure 1.5). In addition, the North Atlantic Oscillation and the Southern Annular Mode are two atmospheric circulation patterns that can have a major impact on the ice sheet mass balance (Chylek, Box, and Lesins, 2004; Johannessen

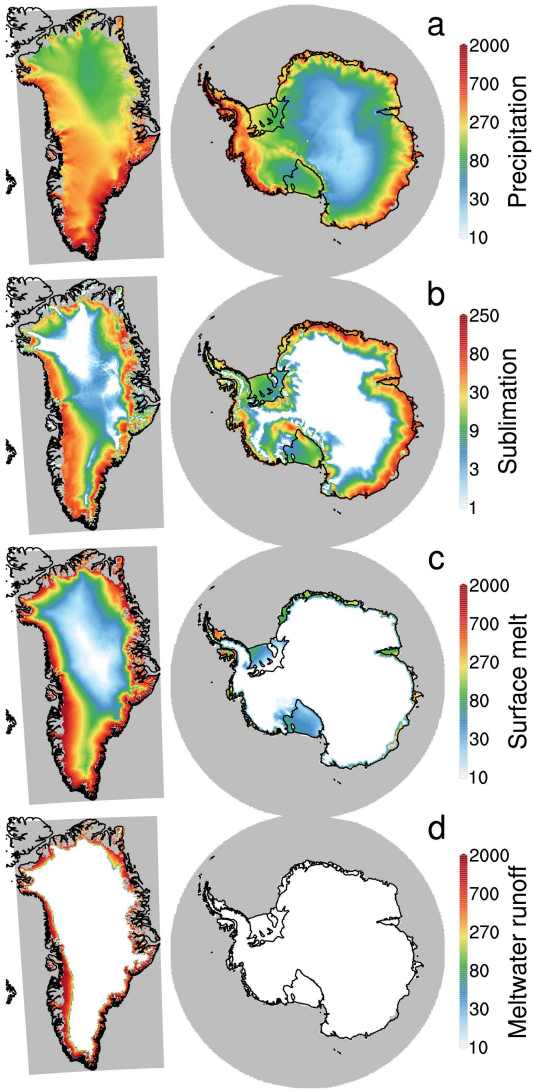


Figure 1.5: **Annual mean (mm w.e./year, 1980–2015) SMB Regional Atmospheric Climate Model (RACMO2) modelled components in GrIS (left) and AIS (right). (a) Precipitation; (b) Sublimation (surface + drifting snow); (c) Surface melt rate; (d) Meltwater runoff. (Lenaerts et al., 2019).**

et al., 2005; Verfaillie et al., 2022; Walker and Gardner, 2017). Due to warming temperatures in the Arctic and Antarctic, the extent of sea ice cover in the waters around both ice sheets has decreased (Bi, Y. Liang, and X. Chen, 2023; Comiso et al.,

2008; Hao et al., 2021; Purich and Doddridge, 2023; Stroeve et al., 2007). This exposes more of the ocean's surface to direct sunlight and warmer air temperatures, which accelerates the melting of the ice sheet's marine-terminating glaciers and contributes to sea level rise (Holland et al., 2008; Shepherd, D. Wingham, and Rignot, 2004; L. Zhang et al., 2022). Both regions are vulnerable to climate change, Greenland's sensitivity to warming temperatures has significant implications for sea level rise and Arctic ecosystems, and although Antarctica has not yet demonstrated that level of mass balance change, its unique characteristics contribute to understanding global climate systems and sea level rise (Fyke et al., 2018; Hanna et al., 2013).

1.2 Hydrology

Positive degree (°C) temperatures cause the melting of snow, ice, and firn (partially compacted snow from previous years, but not yet formed ice) on the AIS and GrIS. These melting conditions lead to the formation of liquid water, hydrology, on the surface of the ice (supraglacial), within the ice (englacial), and under the ice along the bed (subglacial). The composition of this meltwater is decided by the characteristics of the ice sheet. These characteristics decide if the melted water will refreeze within the snowpack or become runoff, if it will accumulate in supraglacial lakes or move across or through the ice sheet in channels, crevasses, or moulins, thus playing an important role in determining the highly dynamic transient hydrological system (Figure 1.6). Here, I detail how liquid water influences the ice sheets of Antarctica and Greenland and why it is critical to know more about the behaviour of this phenomenon in a warming world. I will begin by introducing the components of supraglacial, englacial, and subglacial hydrology. Section 1.3 will focus on the importance of supraglacial hydrology and its effects on ice sheet

systems. Section 1.4 will examine the known networks and extent of hydrology. Section 1.5 will analyse the available remote sensing data. Finally, Section 1.6 will investigate the techniques used to map supraglacial hydrology.

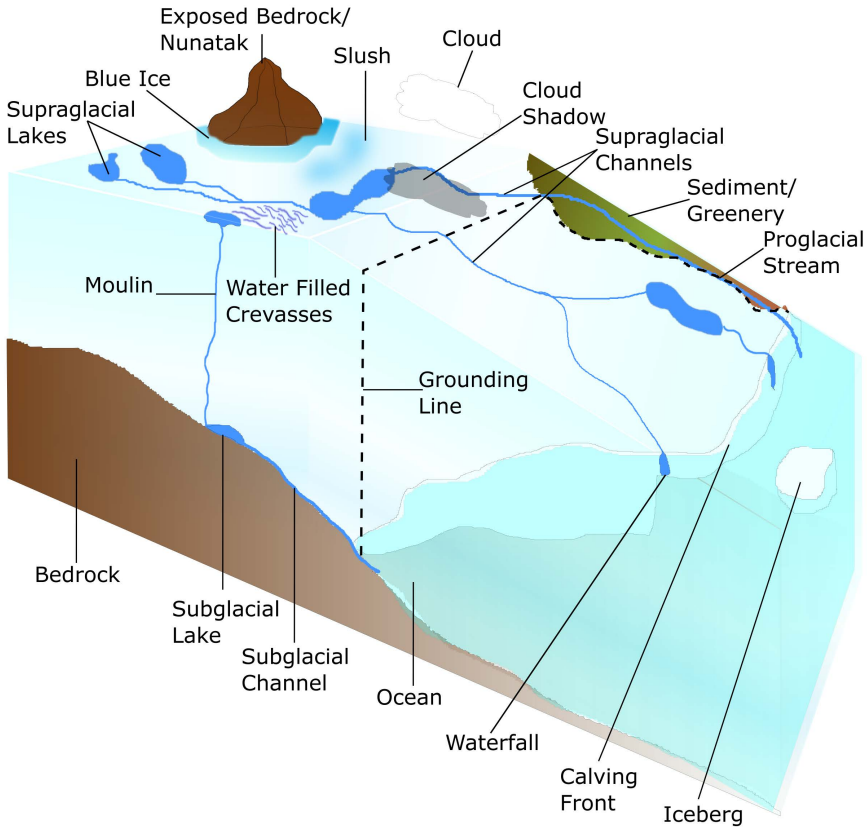


Figure 1.6: **Visualisation of the interconnected hydrological system.**

1.2.1 Supraglacial hydrology

When spring arrives, the sun's radiation and warmer air cause the surface of the ice and snow to heat up, leading to melting when there is enough energy available (Langley et al., 2016). The start and end of the melt seasons differ geographically on each ice sheet and can be affected by factors such as latitude, altitude, and local weather patterns (Johansson, Jansson, and Brown, 2013; Langley et al., 2016;

L. Liang et al., 2023; McMillan et al., 2007; Sundal et al., 2009; Trusel, Frey, and Das, 2012). Initially, the melt is absorbed into the snowpack, experiencing daily refreeze cycles on the surface and persistent melt deeper within the firn (Colliander et al., 2023). As the summer months bring higher temperatures, a higher melt intensity leads to a higher surface meltwater production (Trusel, Frey, and Das, 2012). Furthermore, when the air temperature is above 0 °C, there is a strong correlation between the surface air temperature and the melting rate on the surface of the ice (Braithwaite, 1995). When the meltwater gathers in undulations on the surface of the ice sheet, supraglacial lakes form (Bell et al., 2018; Box and Ski, 2007; Echelmeyer, Clarke, and Harrison, 1991; Langley et al., 2016; Selmes, Murray, and James, 2011). The production of meltwater is ultimately dependent on a complex combination of factors and conditions. During summer, increasing air temperatures cause surface melting (Langley et al., 2016). Solar radiation is an important factor, with sunlight being absorbed by the surface of the ice sheet, particularly when the albedo decreases due to melting (Leidman et al., 2021). Atmospheric circulation patterns can bring warm or cold air masses, which can influence temperature and precipitation on ice sheets (Liu, Wang, and Jezek, 2006). Furthermore, the wind redistributes heat and moisture and can increase sublimation and heat exchange (Laffin et al., 2022; Laffin, 2022).

As temperatures start to cool in autumn, the intensity of surface melting decreases and eventually stops. Supraglacial lakes begin to freeze over, and the surface of the ice sheet largely returns to a frozen state. However, the presence of wintertime buried lakes has recently been discovered (Dunmire et al., 2020; Dunmire et al., 2021; Koenig et al., 2015; K. E. Miles et al., 2017). During the winter months, the temperatures in the polar regions are usually very cold, often below freezing. As a result, the surface of ice sheets and glaciers remains frozen, with little or no melting (Langley et al., 2016; L. Liang et al., 2023).

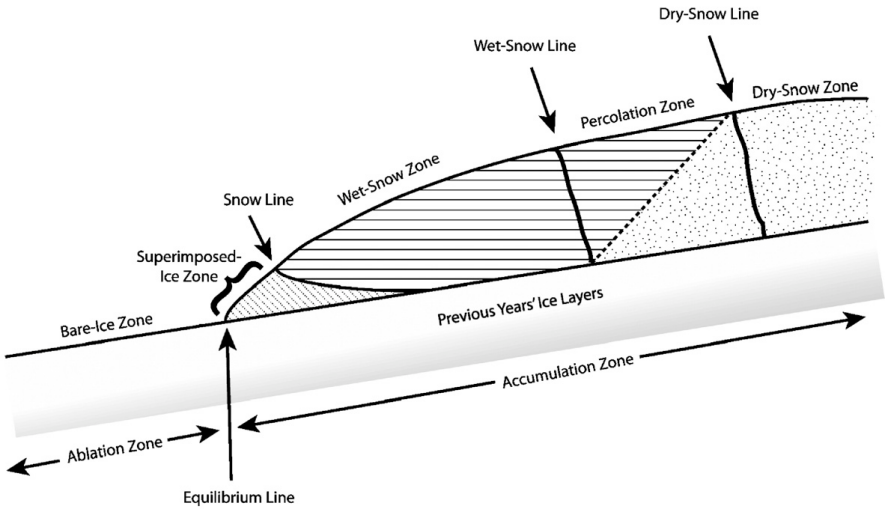


Figure 1.7: **Ice sheet accumulation (made up of wet-snow, percolation, and dry-snow zones) and ablation zones (M. Payne and A. Nolin, 2008).**

The surface of the ice sheets is divided into two distinct areas: the accumulation and ablation zones. The boundary between them is known as the Equilibrium-Line Altitude (ELA), which is a critical concept in glaciology. It is the elevation on a glacier or ice sheet where the mass balance is zero over a long period, usually a year. In other words, it is the line on a glacier where the amount of ice accumulated is equal to the amount of ice lost (Box et al., 2022; Braithwaite and Raper, 2009). The accumulation zone, located above the ELA, is composed of three distinct regions (Figure 1.7): the interior or dry-snow zone, where little melting occurs and hydrological activity is minimal; the percolation zone, where a limited amount of meltwater is produced during the melt season and infiltrates the snow and firn layers; and the wet snow zone, where the snow from the previous winter is completely saturated with meltwater (A. W. Nolin and M. C. Payne, 2007). The ablation zone, where more ice is lost than gained, is where the snow from the previous winter melts completely during the summer and exposes bare ice (Figure 1.7, A. W. Nolin and M. C. Payne, 2007).

In order for meltwater to accumulate on the surface of an ice sheet, it must be impeded by saturated firn or impermeable ice formed from refrozen water in the englacial environment (K. E. Alley et al., 2018; J. Harper et al., 2012; B. Hubbard et al., 2016; Lenaerts et al., 2017). The water then percolates downward through the porous snow and firn, following the preferred flow paths (Humphrey, J. T. Harper, and Pfeffer, 2012; Pfeffer and Humphrey, 1998). Once within the firn layer, the meltwater can be retained as liquid water (Humphrey, J. T. Harper, and Pfeffer, 2012), stored in perennial firn aquifers (Forster et al., 2014; Machguth et al., 2018), or refrozen (J. Harper et al., 2012; Machguth et al., 2018; Pfeffer, Meier, and Illangasekare, 1991). In Greenland, water refreezing can create three different types of features: thin ice lenses ($<$ 0.1 m thick), ice layers (0.1-1.0 m thick), or ice slabs (over 1 m to several m thick) (Culberg, Schroeder, and Chu, 2021; de la Peña et al., 2015; MacFerrin et al., 2019; Machguth et al., 2016). In Antarctica, repeated freezing and thawing cycles can lead to an increase in grain size and the potential formation of ice layers (Picard and Fily, 2006). Ice is impermeable to water flow, thus slowing or even stopping drainage and causing surface runoff, which can form supraglacial lakes, channels, or slush flows (Onesti and Hestnes, 1989). Supraglacial hydrology is the study of the intricate network of lakes and channels on the surface of the ice sheet (Figures 1.6, 1.9, 1.8). In the following sections, I will give an extensive overview of the elements and formation of the supraglacial hydrological network. I will then discuss how surface meltwater can move over the ice surface and directly enter the ocean (Bell et al., 2018), or be directed into sub- and englacial environments through fractures, crevices, and moulins (Hoffman et al., 2018; McGrath et al., 2012; van der Veen, 2007). In these areas, the meltwater can refreeze or eventually be routed into the ocean.

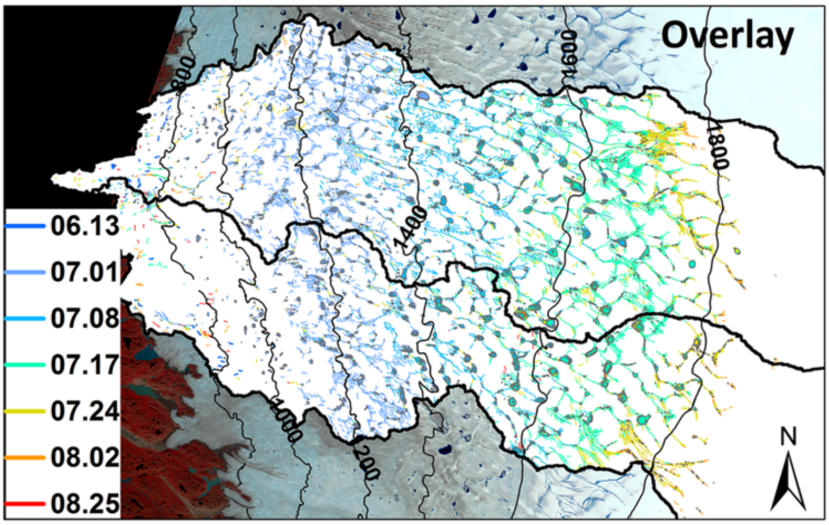


Figure 1.8: **A temporal composite of seven dates revealing seasonal evolution of supraglacial lakes and rivers on the southwest Greenland Ice Sheet as mapped from seven Landsat 8 OLI satellite images acquired throughout the 2015 melt season. Isortoq and Water river basins are overlaid (K. Yang et al., 2021).**

1.2.1.1 Supraglacial lakes

Supraglacial lakes, or SGLs, form when meltwater accumulates in depressions on the surface of an ice sheet (Echelmeyer, Clarke, and Harrison, 1991). The position of Surpraglacial Lakes (SGL)s on grounded ice is determined by the underlying topography of the bedrock, and they tend to form in the same or similar places in successive years (Bell et al., 2018; Box and Ski, 2007; Echelmeyer, Clarke, and Harrison, 1991; Langley et al., 2016; Selmes, Murray, and James, 2011). The depth and size of the lakes are affected by the amount of water in the system, the lake catchments, and the speed of ice flow (J. F. Arthur et al., 2020a; Turton et al., 2021). Slower and thicker ice is more resistant to crevassing and produces smoother and wider undulations than fast-moving ice (Gudmundsson, 2003). As a result, larger and deeper lakes form on slow-moving grounded ice, since they remain in the same

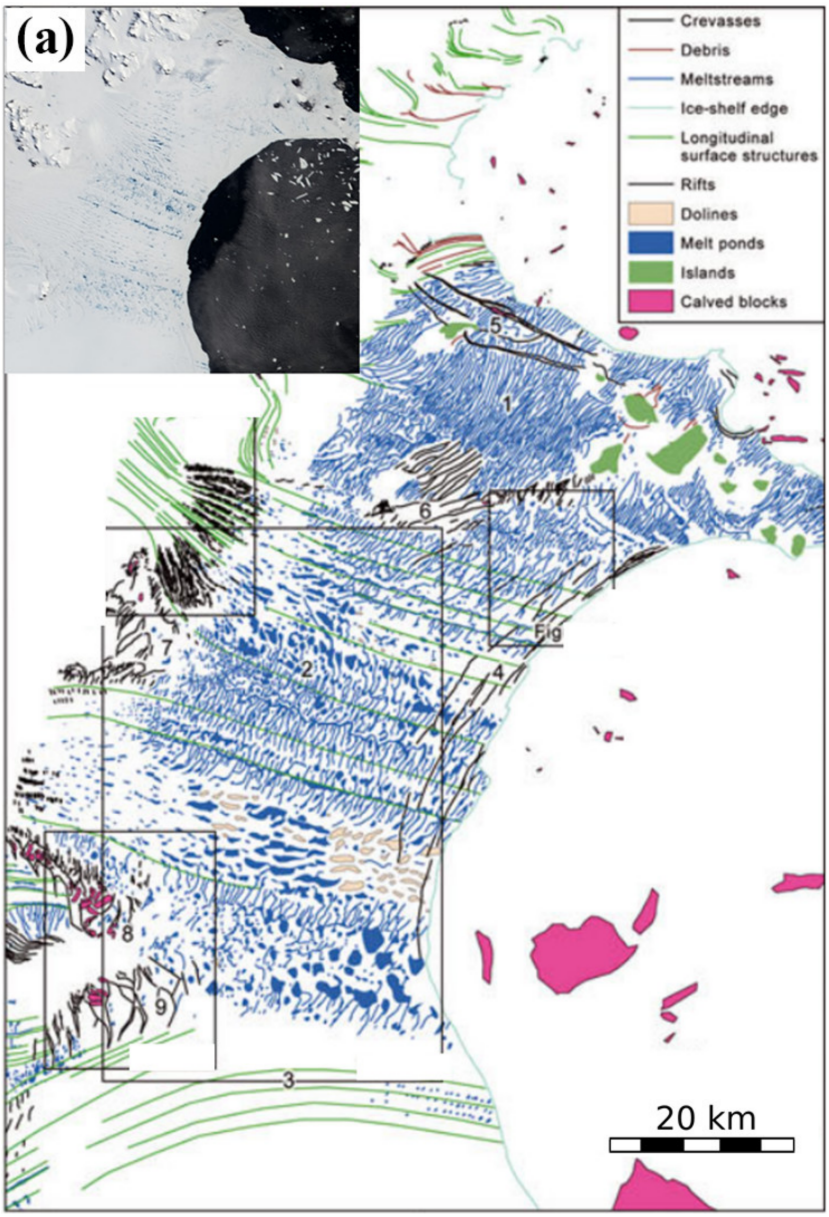


Figure 1.9: **Larsen B Ice Shelf surface structures in February 2002 including an extensive network of SGLs and streams that extend to the ice shelf calving front, alongside ice dolines indicating possible drained lakes (J. F. Arthur et al., 2020a).**

surface depressions longer, collecting more water from SGLs and expanding through lake bottom ablation (Banwell et al., 2014; Das et al., 2008; A. A. Leeson

et al., 2012; Macdonald, Banwell, and MacAyeal, 2018; Sergienko, 2013; Tedesco et al., 2012). On floating ice, lakes form in surface depressions that depend on and move with ice flow (Macdonald, Banwell, and MacAyeal, 2018). Depressions in the ice are caused by spatial and temporal changes in its flexure, flow, and thickness (Banwell et al., 2019). These are regulated by a combination of the position of the basal channels and crevasses (McGrath et al., 2012), the flow stripes, the suture zones, and the shear margins (Banwell et al., 2014; Bell et al., 2017; Ely et al., 2017; Glasser and Gudmundsson, 2012).

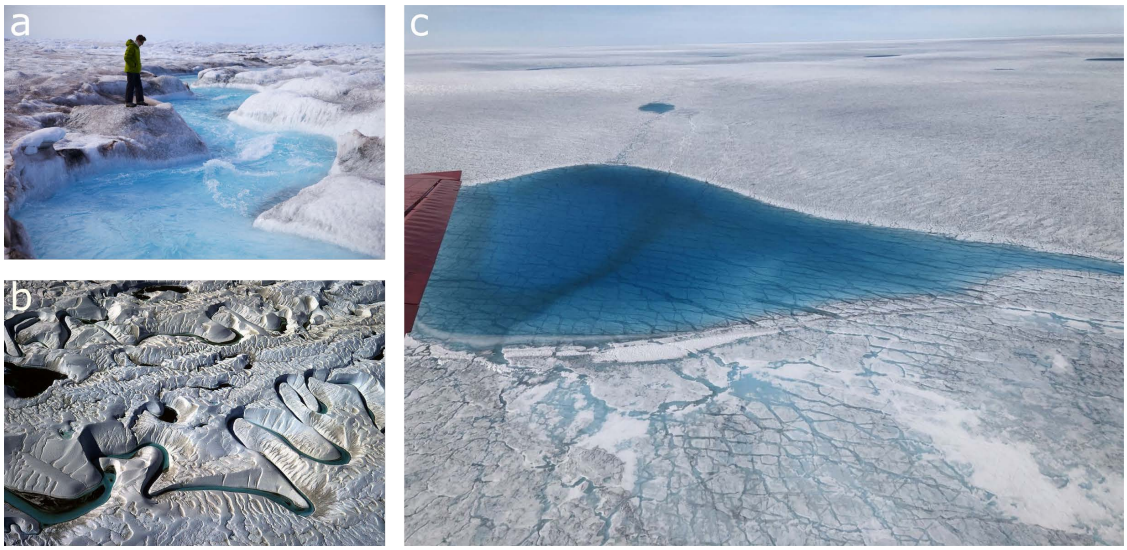


Figure 1.10: **Photos of the interconnected hydrological system. a) Person observing surface runoff in a supraglacial stream on GrIS (image credit: Sara Penrhyn-Jones), b) meandering supraglacial channels at Vibeke Gletscher in east Greenland (image credit: M. Hambrey), c) Supraglacial lake as seen from the onboard airborne Polar 5 survey (image credit: Angelika Humbert).**

Katabatic and föhn winds are associated with SGL activity on both ice sheets. Katabatic winds form in the cold and dry conditions of the interior of the ice sheet (Laffin, 2022), while föhn winds are created when cool, moist air is forced over a

mountain and releases latent heat and precipitates during ascent (Laffin, 2022). These winds reduce moisture and prevent cloud formation, leading to increased surface insolation and heating (Mioduszewski et al., 2016; Vihma, Tuovinen, and Savijärvi, 2011). The intensity and frequency of the föhn winds determine the extent and distribution of SGL and ponding in Antarctica (Datta et al., 2019; Laffin et al., 2021; Laffin et al., 2022; Laffin, 2022; Turton et al., 2020; Vihma, Tuovinen, and Savijärvi, 2011) and Greenland (Mattingly et al., 2021; Mioduszewski et al., 2016; M. R. van den Broeke, Duynkerke, and Oerlemans, 1994; Wheel, Christoffersen, and Mernild, 2020). Furthermore, winds drive snow erosion, which exposes blue ice areas and intensifies surface melting (Kingslake et al., 2017; Lenaerts et al., 2017). However, the regions with the highest surface melt do not always correspond to the areas with the highest extent and distribution of SGL activity (J. F. Arthur et al., 2020a).

On grounded ice, SGLs are more likely to occur in areas with lower elevations and shallower slopes on both ice sheets (Stokes et al., 2019; Turton et al., 2021). At higher elevations in Antarctica, proximity to blue ice and nunataks is a significant factor in the occurrence of SGLs (Stokes et al., 2019). In Greenland, SGL coverage starts near the margin and extends further inland during the melt season (Y.-L. Liang et al., 2012; McMillan et al., 2007; Sneed and Hamilton, 2007; Sundal et al., 2009; K. Yang et al., 2021). In Antarctica, meltwater features are often seen near the grounding line (Kingslake et al., 2017; Lenaerts et al., 2017; Stokes et al., 2019), while most SGLs are located on ice shelves downstream of the grounding line (Stokes et al., 2019).

When the infiltration of water is hindered by saturated firn or refrozen water in the englacial environment, meltwater begins to form ponds (K. E. Alley et al., 2018; J. Harper et al., 2012; B. Hubbard et al., 2016; Lenaerts et al., 2017). These lakes can vary in size, with extents ranging from tens of metres to 17 km^2 (Selmes, Murray, and James, 2011) and depths of 12 m (Box et al., 2012) in Greenland and up to 80 km

long (Kingslake et al., 2017), with an area of 71.5 km^2 (Stokes et al., 2019) and depths of up to 6.8 m in Antarctica (J. F. Arthur et al., 2020a). As more ice melts, these SGLs continue to expand until no more meltwater is generated, at which point they either freeze over or drain. This draining can occur over a period of days to weeks or months, as they slowly overflow their banks or channel away (Hoffman et al., 2011; Tedesco et al., 2013). Alternatively, they can be emptied in a matter of hours to days through hydrofracture, a process in which meltwater flows into fractures on the ice surface, enlarging them and creating a connection between the ice surface and the bed (Banwell, MacAyeal, and Sergienko, 2013; Das et al., 2008; Doyle et al., 2013; Lai et al., 2020; T. Scambos et al., 2009; Stevens et al., 2015; Tedesco et al., 2013). This connection can remain open for the duration of the melting season. In Greenland, it is estimated that between 28% and 45% of all SGLs drain rapidly (Fitzpatrick et al., 2014), and one lake drainage event can cause others to follow suit (Christoffersen et al., 2018). In Antarctica, lake drainage events are rare, but rapid lake drainage has been linked to the disintegration of the Larsen B ice shelf (A. A. Leeson et al., 2020; T. A. Scambos et al., 2004; T. Scambos et al., 2009). Lakes that do not drain can freeze (Dunmire et al., 2020; Langley et al., 2016; A. A. Leeson et al., 2020; Tuckett et al., 2019), or remain liquid in buried or subsurface lakes during winter (Koenig et al., 2015). It has been observed that subsurface lakes can last for multiple melt seasons (Lampkin et al., 2020), and can warm the englacial environment through the transfer of latent heat during winter (T. Phillips, Rajaram, and Steffen, 2010; T. Phillips et al., 2013).

1.2.1.2 Supraglacial channels

Supraglacial channels (Surpraglacial Channels (SGC)) are a system of rivers and streams that transport meltwater across the surface of the ice sheet (Figure 1.10). Rivers are generally larger than streams, as streams are usually tributaries of larger

rivers (Pitcher and L. C. Smith, 2019; L. C. Smith et al., 2015). The formation of SGCs is not fully understood (Irvine-Fynn et al., 2011; Mantelli, Camporeale, and Ridolfi, 2015), but is thought to be related to the rate of incision of the channel compared to the amount of surface melt, the production of meltwater, and the surface topography (Irvine-Fynn et al., 2011; Marston, 1983; A. W. Nolin and M. C. Payne, 2007). It is believed that the evolution of these channels follows three steps: incision, ablation, and meandering (Kostrzewski and Zwolinski, 1995). SGCs usually form in the direction of the ice flow (David Knighton, 1972; Hambrey, 1977), and along the steepest flow direction (Mantelli, Camporeale, and Ridolfi, 2015). Surface topography affects the flow of water, and many channels form in similar positions between the melt seasons (Hagen, Korsen, and Vatne, 1991). SGCs can extend for many kilometres before reaching the ocean, entering the englacial environment, or draining into supraglacial lakes (K. Yang and L. C. Smith, 2016). A recent study has suggested that in Greenland, and probably Antarctica under future warming, a slower type of hydrofracture could be caused by SGC (Chandler and A. Hubbard, 2023). This process involves supraglacial streams that intersect thin surface fractures and drain surface meltwater into the englacial environment. In Greenland, channels can be up to 55 km long (K. Yang et al., 2019a) and vary in width from 0.20 m (Gleason et al., 2020) to tens of metres (Lampkin and VanderBerg, 2014; K. Yang and L. C. Smith, 2016; K. Yang et al., 2019a). They carry meltwater to supraglacial lakes, moulins, or directly off the ice sheet. In Antarctica, supraglacial channels are less common but have been observed to flow up to 5 km in length (Kingslake et al., 2017), with a hydrological network spanning 120 km (Kingslake et al., 2017) and ending in waterfalls 130 m wide (Bell et al., 2017).

1.2.1.3 Supraglacial slush

In addition to meltwater stored as open water, meltwater can be stored in firn pore spaces (Dunmire et al., 2020; Montgomery et al., 2020). Slush is formed when firn pore spaces become saturated, especially in areas where the firn covers former blue ice areas, refrozen lakes, or where refrozen water has formed ice layers at depth within the firn (Dell et al., 2022). Melting and refreezing of the slush reduces the air content in the firn, increasing the density of the firn and increasing the vulnerability of the ice surface to ponding (K. E. Alley et al., 2018; B. Hubbard et al., 2016; Kuipers Munneke et al., 2014). Slush has been found to account for almost two thirds of the total meltwater area on an Antarctic ice shelf (Roi Baudouin Ice Shelf, 64%) in the period between 2013 and 2020 (Dell et al., 2022). In Greenland, slush is found to agree well with runoff limits mapped from higher resolution satellite data (Machguth, Tedstone, and Mattea, 2023; Tedstone and Machguth, 2022). In particular, slush fields have been reported to be a potential source of supraglacial channels (Pitcher and L. C. Smith, 2019).

1.2.2 Englacial hyrdology

Englacial describes the features of the ice sheet system that are located, occur, or form beneath the surface, inside the glacier. Englacial hydrology is concerned with meltwater that is present in the englacial environment through local thawing or draining of supraglacial meltwater. Meltwater that enters the englacial environment through permeable crevasses or moulins (Figure 1.11) helps to drain the meltwater to the bed (Das et al., 2008). Liquid water in the englacial system releases heat as it flows through porous ice, causing localised melting (Seguinot et al., 2020).

Figure 1.11: **Photos of features that facilitate the penetration of surface water into the englacial hydrological system. a) Curved crevasses on the Jakobshavn Glacier (image credit: University of Washington), b) supraglacial lake and water-filled crevasses near Illulissat, Greenland (image credit: James Balog/Extreme Ice Survey), c) A person descends into a moulin on the Greenland Ice Sheet (image credit: Christian Pondella/Red Bull Content Pool).**

1.2.2.1 Crevasses

Cracks or fractures in the ice sheet, known as crevasses, can range in size from millimetres to tens of metres (Holdsworth, 1969). The shape of the crevasse is determined by the stresses within and on the ice, which are affected by temperature, density, structure, and water content (van der Veen, 1998, 1999). At the glacier termini, crevasses can curve or rotate due to varying ice velocities (Figure 1.11b). Compressive forces cause crevasses to form near the surface of the ice sheet. However, during the melt season, meltwater entering crevasses can cause

the fracture to widen and deepen, a process known as hydrofracture. When water fills a crevasse to the point where the water pressure exceeds the fracture strength of the ice, vertical lake drainage occurs (K. E. Alley et al., 2018). The crevasse can spread through the entire thickness of the ice to the bed, forming a moulin through which the lake drains (Das et al., 2008; Hoffman et al., 2018; McGrath et al., 2012).

1.2.2.2 Moulins

Conduits known as moulins form when water drains vertically and are initiated by the ice hydrofracture process. These moulins can range in size from centimetres to tens of metres. Unlike crevasses, moulins can facilitate rapid drainage of supraglacial meltwater (Colgan et al., 2011; McGrath et al., 2011). Studies have found that crevasses and moulins can drain up to 86% of surface meltwater in West Greenland (Koziol et al., 2017). Furthermore, crevasse areas have been increasing in size (10% between 1985 and 2009 in West Greenland (Colgan et al., 2011), while crevassing on the Thwaites and Pine Island ice shelves has also increased in the past decade (Surawy-Stepney et al., 2023)) due to ice sheet thinning, which could lead to an increase in meltwater drainage. Meltwater that enters cracks, crevasses, and moulins on grounded ice drains into subglacial environments (McGrath et al., 2012; van der Veen, 2007).

1.2.3 Subglacial hydrology

The study of water beneath the ice sheet is known as subglacial hydrology. Due to the difficulty of accessing the bed for fieldwork and exploration, our knowledge of subglacial conditions is largely based on indirect observations and modelling. Subglacial meltwater is believed to be stored in lakes, refrozen, or transported along the bed.

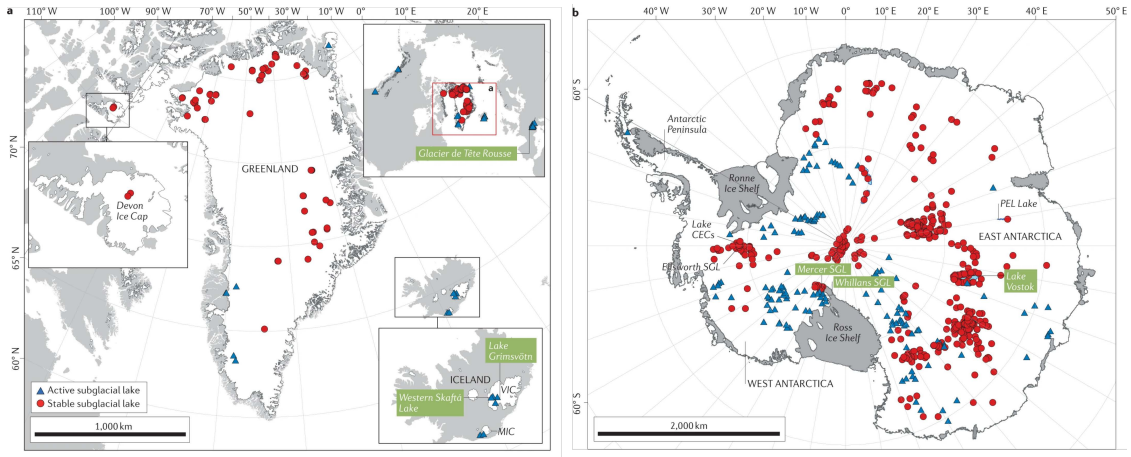


Figure 1.12: **Inventory of known subglacial lakes in a) Northern Hemisphere and Greenland and b) Antarctica (Livingstone et al., 2022).**

1.2.3.1 Basal conditions and water routing

Basal melting generates meltwater in addition to water that enters the subglacial environment through the drainage of Surpraglacial Hydrology Features (SGHF). This process is caused by frictional heat produced when ice slides over the bed, heat generated by supraglacial melt flowing into the subglacial environment, and geothermal heat flux from the Earth's core (Karlsson et al., 2021). These phenomena vary in different areas. In the interior of the ice sheet, the basal ice temperatures are predicted to be below the pressure melting point. In addition, there is little surface meltwater input to the bed due to lack of meltwater at the surface and ice thickness inhibits vertical lake drainage to the bed (MacGregor et al., 2016), while frozen bed conditions inhibit basal sliding (Eisen et al., 2020; Pattyn, 2010). However, around the edge of the ice sheet, where the ice is not as thick, the basal temperatures reach or are close to the pressure melting point, allowing basal melting (Dawson et al., 2022; Engelhardt, 2004; Fahnestock et al., 2001; Joughin et al., 2009; Karlsson et al., 2021).

At the surface, the flow of water is determined by the surface's slope and gravity. Under the immense pressure created by the mass of ice above it, the flow of subglacial water is regulated by gradients in water pressure and the topography of the bed (Fowler, 1987; Iken and Bindschadler, 1986; Schoof, 2010). If the thickness of the ice, and therefore the pressure, is high enough, subglacial water can go against gravity and flow uphill or parallel to the slope of the bed (Lliboutry, 1968). Subglacial meltwater can be drained through efficient channels or inefficient distributed systems, with meltwater spread over a large area of the bed (Davison et al., 2019).

The bed beneath both ice sheets is composed of bedrock and porous sediment regions (Booth et al., 2012; Davison et al., 2019; J. T. Harper et al., 2017; Kulessa et al., 2017; A. M. Smith et al., 2018). Inefficient drainage occurs through porous (Darcian) flow through sediment (Boulton and Jones, 1979), films (sheets of water) (Weertman, 1962), or linked-cavity systems (Kamb, 1987). Channelised drainage systems are more effective than distributed systems. It is suggested that channels can form along the main flow paths in the ablation zone; however, their formation and endurance may be impeded with increasing ice thickness and distance from the margin (Davison et al., 2019).

The summer melting of ice leads to an increase in subglacial water pressure, which alters the dynamics of ice flow. This can cause the formation of temporary channels and conduits that allow water to travel beneath the ice. On the contrary, during the winter months, temperatures drop significantly, resulting in less surface melting, drainage pathways close and the system is poorly connected (Bartholomew et al., 2012). This causes the subglacial water systems to become less active. Despite this, widespread wintertime subglacial water storage has been identified beneath Greenland (Chu et al., 2016).

1.2.3.2 Subglacial lakes

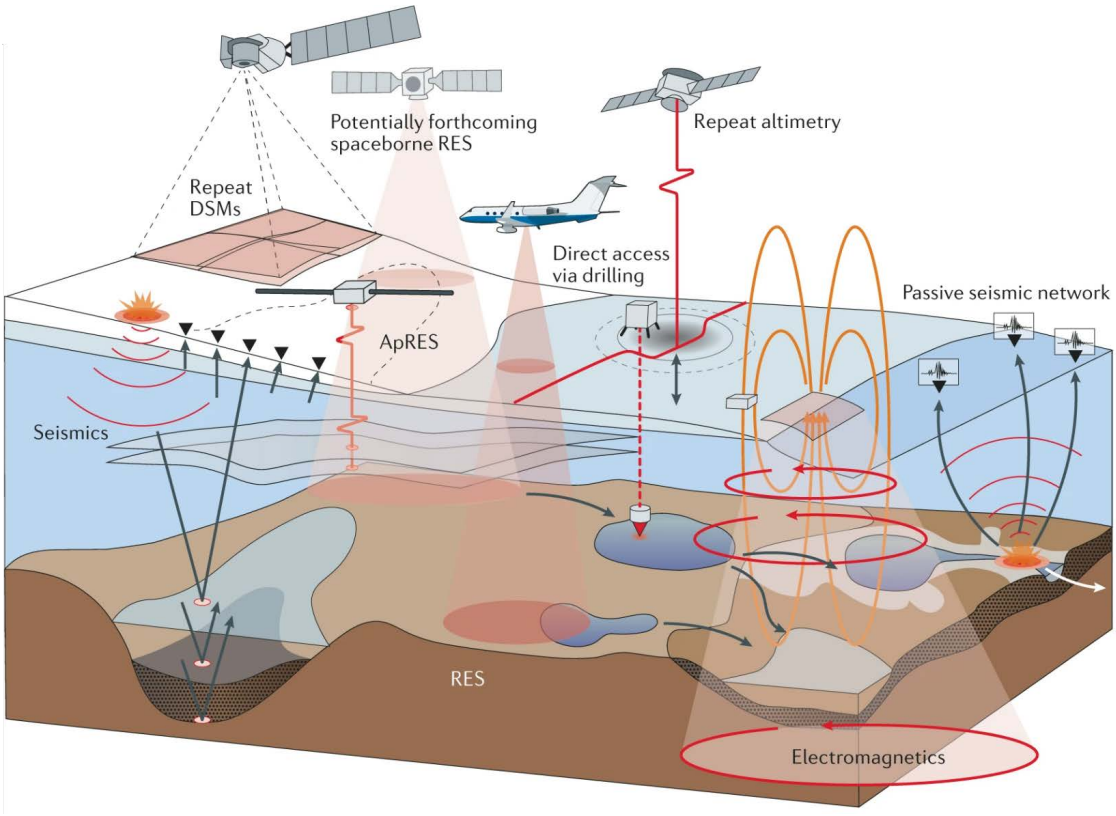


Figure 1.13: **Geophysical methods and satellites used to identify subglacial lakes, sample their environment, and monitor their dynamics (Livingstone et al., 2022).**

Subglacial lakes form when meltwater collects in depressions in the topography of the bed, gets trapped between the ice and the bed (Gilbert et al., 2012), or ice flows over regions of higher basal friction (known as sticky spots) (Sergienko and C. L. Hulbe, 2011). Subglacial lakes exhibit two distinct types of behaviour: active and stable. Active lakes drain along flow paths, forming connected networks, where drainage events in one lake can cause filling and subsequent drainage in adjacent lakes (Flament, Berthier, and Rémy, 2014; Fricker and T. Scambos, 2009). It has been observed that 80% of known lakes are stable, indicating that subglacial

systems, including these lakes, are closed or have balanced inflow and outflow (Livingstone et al., 2022). However, it is possible that we do not have enough data over time to confirm that they are active.

As accessing and exploring the subglacial environment directly is challenging, identifying and characterising lakes directly is infeasible. Techniques such as Radio-Echo Sounding (RES), swath radar technology, repeat altimetry, digital surface models, and direct access via drilling are used to sample lakes (Livingstone et al., 2022). Despite the fact that Greenland has a more extensive area of surface melting, Antarctica has ten times the amount of subglacial lakes (Livingstone et al., 2022). There are 773 known subglacial lakes below ice in the world (Figure 1.12), 675 of these lakes have been identified in Antarctica, 64 in Greenland, and an additional 34 under ice in other places (Livingstone et al., 2022).

1.3 Importance of supraglacial hydrology

Surpraglacial Hydrology (SGH) is thought to play a significant role in ice sheet mass balance through processes that affect supra, en, and subglacial environments (Figure 1.14). In the following sections, I provide a summary of the principal processes: increased runoff and surface thinning, decreased albedo and feedback loops, ice shelf collapse, the impact on basal conditions, and finally, cryohydrologic warming.

1.3.1 Increased runoff and surface thinning

In Antarctica, runoff and thinning are mainly observed on ice shelves, while surface melting is a common occurrence in most of the Greenlandic ice sheet (Bell et al., 2018). Meltwater stored in lakes, which does not drain to the sub- or englacial

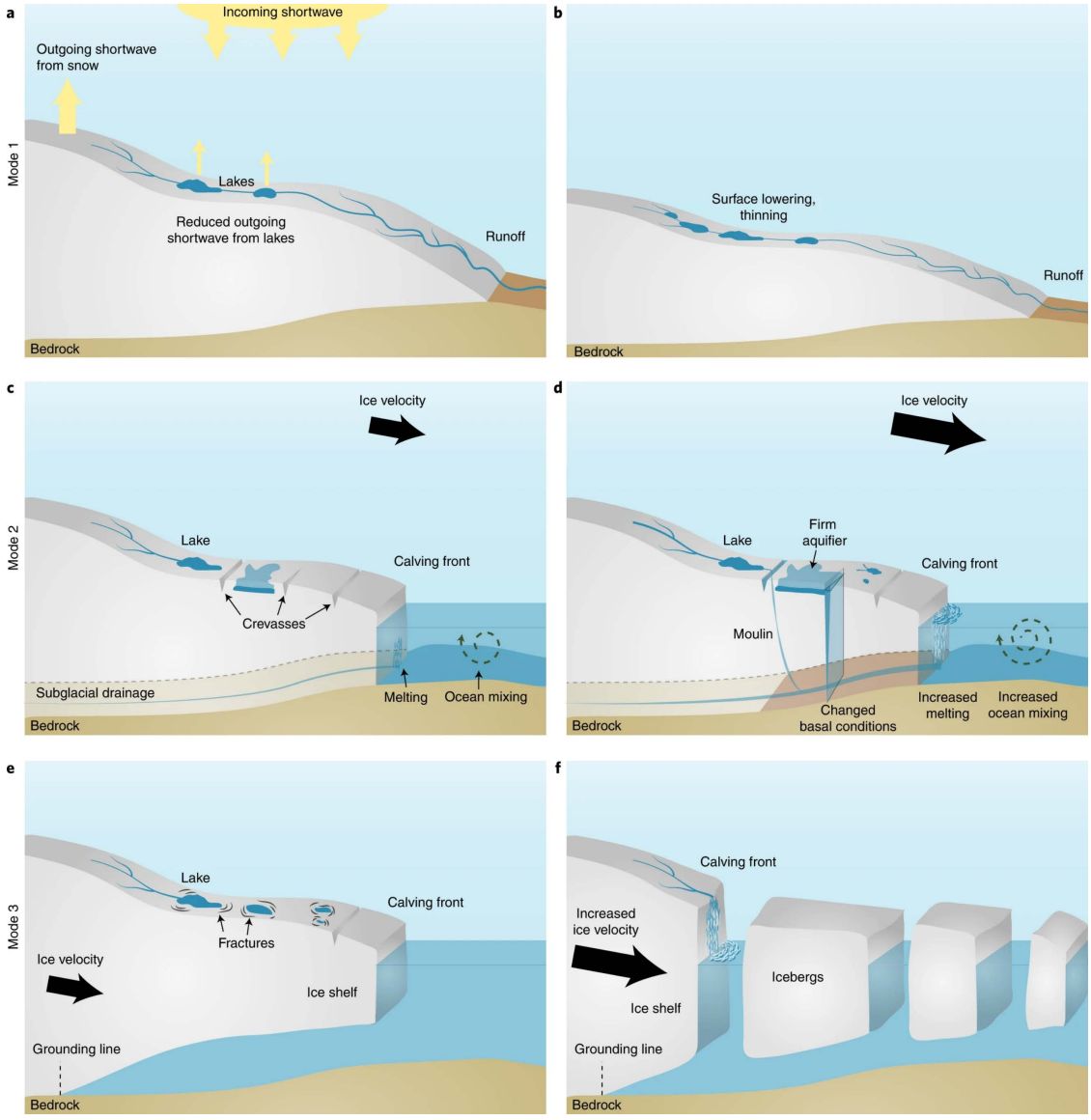


Figure 1.14: **Three modes of surface melt impact on ice-sheet mass balance: a, b) surface melt leading to direct surface runoff and thinning, c, d) injection of surface meltwater into the subglacial environment, e, f) meltwater-induced ice-shelf collapse (Bell et al., 2018).**

environments or refreeze, may be discharged directly from the ice through supra- and pro-glacial channels, thus contributing to global sea level rise. Prior to 2006, the

amount of mass lost from Greenland due to surface melting and runoff was equivalent to the mass lost from dynamical imbalance processes (M. van den Broeke et al., 2009). However, since then, the mass lost from surface melting and runoff has exceeded the mass lost from ice dynamics, with up to 84% of the total annual mass loss from the GrIS being attributed to these processes (Enderlin et al., 2014; M. van den Broeke et al., 2009). During the period 2003-2013, the southeast and northwest regions of Greenland accounted for 70% of the total mass loss (280±58 Gt/yr), while the 54% of the losses suffered in the southwest were due to a decrease in SMB (Velicogna, Sutterley, and M. R. van den Broeke, 2014). In Antarctica, most of the mass loss is caused by changes in ice dynamics (Rignot et al., 2008b, 2019; van de Berg et al., 2006; Velicogna, Sutterley, and M. R. van den Broeke, 2014). The Amundsen Sea sector and the Antarctic Peninsula are responsible for 64% and 17%, respectively, of the total loss (180±10 Gt/yr, Velicogna, Sutterley, and M. R. van den Broeke, 2014). Queen Maud Land, East Antarctica, is the only sector that has experienced a significant mass gain due to a local increase in SMB (63±5 Gt/yr, Velicogna, Sutterley, and M. R. van den Broeke, 2014). As temperatures and melting rates in Antarctica continue to rise, it is predicted that mass loss due to surface runoff and thinning will also increase (Bell et al., 2018).

1.3.2 Albedo & feedback loops

The albedo of a surface is the ratio of the amount of solar radiation that is reflected from it compared to the amount that is received. It is usually expressed as a value between 0 and 1 or as a percentage. The higher the albedo, the more solar radiation is reflected. White surfaces, such as snow and ice, reflect more radiation than darker surfaces, such as supraglacial meltwater, cryoconite, or blue ice (see Figure

1.15 in Bell et al. (2018), Leidman et al. (2021), Lenaerts et al. (2017), and Rennermalm et al. (2013)). Supraglacial hydrology decreases the higher albedo of the ice surface, increasing the absorption of the incoming solar radiation, and establishing a positive feedback loop that can enhance further melting (A. A. Leeson et al., 2015; Lüthje et al., 2006; Tedesco et al., 2012). Although the albedo decreases with increasing depth of the meltwater, the rate of decrease decreases rapidly with increasing depth (Lüthje et al., 2006). The fraction of radiation that is reflected or absorbed by supraglacial meltwater is bound by Beer's law (Equation 1.3, Philpot, 1989) which can be used to estimate the depth of supraglacial water from reflectance values (Banwell et al., 2014; Pope et al., 2016; Sneed and Hamilton, 2007; Tedesco and Steiner, 2011; Williamson et al., 2018).

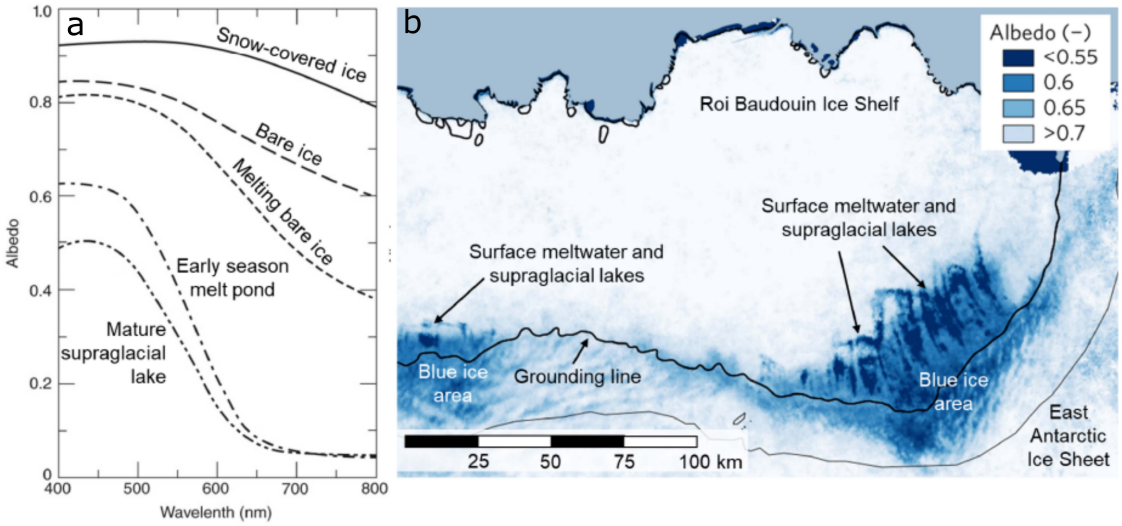


Figure 1.15: **a) Spectral albedo for supraglacial surfaces dependent on wavelength (Tedesco, 2015), b) Summer albedo derived from MODIS imagery (Lenaerts et al., 2017) (J. F. Arthur et al., 2020a).**

Snow-free (or blue) ice, exposed bedrock, and cryoconite (sediment found on ice sheet surfaces composed of mineral and biological material) can reduce the albedo

of the surface compared to surrounding snow and ice (A. A. Leeson et al., 2015; Leidman et al., 2021; Lüthje et al., 2006; Orheim and B. Lucchitta, 1990; Rennermalm et al., 2013; Tedesco et al., 2012). This can lead to local melting in areas where katabatic winds cause snow erosion, such as around the grounding line (Lenaerts et al., 2017), making Antarctic ice shelves vulnerable to increased melting (Winther et al., 1996).

1.3.3 Effect on ice shelves

The Antarctic ice sheet has been losing mass at an accelerated rate in recent years, mainly due to increased melting of ice shelves caused by the ocean and the dynamic response of the outlet glaciers (B. Smith et al., 2020). However, atmospheric warming has also had an effect on ice shelves, such as the disintegration of the Larsen B ice shelf in 2002, which has been associated with an increased presence of supraglacial lakes in the years before its collapse (Bell et al., 2018; A. A. Leeson et al., 2020). The refreezing of meltwater in the firn pores has been proposed to cause the collapse of ice shelves due to its effect on surface permeability and the ability of the water to form supraglacial lakes (Kuipers Munneke et al., 2014). In the months and years before their collapse, the Prince Gustav and Larsen A ice shelves were host to widespread SGL coverage (Cooper, 1997; A. A. Leeson et al., 2020). Meltwater flows laterally, forming channels and filling surface indentations and in some cases draining through the ice shelf. Alteration in local load due to movement and accumulation of water on the shelf can cause flexural stresses that are strong enough to create fractures. Meltwater can flow into and expand these fractures through hydrofracturing, allowing rapid drainage of SGLs (T. A. Scambos et al., 2000). Rapid drainage of supraglacial lakes has been suggested as a mechanism to weaken floating ice shelves to the point of collapse (K. E. Alley et al., 2018; Banwell,

MacAyeal, and Sergienko, 2013; Banwell et al., 2019; Glasser and T. A. Scambos, 2008; A. A. Leeson et al., 2020; T. Scambos, C. Hulbe, and Fahnestock, 2003; Trusel, Pan, and M. Moussavi, 2022).

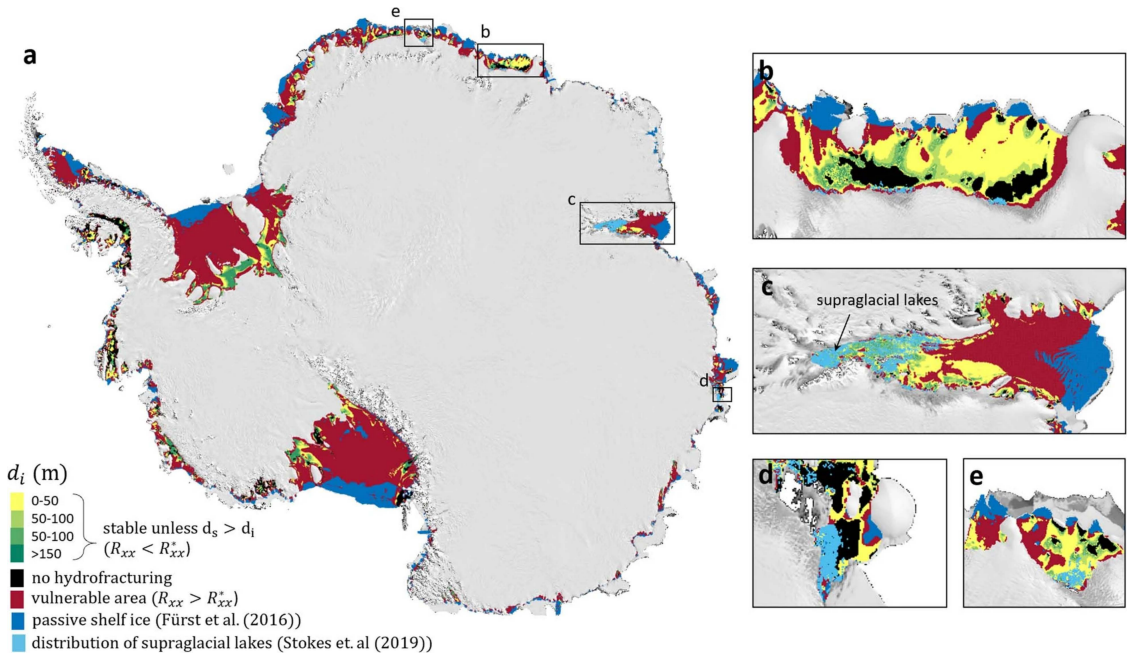


Figure 1.16: **SGLs mapped over East Antarctica during 2017 overlaid on a vulnerability map of Antarctica's ice shelves (Lai et al., 2020).**

Laterally confined ice shelves generate a resistive force that slows the flow of ice into the ocean, a process known as buttressing (Gudmundsson, 2013; van der Veen, 1998). When an ice shelf breaks up, the buttressing effect is reduced, leading to an increase in the discharge of ice from the upstream glaciers and a corresponding rise in sea level (De Angelis and Skvarca, 2003). For example, following the collapse of the Larsen B ice shelf in 2002, the Hektoria, Green, and Evans glaciers accelerated by up to 8 times their original velocity (Rignot et al., 2004). The stability of ice shelves across Antarctica varies due to different stress regimes, supraglacial lake coverage, and local climatological conditions (Figure 1.16, Lai et al., 2020). Many of

the ice shelves that usually experience widespread coverage of the surface glacial lake (SGL) are resistant to hydrofracture due to low pressures. However, 60 ± 10 percent of the ice shelves (in terms of area) buttress upstream ice and are susceptible to hydrofracture if they are flooded with water (Lai et al., 2020). Surface meltwater is present on many of Antarctica's ice shelves (J. F. Arthur et al., 2020b; Dell et al., 2022; M. Dirscherl et al., 2020; Stokes et al., 2019), and although SGH has been systematically mapped on the EAIS, no dedicated lake inventory for the entire WAIS and AP exists.

1.3.4 Effect on basal conditions

The drainage of SGL features through hydrofracture is not only limited to floating ice shelves. Although it is uncertain if this process occurs on grounded ice in Antarctica, it has been extensively documented in Greenland (K. E. Alley et al., 2018; R. B. Alley et al., 2005; van der Veen, 2007). This drainage has a noticeable effect on the basal conditions of glaciers, with rapid delivery of surface water to the bed reducing basal friction and temporarily increasing ice flow velocities by up to an order of magnitude (Tedesco et al., 2013). The correlation between surface melting and ice velocity is largely determined by the structure and efficiency of the subglacial drainage system, which changes over time and space (Bartholomew et al., 2010). Therefore, the presence of surface meltwater can cause seasonal variations in the motion of the ice due to positive feedback between surface melting and ice velocities (Parizek and R. B. Alley, 2004; Shepherd et al., 2009). A more efficient drainage system can reduce the basal lubrication effect of external meltwater inputs by allowing larger discharges in discrete channels (Sundal et al., 2011). Ice acceleration depends on the amount of water that enters the system, as a large flow can exceed the hydrological capacity, resulting in an increase in velocity

as surface melting begins to increase at the beginning of the season (Bartholomew et al., 2011). As the rate of surface melting stabilises, the acceleration stabilises as well. The drainage system, which transports water from the ice surface to the edge of the ice sheet via the bed, develops and expands during the melting season (Bartholomew et al., 2011), allowing ice sheet accelerations of 35% per positive degree day of melting during late summer (Shepherd et al., 2009).

Supraglacial lakes have been observed to drain through 1000 m thick ice, causing seismicity, transient acceleration, ice sheet uplift, and horizontal displacement, while subsidence and deceleration occurred during the subsequent 24 hours (Das et al., 2008). It has been suggested that lake drainage events in Greenland are preceded by hourly periods of ice sheet uplift and increased basal slip (Stevens et al., 2015). This is believed to be due to the injection of meltwater into the bed through neighbouring moulin systems (Stevens et al., 2015). Similar mechanisms have been hypothesised to occur in East Antarctica (Langley et al., 2016). A 2019 study has provided evidence that five glaciers on the Antarctic Peninsula (Drygalski, Hektoria, Jorum, Crane and Cayley) experienced near-simultaneous speed-up events in March 2017, November 2017 and March 2018 (Tuckett et al., 2019), suggesting that surface meltwater may have entered the subglacial hydrological system. Injection of water into the subglacial environment may lead to increased ocean melting at the calving fronts of both ice sheets (Bell et al., 2018; Choi et al., 2021) and in ice shelf cavities in Antarctica (Jenkins, Nicholls, and H. F. J. Corr, 2010; Stewart et al., 2019). In addition to the impact of subglacial water on modulating ice flow, water discharged into marine settings has the potential to influence ocean circulation, biogeochemical cycles, ice-ocean interactions, and advection of geothermal heat flux (Gooch, Young, and Blankenship, 2016; Liljedahl et al., 2021).

1.3.5 Cryohydrologic warming

Cryohydrologic warming is a process in which latent heat from meltwater is transferred to the environment when it refreezes during winter (T. Phillips, Rajaram, and Steffen, 2010; T. Phillips et al., 2013). This phenomenon can alter the rheology of ice in crevasses and firn aquifers (B. Hubbard et al., 2016). A study of the Larsen C Ice Shelf found that the englacial layer (firn) was $\sim$ 10 °C warmer and $\sim$170 kg m^{-3} denser than expected, due to intense melting and intermittent ponding (B. Hubbard et al., 2016). This suggests that cryohydrologic warming not only releases latent heat, but also changes the density of ice columns (B. Hubbard et al., 2016). Therefore, it is important to understand the distribution of supraglacial hydrology to assess its effect on the mass balance of the ice sheets.

1.4 Known distribution & extent of supraglacial hydrology

Existing inventories, systematic surveys, and datasets of Antarctic and Greenlandic supraglacial hydrology are rare. Extensive studies which do exist examine SGHF either on a small spatial but long temporal scale or a snapshot covering a large area. Here, I provide an overview of the data sets and research papers that have been published about SGH in both AIS and GrIS.

1.4.1 Antarctica

There is no complete survey of supraglacial hydrology extent for the AIS. However, a comprehensive mapping of the East Antarctic Ice Sheet revealed more than 65000 supraglacial lakes in east Antarctica (Stokes et al., 2019) in January 2017 (Figure 1.17). These lakes covered a total area greater than 1300 km^2 and were found

mainly in coastal regions, at elevations less than 100 m above sea level, and on surface slopes less than 1°. Although this study provided a comprehensive snapshot of east Antarctic lake distribution, it does not reveal how these lakes have evolved over time; both in terms of their long-term (decadal scale) evolution and their short-term (seasonal) variability. Evidence of the long-term endurance of SGLs and surface streams has been documented on the Shackleton Ice Shelf since 1947 (Kingslake et al., 2017). Additionally, SGH activity has been observed sporadically over a decade-long period across the EAIS (Figure 1.17, Bell et al., 2017; Kingslake et al., 2017). Unfortunately, there are no data to indicate whether the occurrence, area, or volume of these features is increasing or decreasing.

Lakes have commonly been identified on the ice shelves of the Antarctic Peninsula, including George VI, Larsen B, Larsen C, and Wilkins (J. F. Arthur et al., 2020a; Glasser and T. A. Scambos, 2008; Kingslake et al., 2017; LaBarbera and MacAyeal, 2011; A. A. Leeson et al., 2020; T. Scambos, C. Hulbe, and Fahnestock, 2003; D. G. Vaughan et al., 1993). However, to date, no systematic study of supraglacial lake evolution over time has been performed on either the EAIS or AP. Furthermore, even less is known about the widespread distribution, density, and seasonal evolution of supraglacial lakes and channels on the WAIS. Studies have demonstrated that SGLs form on the Ford Ranges and Pine Island Glacier (Figure 1.17, J. F. Arthur et al., 2020a; Kingslake et al., 2017) and are involved in the collapse of Larsen-B (A. A. Leeson et al., 2020). A recent study revealed that the duration and extent of surface melt on the northern George VI Ice Shelf during the austral summer of 2019 / 2020 were extraordinary compared to the 31 preceding summers of significantly lower melt. Interestingly, SGH activity is visible from satellite imagery on Sulzberger, Nickerson, and Dotson ice shelves (J. F. Arthur et al., 2020a), yet no comprehensive survey has been conducted to date. Supraglacial channels have been observed crossing the grounding line from Shackleton Glacier onto the

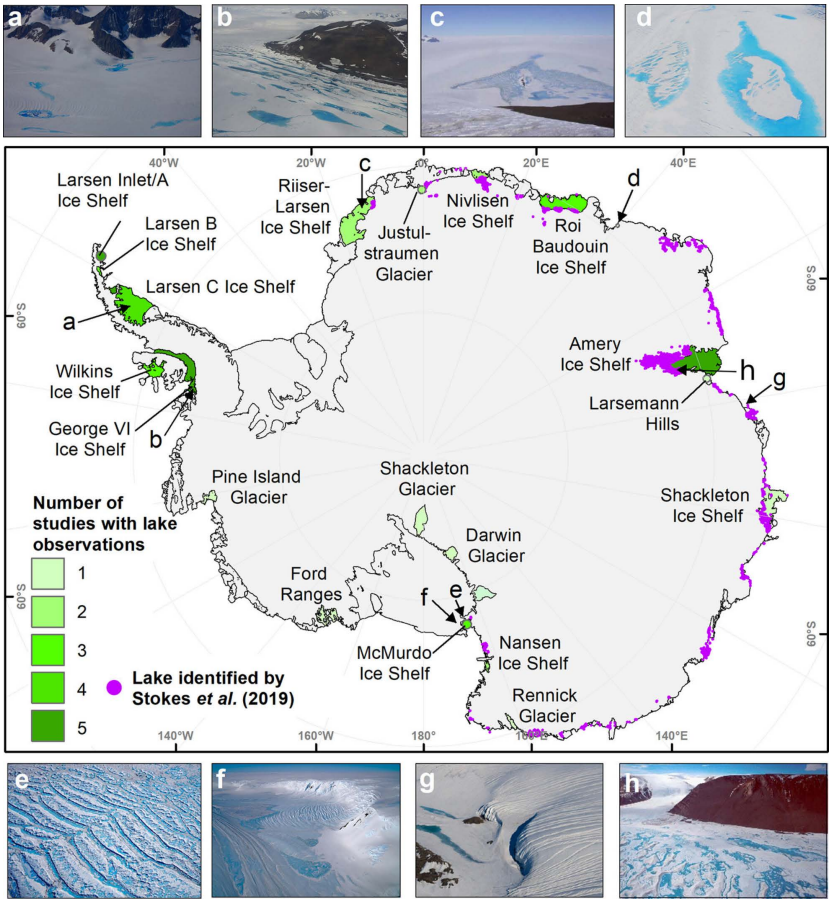


Figure 1.17: **Locations around Antarctica where supraglacial lakes have been observed, together with examples: (a) Larsen C Ice Shelf, (b) George VI Ice Shelf, (c) Riiser-Larsen Ice Shelf, (d) Langhovde Glacier, (e) Ross Archipelago, (f) McMurdo Ice Shelf, (g) Sørsdal Glacier, (h) Mawson Glacier. The locations of more than 65000 lakes identified on the EAIS are pink (J. F. Arthur et al., 2020a; Stokes et al., 2019). Image credits: Martin Truffer (a), Frithjof C. Küpper (b), Matti Leppäranta (c), Takehiro Fukuda (d), NASA Operation IceBridge (e), Chris Larsen, NASA Operation IceBridge (f), Sarah Thompson (g), and Richard Stanaway (h)).**

Ross Ice Shelf, the Amery Ice Shelf, Shackleton Ice Shelf and various other locations on the EAIS (Bell et al., 2018; Kingslake et al., 2017).

1.4.2 Greenland

In Greenland, most of the studies of SGH to date have generally aimed to develop our process-level understanding of the seasonal evolution of supraglacial lakes and have focused on one or more study sites associated with specific glaciological basins (Y.-L. Liang et al., 2012; McMillan et al., 2007; Sundal et al., 2009; Williamson et al., 2018). Small-scale studies detail the existence of lakes on a wide spatial scale, for example, on the Ryder Glacier (Otto, Holmes, and Kirchner, 2022) and in the northwest (Gledhill and Williamson, 2018), northeast (Hochreuther et al., 2021; Schröder et al., 2020; Turton et al., 2020), southeast (Jiang et al., 2022; Stevens et al., 2022), southwest (K. Yang et al., 2021), and west (Chudley et al., 2019; Yuan et al., 2020) of the ice sheet. Research carried out on a larger scale showed that lakes were draining rapidly, with considerable variations in the presence of fast-draining lakes in the region. Furthermore, a correlation was observed between sudden lake drainage and dynamic mass loss (Selmes, Murray, and James, 2011) and is the most recent study that provides seasonal analysis (2009 melt season). However, the sensor used had a low resolution, which could have caused it to overlook parts of the SGH system, and the ice sheet has experienced a series of high melting years since then, such as 2011, 2012 and 2019 (Otosaka et al., 2022). An examination of the 2017 melt season revealed 4530 ice marginal lakes that form at the fringes of glaciers and ice sheets where flow is impeded or limited (How et al., 2021). Further research has investigated the yearly and long-term changes of SGH on the GrIS, discovering SGLs in new parts of the northeast sector of the ice sheet and an expansion further inland to higher altitudes (Howat et al., 2013; Igneczi et al., 2016), but did not conduct a comprehensive study across the entire ice sheet. A systematic study of supraglacial lakes has been carried out on an interannual timescale covering the entire Greenland ice sheet during the period 2016-2018 (Figure 1.18, Hu et al., 2022).

This research revealed that SGH activity occurs around most of the edge of the ice sheet, and the majority of SGLs located in the southwest, northeast, and west of the ice sheet (Hu et al., 2022).

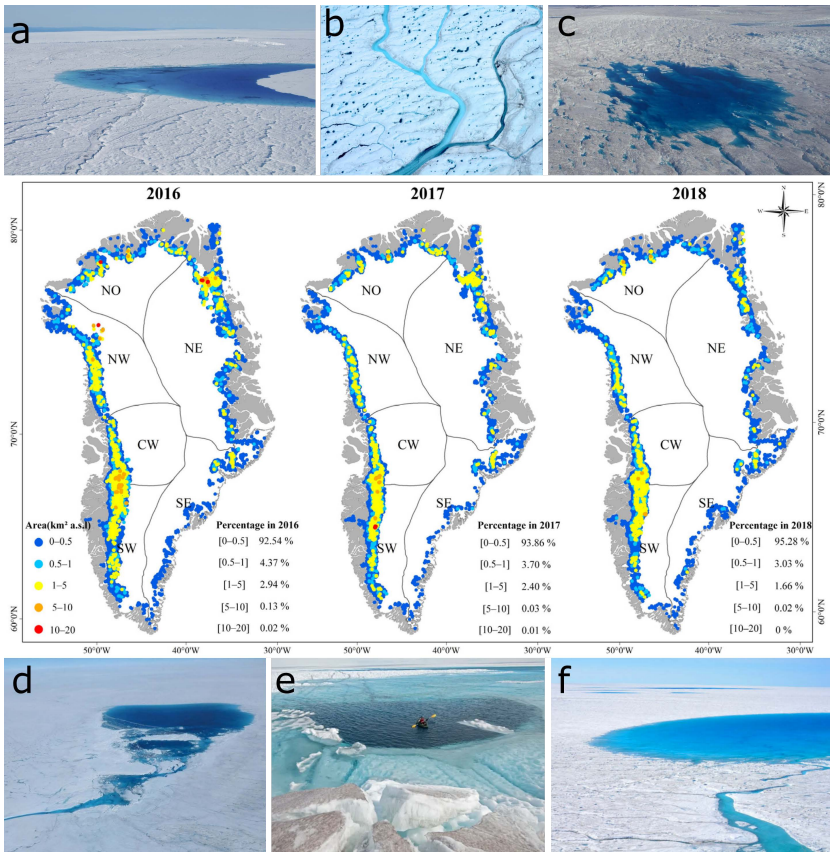


Figure 1.18: **Temporal and spatial distribution of SGLs in Greenland during the melt seasons from 2016 to 2018. The features are separated by area. Adapted from (Hu et al., 2022). Examples of supraglacial lakes and channels on GrIS, Image credits: a), b), f) Laura Stevens, c) Niklas Neckel, d) Marco Tedesco, and e) James Balog, National Geographic.**

Detailed studies reveal supraglacial channels on the GrIS, for example, in the Isortoq and Watson River basins (K. Yang et al., 2021), northwest Greenland (K. Yang et al., 2019a), and southwest Greenland (Gleason et al., 2016; Gleason et al., 2020;

Leidman et al., 2021, 2023; L. C. Smith et al., 2015). However, these studies are based on very high resolution commercial data, which can resolve to sub-metre resolution but are limited by data coverage and availability at the ice sheet scale, or concern only small study sites.

1.5 Remote sensing data

The practise of detecting and monitoring physical characteristics by collecting reflected and/or emitted radiation from satellites or aircraft is known as remote sensing. This field is usually divided into active (a signal sent out from the monitoring device which is reflected by the target object and detected by the sensor, e.g. radar) and passive (the reflection of sunlight detected by the sensor, e.g. optical imagery) remote sensing. To begin with, I explain the various data sources that have been used in the literature to map supraglacial hydrology. In Section 1.6, I examine the different techniques that have been used to identify supraglacial hydrology. In this review, I will be concentrating solely on satellite data, as other data sources, such as field studies, are not pertinent to the techniques I am using in my thesis.

1.5.1 Optical & multispectral imagery

Multispectral satellite imagery is advantageous for mapping ice sheet surface hydrology because open water features, such as lakes and channels, have a distinct spectral signature in the available wavelengths. This makes it possible to identify supraglacial lakes on the Greenland and Antarctic Ice Sheets (Figure 1.19). Studies of supraglacial channels are less common and are usually limited to regional-scale studies due to the need for very high-resolution data (e.g., L. C. Smith et al., 2015).

Additionally, multispectral imagery can be used to estimate lake depth (Section 1.6.4) by applying a radiative transfer model (Philpot, 1989). However, the main limitation of multispectral imagery is its sensitivity to clouds, which can obscure the surface when present. This is especially problematic around ice sheet margins, which are often cloudy due to orographic effects, and can significantly reduce the temporal sampling of optically derived feature datasets, such as supraglacial lake inventories (A. A. Leeson et al., 2013).

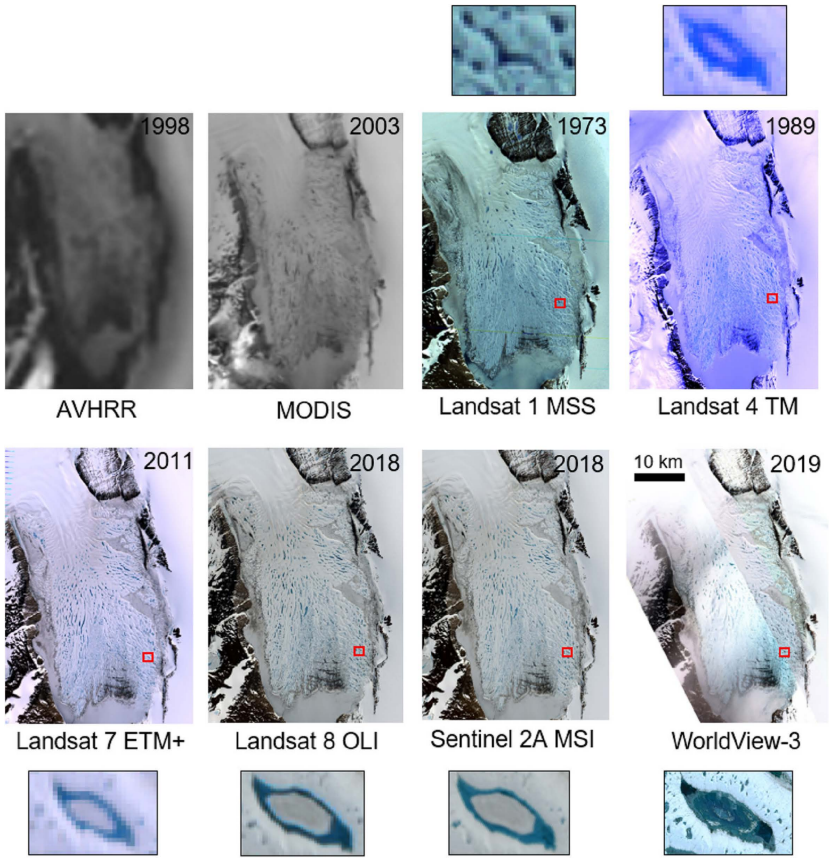


Figure 1.19: **Examples of the evolution of satellite image sensor resolution and detection of supraglacial lakes on Beaver Lake, adjacent to Amery Ice Shelf, East Antarctica (J. F. Arthur et al., 2020a).**

1.5.1.1 Coarse resolution sensors

Coarse resolution sensors are those with a spatial resolution of less than 250 metres. The Moderate Resolution Imaging Spectroradiometer (MODIS) sensors on board the Aqua and Terra satellites have a spatial resolution ranging from 250 to 1000 metres and a temporal resolution of 1 to 2 days (Table 1.2). This imagery has been used to map supraglacial lakes on the GrIS and AIS in many studies (Box and Ski, 2007; Fitzpatrick et al., 2014; B. Hubbard et al., 2016; Johansson and Brown, 2012; A. A. Leeson et al., 2012; A. A. Leeson et al., 2013; Lenaerts et al., 2017; Y.-L. Liang et al., 2012; Macdonald et al., 2019; Selmes, Murray, and James, 2011; Spergel et al., 2021; Sundal et al., 2009; Williamson et al., 2017). Additionally, Advanced Very High Resolution Radiometer (AVHRR) (Table 1.2), a multispectral scanner with a spatial resolution of 1 km and sub-daily repeat coverage at nadir, has been used to identify supraglacial lakes (Greuell, Reijmer, and Oerlemans, 2002; T. A. Scambos et al., 2000; Steffen, Abdalati, and Stroeve, 1993). These sensors are beneficial for studies requiring high temporal frequency, but their relatively poor spatial resolution makes it difficult to resolve features smaller than $<$0.062 km^2 for MODIS and $<$1.18 km^2 for AVHRR (Figure 1.19, J. F. Arthur et al., 2020a). To gain a more comprehensive understanding of SGH on Earth's ice sheets, higher spatial resolution is needed.

1.5.1.2 Medium resolution sensors

Numerous studies have evaluated the activity of SGHF using sensors of medium resolution. I consider medium resolution sensors to have a spatial resolution of between 10 and 250 meters. In early studies, SGLs were identified from Landsat-1, 2, 3, 4, and 5 (60-100 meters); however, their area could not always be precisely determined (Table 1.2, Orheim and B. K. Lucchitta, 1987; Winther et al., 1996). The

Table 1.2: **Optical satellite sensors and their respective launch dates, end of mission dates, revisit time, and spatial resolution.**

Sensor	Launch Date	End Date	Revisit Time [days]	Spatial Resolution [m]
MODIS Aqua	05/2002	Ongoing	1-2	250-1000
MODIS Terra	12/1999	Ongoing	1-2	250-1000
AVHRR	01/1970	Ongoing	0.5	1100
Landsat 1	07/1972	01/1983	18	60
Landsat 2	01/1975	02/1982	18	60
Landsat 3	03/1978	03/1983	18	60
Landsat 4	07/1982	06/2013	16	30
Landsat 5	03/1984	06/2013	16	30
Landsat 6	10/1993	Launch Failure	-	-
Landsat 7	04/1999	09/2021	16	30
Landsat 8	02/2013	Ongoing	16	30
Landsat 9	09/2021	Ongoing	16	30
Sentinel-2A	06/2015	Ongoing	5	10-60
Sentinel-2B	03/2017	Ongoing	5	10-60
ASTER	12/1999	Ongoing	4-16	15-90
WorldView-2	10/2009	Ongoing	1.1	0.46-1.8
IKONOS	09/1999	Ongoing	1-3	0.82-3.2
PlanetScope	01/2013	Ongoing	1-3	3

launches of Landsat-7, 8, and 9 (with resolutions of 15 to 30 metres) have enabled the repeat monitoring of supraglacial lakes for small-scale studies (e.g. Bell et al., 2017; Dell et al., 2022; Gledhill and Williamson, 2018; Halberstadt et al., 2020; Kingslake et al., 2017; Langley et al., 2016; K. E. Miles et al., 2017; M. Moussavi et al., 2020; Pope et al., 2016; Stokes et al., 2019; Williamson et al., 2018 and many others). The two Landsat missions that are still in operation (Landsat-8 and Landsat-9) are designed to provide 8-day repeat coverage of any part of the world (Table 1.2). They are equipped with 8 multispectral bands (with a resolution of 30

metres), a panchromatic band (15 metres) and 2 Thermal Infra-Red Sensor (TIRS) bands (100 metres).

The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) sensor, which is on board the Terra satellite, has 14 multispectral bands with a resolution of 15-90 m and covers the visible to thermal infrared range (Table 1.2). It has a revisit time of 1 to 2 days and has been used to observe the seasonal changes and drainage events of lakes in East Antarctica (Langley et al., 2016), Jakobshavn Isbrae, Greenland (Georgiou et al., 2009), and Larsen-B, Antarctic Peninsula (Glasser and T. A. Scambos, 2008). National Aeronautics and Space Administration (NASA) made ASTER data publicly available in 2016, however, all Short-Wave Infra-Red (SWIR) data collected by the sensor since 2008 have been deemed unusable.

The European Space Agency (ESA) has launched two twin sensors, Sentinel-2A and 2B, which offer a higher spatial resolution of 10 metres for red, green, blue, and near-infrared (NIR) bands, with a 5-day revisit period (Table 1.2). Additionally, they contain nine bands spanning the visible to short-wave infrared (SWIR) ranges, with a resolution of 20 to 60 metres. Similar to the Landsat missions, Sentinel-2 has been used to effectively estimate SGH activity on both ice sheets (M. Dirscherl et al., 2020; M. C. Dirscherl, Dietz, and Kuenzer, 2021; Hu et al., 2022; Stokes et al., 2019; Williamson et al., 2018). The higher spatial resolution of Sentinel-2 allows for the delineation of supraglacial channels (SGCs) (Y. Li et al., 2022; L. C. Smith et al., 2015, 2021; K. Yang et al., 2014; K. Yang et al., 2019a, 2021).

In the past eight years, Landsat-8 and Sentinel-2 have been combined to increase the number of cloud-free scenes that can be used to map SGH activity (Kingslake et al., 2017; A. A. Leeson et al., 2015; K. E. Miles et al., 2017; M. Moussavi et al., 2020; Stokes et al., 2019). The combination of sensors with different spatial and radiometric resolutions can be difficult, as it can lead to geometric misalignments

and scaling issues. Furthermore, Sentinel-2 captures images in 12-bit radiometric resolution, while Landsat-8 uses 16-bit, so merging these datasets requires radiometric consistency to prevent information loss and maintain data quality. Additionally, it can be difficult to reconcile the temporal differences in data acquisition between the two sensors, so caution should be taken when directly comparing the results obtained from each sensor. These sensors, which are free and easy to access, are a great asset for polar science; however, there are also higher-resolution commercial sensors available.

1.5.1.3 Fine resolution sensors

Commercial satellite imagery provides much finer spatial and temporal resolution data products than ever before (Table 1.2). SGHF have been identified in Worldview-2 (1.84 m resolution, 1.1 day revisit period), IKONOS (3.28 m, 3 day revisit period) and PlanetScope satellite constellation imagery (0.8-5 m resolution, daily revisit period), as reported in (Bolch et al., 2008; Datta and Wouters, 2021; Jawak and Luis, 2014; Juen et al., 2014; M. S. Moussavi et al., 2016; Pope et al., 2016; K. Yang and L. C. Smith, 2013; K. Yang et al., 2017). Despite the improved spatial and temporal resolution, the cost and data availability of these sensors are restrictive, limiting their use in large-scale studies, and thus they have only been used in small-scale studies.

1.5.2 Radar

Radiation emitted by Radio Detection and Ranging (RADAR) devices is less affected by clouds than optical images, thus providing a way to map surface hydrology even in cloudy weather, winter, and when the solar zenith angle reduces the spectral distinctions between SGLs and their environment.

1.5.2.1 Synthetic Aperture Radar

Synthetic Aperture Radar (SAR) is a type of radar device used to detect lakes in imagery. Unlike optical approaches, SAR relies on the backscatter signature of the lake being different from the surrounding ice. The signal is more complex in SAR, since the backscattered signal of the ice is more variable (M. Dirscherl et al., 2021; Jiang et al., 2022), while lakes always appear darker than the surrounding ice in optical imagery. In Antarctica, studies have found that lakes appear dark against bright in winter and bright against dark in summer, with a period in autumn and spring where the signal is indistinct (A. A. Leeson et al., 2020). In Greenland, the signature of lakes is more uncertain; studies suggest that in the percolation zone in southwest Greenland, liquid water appears dark against bright, and fully refrozen lakes appear bright against dark (K. E. Miles et al., 2017).

In the past, ERS-1/2, Radarsat, and Envisat SAR have been employed to detect supraglacial lakes on both ice sheets (Fricker et al., 2002; Johansson and Brown, 2012; B. K. Lucchitta and Rosanova, 1998; Luckman et al., 2014; H. A. Phillips, 1998; Rott, Skvarca, and Nagler, 1996). More recently, Sentinel-1 has been used for this purpose (M. Dirscherl et al., 2021; Jiang et al., 2022; K. E. Miles et al., 2017; Schröder et al., 2020). However, SAR imagery is only useful for mapping areas and not depth due to the strong attenuation of microwave signals in water, which prevents the radar from penetrating the water surface and measuring the depth. Additionally, due to the penetration of the microwave signal into the snow, SAR imagery can be used to map subsurface and refrozen supraglacial lakes (Benedek and Willis, 2021; Koenig et al., 2015; K. E. Miles et al., 2017; Schröder et al., 2020), allowing for the examination of their wintertime development (Figure 1.20).

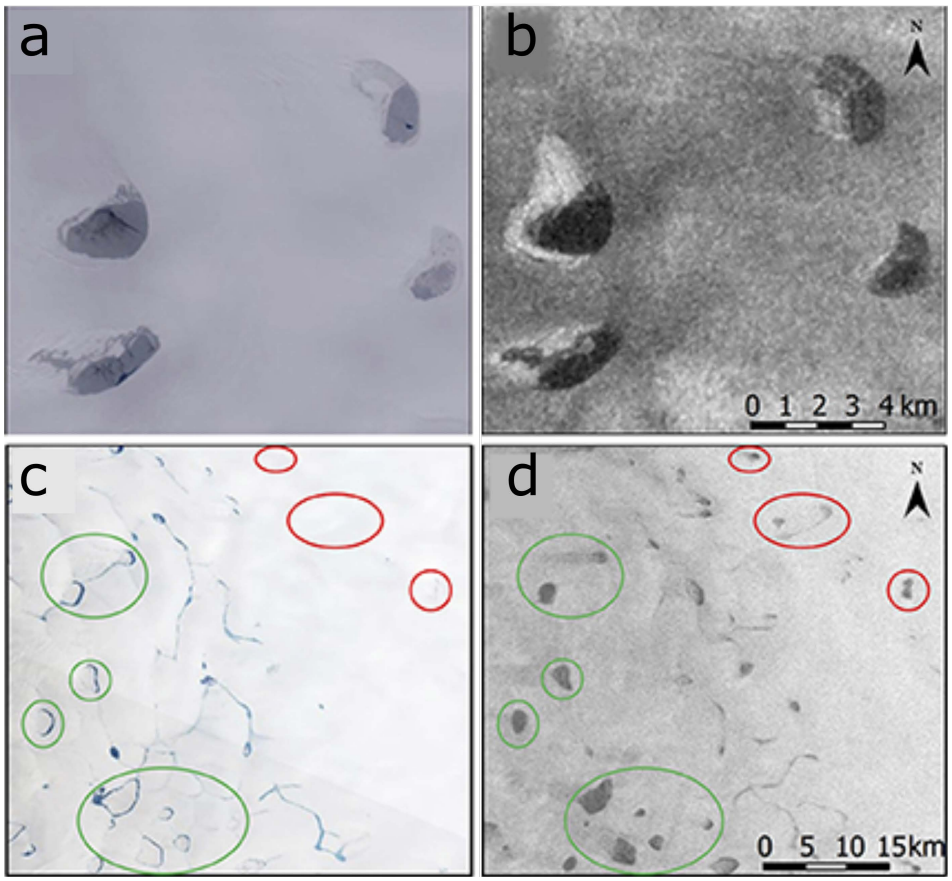


Figure 1.20: **Supraglacial lake examples in Landsat-8 RGB imagery (a, c) and Sentinel-1 HH polarisation (b, d). a) and b) show examples of partially refrozen lakes on days 235 and 232 of the 2015 melt season, while c) and d) reveal subsurface lakes on day 212 and, much later, day 239 of the 2015 melt season (K. E. Miles et al., 2017).**

1.5.2.2 Lidar and radar altimetry

Laser Imaging, Detection, and Ranging (LIDAR) is a technique for measuring the distance between a sensor and an object by timing the reflection of laser pulses from the object's surface to the receiver. Water surfaces reflect light horizontally and in a highly concentrated manner, which can be used to accurately determine the area of

supraglacial lakes. Additionally, some photons can penetrate shallow lakes (or ice and snow) to estimate the depth of the lake (Figure 1.21) and bathymetry (Fair et al., 2020; Studinger et al., 2022) after accounting for refraction in the water column (to detect buried lakes Koenig et al., 2015). SGHF have been identified using multiple lidar sources, such as NASA's Airborne Topographic Mapper from Operation IceBridge, ICESat-2, and other lidar missions (Datta and Wouters, 2021; Fair et al., 2020; Georgiou et al., 2009; Koenig et al., 2015; McMillan et al., 2007; Shean et al., 2019; Studinger et al., 2022). Radar altimeters work similarly to laser altimeters, but emit radio waves instead of laser pulses. For example, CryoSat-2 has been used to measure the height of supraglacial lakes in western Greenland (Gray et al., 2016).

1.6 Methods for mapping supraglacial hydrology

The most reliable way to map the lake region using satellite or airborne imagery is through manual digitisation (A. A. Leeson et al., 2013). However, this is a laborious process and is not suitable for large-scale monitoring. Automated approaches are much more efficient but tend to produce errors in the delineation of lakes. To ensure accuracy while mapping on a large scale, automated techniques are often combined with manual post-processing. Recent advances in applying machine learning algorithms to satellite imagery have enabled automated approaches to classify supraglacial hydrology without the need for laborious post-processing. Mapping supraglacial channels has been done using processes similar to those used for lake detection. However, accurately delineating these features is difficult due to narrow channels, varying contrast between meltwater and surrounding ice or snow, and complex drainage patterns (McGrath et al., 2011; L. C. Smith et al., 2017; K. Yang and L. C. Smith, 2013; K. Yang et al., 2017). Various techniques have been developed for mapping supraglacial hydrology with non-optical remote sensors.

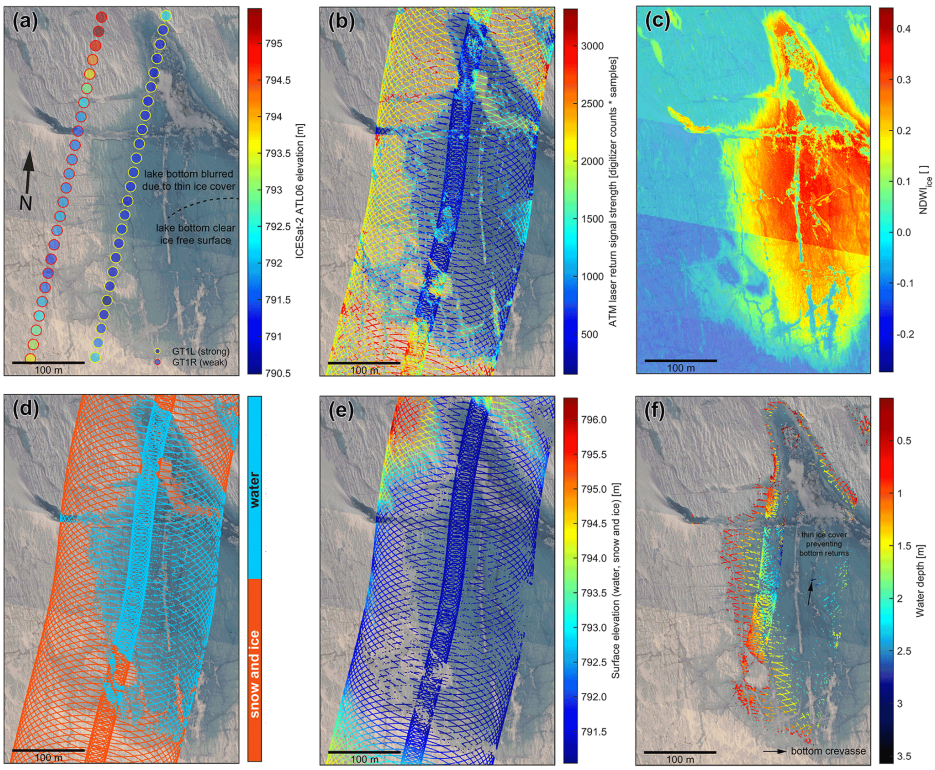


Figure 1.21: **a) A mosaic of five natural colour images of a supraglacial lake from the Airborne Topographic Mapper (ATM) on 15 May 2019, (b) Return signal strength of the ATM laser footprints, (c) $NDWI_{ice}$ plotted on the image mosaic, (d) Surface classification of ATM laser footprints, (e) Surface elevation of laser footprints over snow and ice, (f) Water depth of the hydrological feature (Studinger et al., 2022).**

For example, drone imagery has been used to generate high-resolution (0.01 m) Digital Elevation Model (DEM) to detect channels using flow routing techniques on ice surfaces (Karlstrom and K. Yang, 2016; Rippin, Pomfret, and King, 2015). However, these methods are beyond the scope of the thesis and are not discussed. Here, I evaluate manual, thresholding, and machine learning techniques to classify supraglacial hydrology in optical satellite imagery. In addition, I evaluate the

techniques used to quantify the depth of lakes using remote sensing techniques.

1.6.1 Manual approaches

Manual digitisation of lakes (Figure 1.22) and channels from a combination of true colour composites, band combinations or ratios (e.g., Normalised Difference Water Index (NDWI)), and standard image enhancement procedures (e.g. histogram equalisation and contrast stretching) has been implemented for small-scale studies (Fitzpatrick et al., 2014; Lampkin and VanderBerg, 2014; Langley et al., 2016; K. Yang and L. C. Smith, 2016). However, due to the time-consuming nature of this method, only 150 satellite images were assessed in a previous study (Langley et al., 2016), and it is not an appropriate method for large-scale studies.

1.6.2 Thresholding

Classifying an image as lake or non-lake can be done by using static thresholding, which applies an optimised threshold to a single image. The NDWI technique is used to differentiate open water features and make them more visible in remote sensing images. The original NDWI, $NDWI_{GNIR}$, uses near-infrared radiation and visible green light to emphasise open water and reduce the presence of soil and vegetation (Equation 1.1, McFeeters, 1996), due to the higher reflectance of Near Infra-Red (NIR) of surrounding terrestrial features compared to water. The $NDWI_{GNIR}$ has been used to detect supraglacial hydrology on both ice sheets (Datta and Wouters, 2021; Stokes et al., 2019), and its accuracy has been confirmed by comparing it to manually delineated features (Figure 1.22). The total lake area for each sample in the study had an areal difference of less than 0.5% (Stokes et al., 2019).

$$NDWI_{\mathrm{GNIR}} = \frac{GreenBand - NIRBand}{GreenBand + NIRBand} \tag{1.1}$$

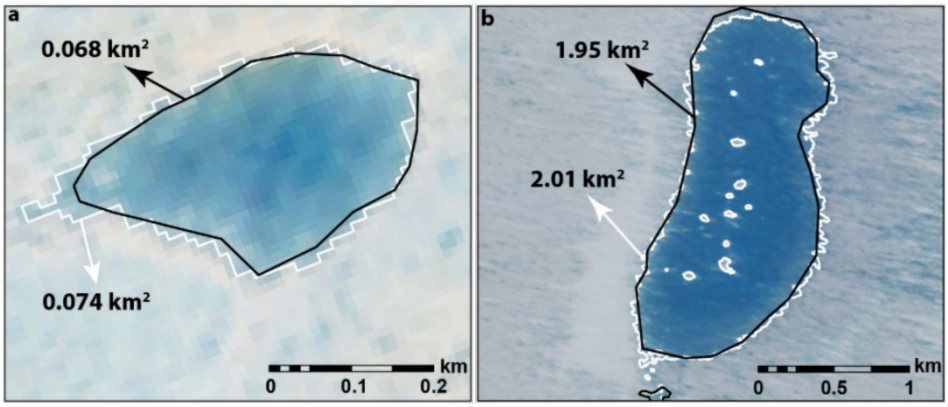


Figure 1.22: **NDWI classification (white outlines) versus manual digitisation (black outlines) from Sentinel-2 imagery for two SGLs on the Amery Ice Shelf (Stokes et al., 2019, Supplementary information).**

However, this approach has been described as unsuitable for glacial environments, as melting ice, snow, and firn also have low reflectivity in the NIR wavelengths (K. Yang and L. C. Smith, 2013). Consequently, $NDWI_{BR}$ (Equation 1.2, also known as $NDWI_{ice}$) was developed to differentiate supraglacial water features on the Greenland ice sheet from WorldView-2 images (K. Yang and L. C. Smith, 2013). A method to delineate actively flowing streams in high-resolution satellite imagery, which uses spectral and pattern information, was based on the initial $NDWI_{BR}$ approach and has a reported overall accuracy of 85.2%. Relying solely on the overall accuracy for the classification of SGLs may not provide a full understanding of the model performance as a result of the class imbalance in satellite imagery. In the context of SGLs, these water bodies are often a minority compared to non-water features such as ice and terrain. As a result, a model that predicts "no lake" for every pixel can achieve a high overall accuracy by correctly

classifying the majority class, but fails to detect the important minority class, SGLs. Therefore, a high overall accuracy metric can be deceptive, masking the model's inability to accurately identify the target lakes. When working with unbalanced data sets, a metric that penalises inadequate performance of the minority class can be a better indicator of model performance, such as the F1 score. The F1 score is a metric used in classification tasks that combines precision (the ability of a model to avoid false positives) and recall (the ability to identify relevant classes) into one value, offering a balanced assessment of a classifier's overall performance. $NDWI_{BR}$ has been used on GrIS to distinguish between lakes, channels, and slush in data from various sensors (Fitzpatrick et al., 2014; K. E. Miles et al., 2017; Williamson et al., 2018; K. Yang et al., 2021), with Root Mean Square Error (RMSE) values ranging from 0.007 km^2 to 0.32 km^2 compared to the results of other models. Generally, the thresholds on $NDWI_{BR}$ are in the range of 0.2 to 0.5 (J. F. Arthur et al., 2020a). SGHF have been studied on the EAIS using $NDWI_{BR}$ with thresholds from 0.12 to 0.25. This has been used to distinguish between shallow water/slush and medium deep water, with misclassifications ranging from 11 to 23% for lakes and partially snow or ice-covered lakes (Bell et al., 2017; Jawak and Luis, 2014).

$$NDWI_{\mathrm{BR}} = \frac{BlueBand - RedBand}{BlueBand + RedBand} \tag{1.2}$$

The NDWI index has been employed to map supraglacial lakes using Sentinel-2, Landsat-8, MODIS, ASTER, and other auxiliary data (Stokes et al., 2019; Williamson et al., 2018; K. Yang and L. C. Smith, 2013). The use of extra filters based on individual image bands has been demonstrated to enhance the effectiveness of NDWI thresholding techniques, reducing misidentification of lakes (Fitzpatrick et al., 2014; M. Moussavi et al., 2020), with accuracies of more than 94% reported (M. Moussavi et al., 2020). These filters include additional thresholding to eliminate misclassified rock and cloud pixels. Red/blue band thresholding, which uses a

ratio of blue to red surface reflectance values, is an additional static thresholding approach used to define water pixels in a scene (Banwell et al., 2014; Box and Ski, 2007). This method exploits the stronger attenuation through water of red light, as compared to blue light, within the optical part of the spectrum. A threshold placed on the ratio of blue to red reflectance is then used to identify water pixels (Banwell et al., 2014; Pope et al., 2016).

Static thresholding techniques are generally reliable in distinguishing SGLs, yet they can cause the exclusion and misidentification of small streams and channels (L. C. Smith et al., 2015). Manual adjustments and tuning of methods are often necessary to meet certain scientific objectives, and the effectiveness of these techniques can vary significantly in different parts of the same ice sheet and even within the same image tile (Hochreuther et al., 2021; Watson et al., 2018). Dynamic thresholding, which applies a different threshold to each region or data tile based on its optical, glaciological, or climatological properties, provides a way to reduce noisy misclassified areas (K. E. Miles et al., 2017; Williamson et al., 2017). This approach involves dividing the image into smaller windows around each lake, making it simpler to distinguish between lake and non-lake. Possible problems that may arise include determining the most suitable window size and the danger of misidentifying in intricate parts of the image, such as when a lake is close to bedrock, blue ice, highly crevassed areas, or cloud shadow. Adaptive classification has been employed to recognise and map SGL features using MODIS imagery. This object-orientated method takes into account information from groups of pixels, rather than individual pixels, allowing the inclusion of size, shape, and contextual information from lakes (Johansson and Brown, 2013). The surface reflectance values of the red and blue bands were used to extract lake information, such as their morphology and environment. This multiclass approach enables SGLs to be identified by their length, shape, and reflectance. While, an approach has been

developed to improve the delineation of SGCs from multispectral imagery in Greenland (K. Yang et al., 2017). Automated approaches to lake digitisation have been shown to generate significant errors (omitting, for example, 29% - 48% of lakes compared to manual approaches, A. A. Leeson et al., 2013). Therefore, more recently, machine learning algorithms have been developed to combat this.

1.6.3 Machine learning

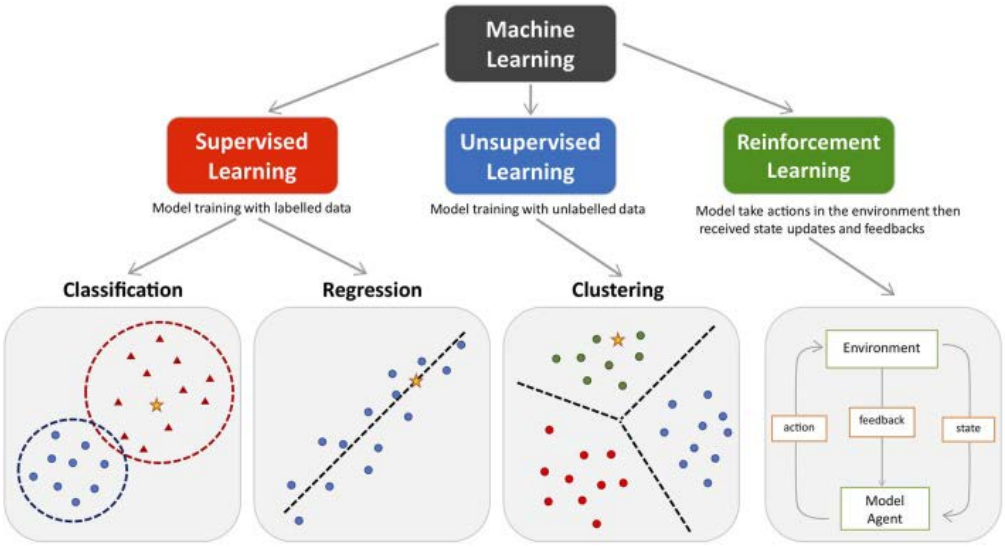


Figure 1.23: **Machine learning techniques encompass supervised learning, such as classification (e.g. random forest) and regression (e.g. linear regression), unsupervised learning, such as clustering, and reinforcement learning, which improves model performance by interacting with the environment. The coloured dots and triangles in the figure represent the training data, while the yellow stars represent the new data that can be predicted by the trained model (Peng et al., 2021).**

Machine Learning (ML) is a branch of artificial intelligence that focusses on the use of data and algorithms to emulate how humans learn, improving performance

through experience. Deep learning is a subset of ML based on artificial neural networks (simply, neural networks). Neural networks are interconnected groups of nodes modelled after the human brain. They are composed of many layers of artificial neurons that work together to build relationships and process information to make a prediction (Figure 1.25).

There are three different types of learning used in machine learning and deep learning: supervised, unsupervised, and reinforcement learning (Figure 1.23). Supervised algorithms require labelled data to be inputted. The model is trained on the desired output and the prediction is compared with the label, with adjustments made depending on the results. This process is repeated until the desired accuracy is achieved or the learning plateaus. Unsupervised learning does not require the input of labels. The model is able to compute an answer without the help of labels and can uncover unknown patterns in the data. Reinforcement learning uses trial-and-error approaches to reward the algorithm when it produces a desired outcome and penalise it when undesired outcomes are reached. Machine learning has been employed for the analysis of satellite imagery since the 1990s (e.g., Charlebois, Goodenough, and Matwin, 1993; Huang, Jensen, and Mackey, 1995). However, it has only been applied more recently to supraglacial hydrology (e.g., M. Dirscherl et al., 2020; Halberstadt et al., 2020; Yuan et al., 2020). Within the context of ML for mapping SGHF, 5 techniques have been used, k-means clustering, Random Forest (RF), Convolutional Neural Network (CNN), Support Vector Machine, and U-NET.

1.6.3.1 Classical machine learning

In Antarctica, k-means clustering techniques, a type of unsupervised algorithm, have been used to train supervised classifiers to map SGLs (Dell et al., 2022; Halberstadt et al., 2020; C. S. R. Smith, 2022). The k-means clustering algorithm,

which is the only supervised classification algorithm available in Google Earth Engine, is a popular choice among the research community due to its robustness; therefore, it is a suitable tool for generating training data (Dell et al., 2022). It is essential to validate the results when using a k-means clustering algorithm to label SGLs. This method does not have any prior knowledge of the characteristics of SGLs, which could lead to incorrect cluster assignments or splitting a single SGL into multiple clusters. K-means assumes that clusters are spherical and of the same size, and requires the number of clusters to be specified in advance, which can be difficult. Furthermore, its sensitivity to initial conditions can cause inconsistent results. To enhance accuracy, it is suggested to combine unsupervised clustering with manual validation and post-processing for reliable SGL identification. However, these studies (Dell et al., 2022; Halberstadt et al., 2020; C. S. R. Smith, 2022), classify SGLs with overall accuracies of more than 90% reported.

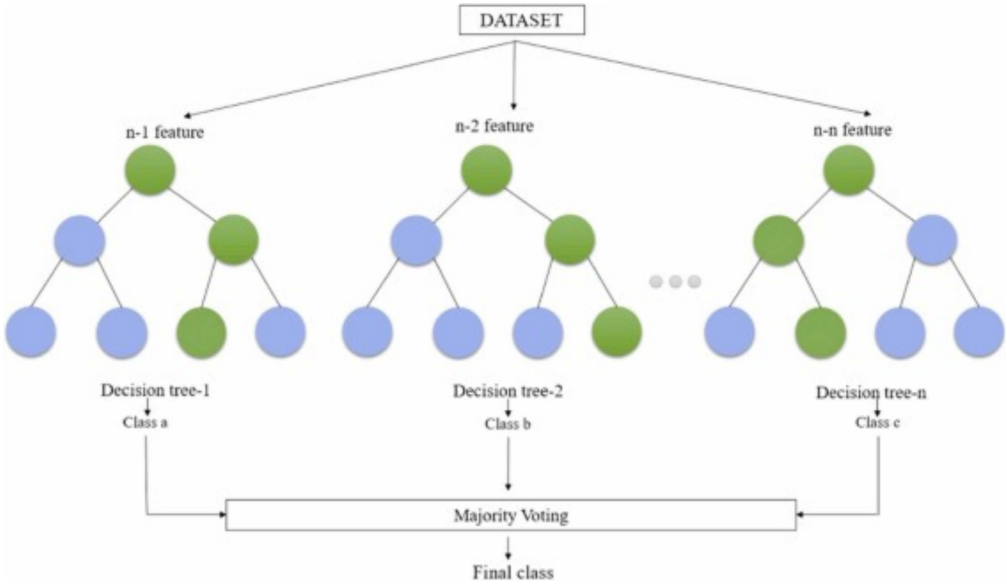


Figure 1.24: **Visualisation of random forest algorithm (Kibria and Matin, 2022).**

Random Forest (RF) is a supervised learning algorithm that requires labelled training data to make predictions (Breiman, 2001). On the other hand, k-means

clustering is an unsupervised learning technique used for data grouping and segmentation. RF combines the predictions of multiple decision tree classifiers, referred to as a *forest*, to enhance the overall classification precision (Figure 1.24). Each decision tree is constructed using a random subset of training data and a random subset of features for each node split. This variability reduces the risk of overfitting, which is especially beneficial when dealing with complex and noisy data, such as optical satellite imagery. During the classification process, each tree in the forest independently assigns a class label (SGL or non-SGL) to a pixel in the satellite image based on the selected features. The final classification is determined by a majority vote among all decision trees. This ensemble approach provides several advantages for SGL classification. Increases the algorithm's capacity to handle high-dimensional data and capture complex relationships between spectral bands, indices, and other relevant features. Compared to other ML classifiers, random forest methods offer comparatively low computation time, parallel processing capabilities, simple parameter tuning, and low risk of overfitting (Belgiu and Drăguţ, 2016; Breiman, 2001; Pal, 2005; Sazonau, 2012). It is made up of a group of unrelated decision trees whose results are combined for a single result (Figure 1.24). The implementation of a RF classifier has allowed the development of a first automated mapping technique of SGHF on Sentinel-2 optical data throughout Antarctica (M. Dirscherl et al., 2020). The RF was trained in 14 training regions and applied in eight spatially independent test regions distributed throughout the Antarctic continent. RF has been exploited to classify ponded SGL and slush in AIS using Landsat-8 (Dell et al., 2022), and SGLs in Greenland from 2016 to 2018 (Hu et al., 2022), however, neither study considered long-term extents of surface water on an ice sheet scale. The performance of the model was evaluated using a variety of metrics, such as the overall accuracy (ranging from 84-98%), the kappa coefficient (0.94), and the F1 score (0.86). The kappa coefficient and the F1 score are not

suitable for comparing the performance of different models, but they are reliable indicators of how a classifier performs on unbalanced datasets.

1.6.3.2 Deep learning

Deep learning algorithms have been applied to the detection and extraction of SGLs in southwest Greenland using CNNs (Yuan et al., 2020). CNNs, one of the various types of artificial neural networks (Figure 1.25), mimic the hierarchical processing of human vision by using convolutional layers to recognise patterns and details in the input image, followed by pooling layers to maintain important information while decreasing computational load. Fully connected layers then make classification decisions based on these features, with activation functions adding non-linearity for pattern recognition. Convolutional, pooling, and fully connected layers are all considered-hidden layers (Figure 1.25). Training CNNs involves supervised learning, adjusting internal parameters to minimise prediction errors. Techniques such as dropout and weight regularisation help avoid overfitting. Convolutional Neural Networks have several advantages over traditional Machine Learning techniques. They can learn relevant features from data without manual feature extraction, making them well-suited for recognising shapes and sizes in SGL. Additionally, they are scale-invariant and can be trained end-to-end. However, they require a large amount of labelled data, complex architectures, and significant computing power. Additionally, they can be difficult to interpret and may overfit with limited data. The authors (Yuan et al., 2020) used CNNs to detect water in Landsat-8 imagery, and then employed morphological and geometrical algorithms to extract the SGLs. This approach has been demonstrated to be more accurate than an adaptive thresholding method, as well as two machine learning methods: Random Forests and Support Vector Machine. The study reported an overall accuracy of 97-99%, recall of 0.95-1.00, and a precision of 0.99-1.00 for CNN,

however, no metric was provided for unbalanced data sets. U-NET is a type of CNN where the image is converted to a vector and, using the same mapping, the vector is reconverted to an image. The distortion of the image is reduced by preserving the original structure. U-NET has been implemented with Sentinel-2 and Sentinel-1 SAR imagery to extract SGL features on GrIS (Di et al., 2021; Lutz, Bahrami, and Braun, 2023). The F1 scores (or Dice coefficients) in both studies were reported to be higher than 0.95.

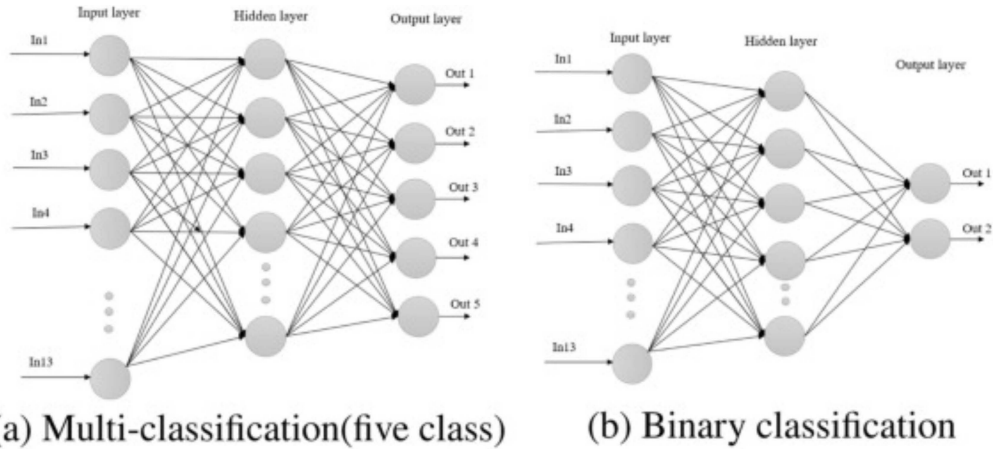


Figure 1.25: **Visualisation of artificial neural network for a) multiple classes, b) binary classes (Kibria and Matin, 2022).**

1.6.4 Methods for estimating supraglacial lake depth & volume

The depth of the SGHF can be measured or estimated using four main methods: in situ sampling, DEM differencing, laser altimetry, and Radiative Transfer Model (RTM). In situ sampling has been used in the literature to validate and/or calibrate other techniques, such as when combined with remote sensing reflectance data (Box and Ski, 2007; Fitzpatrick et al., 2014). However, due to time and financial constraints, in situ measurements are not a practical approach for retrieving depth

on a continental scale. DEM differencing has been used to generate SGL (Das et al., 2008; Lampkin and VanderBerg, 2011) and SGC (Karlstrom and K. Yang, 2016) depth estimates. Additionally, DEMs have been used to calibrate SGL depth estimates (M. S. Moussavi et al., 2016; Pope, 2016). However, DEMs have a relatively low temporal resolution. For example, the most recent version of ArcticDEM's model of the GrIS was created using data collected in 2017 (Porter et al., 2018). Laser altimetry has been shown to be a reliable technique for determining lake depths, as evidenced by its use to validate depths (McMillan et al., 2007), calculate uncertainty (Shean et al., 2019), and estimate depths without the need for additional data (Fair et al., 2020). However, since the laser beam only collects data along the flight path of an aircraft or satellite, depth estimates are only available as transects for a limited number of features (Fair et al., 2020). Radiative transfer models, based on the Bouguer-Lambert-Beer law (Equation 1.3, Philpot, 1989), have been used in the literature to estimate lake depths from optical reflectance in satellite or airborne imagery (Banwell et al., 2014; L. Melling et al., 2023; Pope et al., 2016; Sneed and Hamilton, 2007; Tedesco and Steiner, 2011; Williamson et al., 2018). The attenuation of optical wavelengths in the water column creates a relationship between surface reflectance and depth. These models calculate how light interacts with water and ice, and by monitoring changes in the intensity and spectral properties of the reflected light, the optical depth of the feature can be estimated, which is equivalent to the depth of water or ice in supraglacial lakes. The reflectance values, taken from the red or green bands of optical satellite imagery (Figure 1.26), are used to calculate depth.

$$Z = \frac{\ln A_{\mathrm{d}} - R_{\infty} - \ln R_{\mathrm{w}} - R_{\infty}}{\text{-g}} \tag{1.3}$$

Where:

Z is the depth in metres.

A_d is the reflectance of the ice below the SGHF pixel (bottom albedo).

R_∞ is the reflectance of optically deep water.

R_w is the reflectance of the SGHF pixel of interest.

g is the coefficient for loss of spectral radiance in the water column. $\mathrm{g} \approx 2\mathrm{K_d}$ (K_d is the diffuse attenuation coefficient of the downward light, (Maritorena, Morel, and Gentili, 1994)).

Radiative transfer models allow efficient continental-scale estimations of lake depth from available satellite imagery. Using the same satellite imagery used in the areal classification (Sentinel-2 (S2) and Landsat-8 (L8)) removes the need for auxiliary data, resulting in a streamlined workflow for multiyear continental volume retrieval on GrIS. As the bottom albedo, A_d, value cannot be directly obtained from satellite imagery, it is commonly estimated by taking the mean reflectance value around from a boundary of pixels that form the lake boundary (M. Moussavi et al., 2020; Williamson et al., 2017).

1.7 Identified knowledge gaps, challenges, & solutions for SGH remote sensing

This review has revealed a number of knowledge gaps in our understanding of supraglacial hydrology on the Greenland and Antarctic ice sheets. Here, I detail the knowledge gaps, postulate the challenges that I, and the remote sensing community as a whole, face in improving our understanding of this topic, and suggest potential solutions to these challenges.

Knowledge gaps hinder a thorough understanding of the distribution and

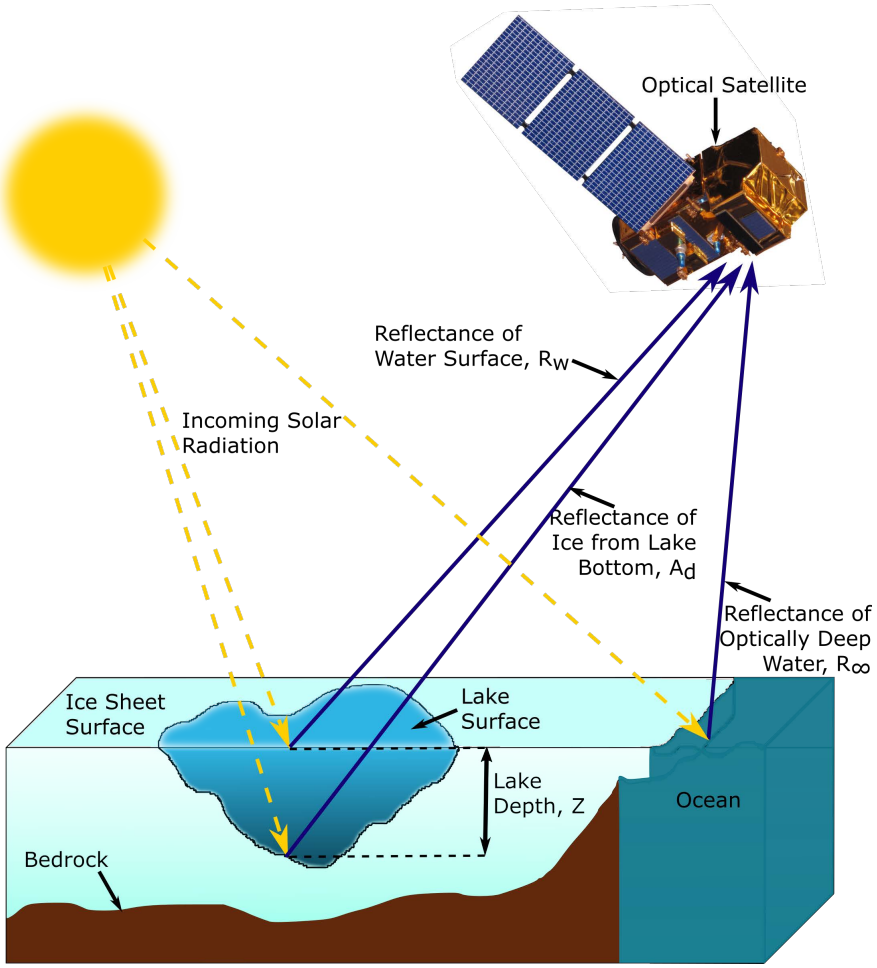


Figure 1.26: **Visualisation of the values which are used in the Radiative Transfer Model to retrieve lake depths.**

temporal evolution of supraglacial hydrology on both ice sheets. For the AIS, there are no large-scale inventories for West Antarctica or the Antarctic Peninsula. In Greenland, inventories are conducted on the ice sheet scale, yet they only take into account a few melt seasons and do not consider interannual seasonal changes in SGH over a longer period of time. The absence of adequate training and validation data sets for machine learning algorithms to measure supraglacial hydrology on an ice sheet scale is clear. Furthermore, while there are approaches used to quantify the

uncertainty in pixel-based classification algorithms (Kwon et al., 2020; Villejo, Illian, and Swallow, 2023; C. Yang et al., 2023), no methods have been applied to the classification of supraglacial hydrology. Therefore, it is difficult to determine whether human-induced climate change has caused an increase in the amount or extent of supraglacial hydrology; if lake and channel features vary seasonally; if hydrological features have an effect on ice dynamics; or if previous estimates for the area and volume of SGHF are reliable.

Major challenges to improving our understanding of the distribution and temporal evolution of SGHF are the lack of usable data, the complexity of combining sensors with different spatial and radiometric resolutions, computer memory and storage limits, correcting for atmospheric effects, the lack of quality training and validation data, the automation of quality assurance procedures, and the quantification of uncertainties.

A lack of data is caused by a variety of elements, such as cloud cover (Figure 1.27), the revisit period, darkness, and the solar elevation angle. However, the use of multispectral satellite imagery is beneficial for mapping meltwater, as it has a unique spectral signature compared to the ice that surrounds it. To address data availability issues, it is common to join optical sensors (Kingslake et al., 2017; A. A. Leeson et al., 2015; B. W. J. Miles, Stokes, and Jamieson, 2017; M. Moussavi et al., 2020; Stokes et al., 2019). Combining products with different spatial resolutions can be a difficult task. To address this, data products can be resampled or pan-sharpened. Landsat-8, for example, has a panchromatic band that can be used for this purpose. The 30 m bands in Landsat-8 can be pan-sharpened to a resolution of 15 m (similar to the 10 m resolution of Sentinel-2) using an Intensity Hue Saturation method (Rahmani et al., 2010). However, pan-sharpening can affect the spectral accuracy of the data, so it should be used with caution. Working with multiple sensors presents an additional challenge: managing large datasets, often

spanning terabytes, requires a reliable computational infrastructure, leading to high costs and logistical complexities. The processing of such data demands substantial memory capacity and computational resources, although the advent of cloud computing, high performance computing clusters, and Google Earth Engine has helped (e.g. Zhu et al., 2022).

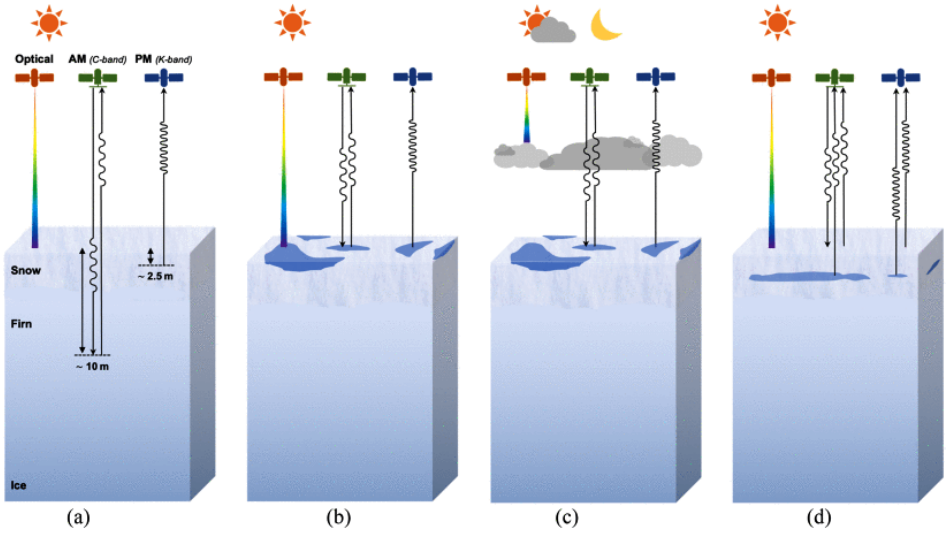


Figure 1.27: **Working principles of optical, active (AM), and passive (PM) microwave sensors under different atmospheric and melting conditions: (a) No melt and sunlight, (b) Surface melt and sunlight, (c) Surface melt and no sunlight due to cloudy and/or nighttime conditions, (d) Near-surface melt and sunlight (Husman et al., 2023).**

Optical satellite imagery captures radiation that is reflected from clouds and the atmosphere, and thus atmospheric correction is often applied. This processing step attempts to remove light reflected from the atmosphere to provide an estimate of surface reflectance. Sentinel-2 Level-1C imagery is provided as a top of atmosphere (TOA) product, while Landsat data are not. There are methods to transform Landsat data into TOA and brightness temperature values (which give insight into

the thermal characteristics of the surface, Band 10 and Band 11 in Landsat-8 imagery), as described by Chander, Markham, and Helder (2009). However, research has suggested that atmospheric corrections can add uncertainty to final products and lead to loss of spectral information, and thus should not be used in machine learning algorithms (Medina-Lopez, 2020).

The lack of in situ data to train, test, and validate any approach is a major challenge in the classification of remote sensing images. Manual and thresholding techniques do not require training data, but for the outputs to be scientifically valid, empirical validation data are necessary. However, machine learning algorithms do require comprehensive and accurate training, testing, and validation data sets. As the available in situ data are insufficient, other means of generating the data must be explored. These include creating synthetic data sets (e.g. Reyes, D'Angelo, and Fraundorfer, 2022), using climate and weather reanalysis datasets (e.g. Zilong, Yubing, and Xiaowei, 2022), and obtaining data sets from satellite imagery using manual or semi-automated approaches. The unsupervised algorithm, k-means clustering, has been used to generate training data for supervised classifiers (Dell et al., 2022; Halberstadt et al., 2020; C. S. R. Smith, 2022). Time-consuming manual intervention is necessary to confirm the labels generated by unsupervised clustering. This can be done by comparing the labels with data from different spectral bands, true colour imagery, or specific spectral indices. Automation of Quality Assurance (QA) processes in remote sensing could be a solution to time-consuming manual post-processing. However, cloud cover, seasonal changes, and sensor limitations can lead to errors or anomalies in the data. The design of algorithms that can detect and mark potential problems while distinguishing between real features and artefacts is difficult, so in studies authors still prefer to do manual QA (e.g., (Lutz, Bahrami, and Braun, 2023; Stokes et al., 2019)).

It is important to assess the uncertainty of remote sensing applications,

especially when assessing the extent and activity of supraglacial hydrology. Sources of uncertainty can include sensor calibration, atmospheric conditions, and data processing. Estimating and propagating these uncertainties accurately is a complex task that requires not only advanced statistical methods but also a comprehensive understanding of the entire remote sensing process. Developing reliable uncertainty models that can generate reliable confidence intervals for supraglacial hydrology estimates is an ongoing challenge in the field. Some studies have attempted to quantify the degree of uncertainty by assigning a percentage to the total area of supraglacial hydrology throughout the ice sheet (e.g. Stokes et al., 2019). This is based on the assumption that the total area of lakes identified from the NDWI model is likely to be comparable to manual digitisation on the ice sheet scale. Other studies have applied uncertainty values on a lake by lake basis, e.g., an uncertainty of a half-pixel width (125 m) to the circumference of each lake (Y.-L. Liang et al., 2012). The supraglacial hydrology on the ice sheets in Antarctica and Greenland changes significantly throughout the year and in different areas. It is essential to have pixel-level uncertainty estimates in order to accurately identify and monitor surface water features. Unfortunately, these estimates are not currently being used.

1.8 Thesis aims, objectives, & structure

In this introduction, I have discussed the main components of the Earth's ice sheets, the interconnected hydrological system, the importance of supraglacial hydrology, the methods and data used to monitor this network, and the challenges remote sensing scientists face in doing so. The melting of the Greenland and Antarctic ice sheets contributes a significant amount to global sea level rise. Although surface melting of the ice sheets has been studied in localised areas, there has been a lack of

long-term or widespread studies that combine these for long-term studies over the entire ice sheets. To address this, this thesis aims to develop a better understanding of the spatial and temporal evolution of the Antarctic and Greenlandic surface hydrological network. This will be attempted by applying conventional methods to previously unsurveyed regions (Chapter 2), using machine learning techniques to generate large data sets to study supraglacial hydrology dynamics at scale (Chapter 3), and developing a novel statistical approach to improve the classification and estimation of uncertainty of supraglacial water bodies (Chapter 4).

My thesis has four primary objectives. (1) To quantify the presence and extent of SGH water bodies across the entire West Antarctic Ice Sheet and Antarctic Peninsula. (2) To evaluate the effectiveness of machine learning to create large SGHF datasets. (3) To identify the distribution and changes of SGH across the entire Greenland Ice Sheet in the last decade. (4) To provide a more accurate estimate of the uncertainty in the classification of supraglacial lakes at the pixel level.

The following three chapters present the results of my thesis, which are dedicated to developing an improved understanding of the continental-scale surface hydrology of the GIS and AIS. In Chapter 2, I describe the work carried out to assess supraglacial hydrology on the West Antarctic Ice Sheet during the 2017 melt season. I developed a dual NDWI approach to quantify the total areal extent of supraglacial hydrology on WAIS and AP for the first time. Additionally, I developed a method to separate different classes of water features (lakes or channels) based on their morphological characteristics. In Chapter 3, I applied ML for the first time, generated the first decadal continental scale dataset of SGH on the Greenland Ice Sheet, and used this to investigate decadal scale trends in the SGHF distribution. In Chapter 4, I develop a Bayesian inference model to analyse the spatial relationship between water pixels for supraglacial lakes with three distinct boundary conditions in west Greenland. This allows me to generate a probabilistic

prediction of the likelihood that pixels are water, thus addressing a limitation in classical machine learning approaches. For the first time, pixel level uncertainty is quantified. Then a binary output is computed by setting a threshold at the point where the probabilistic prediction is more likely to be water than not (0.50). The results demonstrate that this classification approach outperforms static thresholding and performs similarly to the RF and CNN models. In the final chapter of this thesis (Chapter 5), I synthesise the main results, consider the implications of this work in a broader context, and propose potential directions for future research.

Chapter 2

An inventory of supraglacial lakes & channels across the West Antarctic Ice Sheet

Diarmuid Corr[1], Amber Leeson[1,2], Malcolm McMillan[1,3], Ce Zhang[1,2,4], and Thomas Barnes[1].

[1]Centre of Excellence in Environmental Data Science, Lancaster Environment Centre, Lancaster University, Lancaster.
[2]Data Science Institute, Lancaster University, Lancaster.
[3]UK Centre for Polar Observation and Modelling, Lancaster University, Lancaster.
[4]UK Center for Ecology & Hydrology, Library Avenue, Lancaster.

Correspondence:
Diarmuid Corr (dcorr103@gmail.com, d.corr@lancaster.ac.uk)

The following work was published in Earth Systems Science Data on 24th January 2022 (citation: D. Corr et al., 2022). This paper has been edited for self-consistency within the thesis, but otherwise appears here as it was published.

This chapter outlines work carried out as part of the European Space Agency (ESA) funded project 4DAntarctica, in which we assessed supraglacial hydrology on the West Antarctic Ice Sheet. In combination with (Stokes et al., 2019), this forms the first continent-wide assessment helping to quantify the mass balance of Antarctica and its contribution to global sea level rise. I apply thresholds for meltwater classification to satellite images, mapping the extent and manually post-processing to remove false positives. This study provides a high-fidelity dataset to train and validate ML methods.

This paper was edited by Martin Schultz and reviewed by three anonymous referees. DC developed the code, carried out the main body of work and drafted this paper. AL, MM, and CZ provided supervision and contributed extensively to the science, technical details and structure of this paper. TB conducted data processing and manual post-processing and contributed to methodological development. All authors contributed to the manuscript text.

Abstract

Quantifying the extent and distribution of supraglacial hydrology, i.e. lakes and streams, is important for understanding the mass balance of the Antarctic ice sheet, and its consequent contribution to global sea-level rise. The existence of meltwater on the ice surface has the potential to affect Ice Shelf (IS) stability and grounded ice flow through hydrofracturing and the associated delivery of meltwater to the bed. In this study, we systematically map all observable supraglacial lakes and streams

in West Antarctica, by applying a semi-automated Dual-NDWI approach to >2000 images acquired by the Sentinel-2 and Landsat-8 satellites during January 2017. We use a K-Means clustering method to partition water into lakes and streams, which is important for understanding the dynamics and interconnectivity of the hydrological system. When compared to a manually delineated reference dataset on three Antarctic test sites, our approach achieves average values for sensitivity (85.3% and 77.6%), specificity (99.1% and 99.7%) and accuracy (98.7% and 98.3%) for Sentinel-2 and Landsat-8 acquisitions, respectively. We identified 10,478 supraglacial features (10,223 lakes and 255 channels) on the WAIS and Antarctic Peninsula (AP), with a combined area of 119.4 km^2 (114.7 km^2 lakes, 4.7 km^2 channels). We found 27.3% of feature area on grounded ice and 54.9% on floating ice shelves. In total, 17.8% of feature area crossed the grounding line. A recent expansion in satellite data provision made new continental-scale inventories such as these, the first produced for WAIS and AP, possible. The inventories provide a baseline for future studies and a benchmark to monitor the development of Antarctica's surface hydrology in a warming world, and thus enhance our capability to predict the collapse of ice shelves in the future. The dataset is available at `https://doi.org/10.5281/zenodo.5642755` (D. Corr et al., 2021).

2.1 Introduction

The supraglacial hydrological network describes the complex, interconnected system of water movement over the surface of glaciers and ice sheets. Lakes, channels, moulins and crevasses make the network, which forms during the summer months when meltwater is generated at the ice surface. The configuration of the supraglacial hydrological network is transient. It is determined both by the surface topography and the amount of water in the system; greater melt, for example, is likely to lead to deeper and more extensive lakes and channels (Bell et al., 2018; Lüthje et al., 2006; Tedesco et al., 2012).

SGL form when meltwater accumulates in topographic depressions (Bell et al., 2018; Langley et al., 2016). SGL can drain laterally by overflowing their banks or vertically by hydrofracture, when meltwater flows into fractures on the ice surface, increasing the fracture growth (Lai et al., 2020; T. Scambos et al., 2009). Lateral drainage of meltwater can create new channels in the ice surface connecting lakes to other lakes, moulins and crevasses or the ice sheet edge (Bell et al., 2017). Through this interconnected hydrological network, meltwater has been observed to travel over 120 km, and to be redistributed to regions where no melt has occurred locally (Kingslake et al., 2017).

Several recent studies have shown that, contrary to previous understanding, SGL are widespread on Antarctica (Kingslake et al., 2017; Langley et al., 2016; Stokes et al., 2019). A continental-scale inventory has been conducted for EAIS (Stokes et al., 2019), but so far not for WAIS. Lake coverage in West Antarctica has been assessed through small scale ad hoc studies (Banwell et al., 2014; A. A. Leeson et al., 2020; M. Moussavi et al., 2020).

Here, we present a systematic survey of the maximum extent of lakes and large channels on the WAIS and AP during January 2017. Our inventory provides a

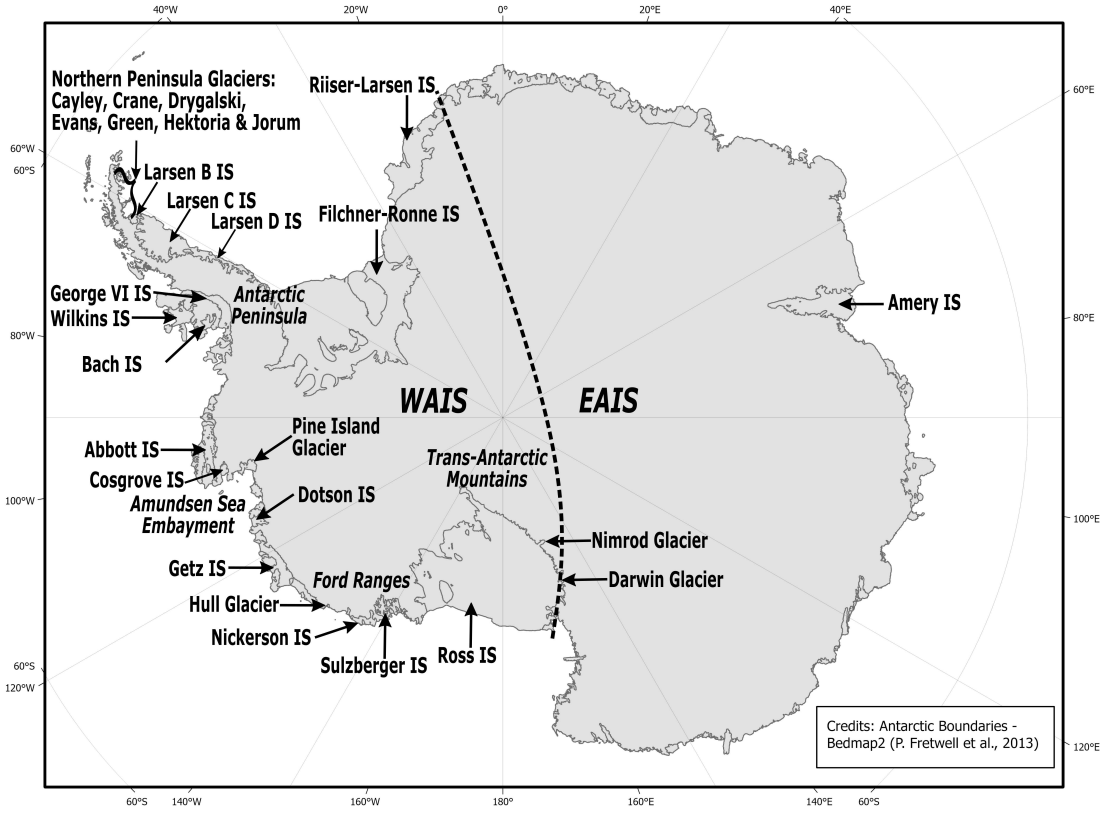


Figure 2.1: **Map of the key locations on the Antarctic Ice Sheet. Antarctic boundaries are according to Bedmap2 (Fretwell et al., 2013). Source: (D. Corr et al., 2022).**

baseline for monitoring future changes. It serves as a training/forcing dataset for other studies, such as those focused upon methodological development or climate and glaciological modelling. High-quality training data are a vital component of ML methodologies. Accurate observations of melt features can act as both boundary conditions and validation for physical models. Knowing the location and characteristics of supraglacial hydrological networks are important on ice sheets because they can alter the location, volume, timing and rate of meltwater drainage (Bell et al., 2018). These provide a mechanism through which climate warming and

associated increases in surface melt might affect the dynamic stability of Earth's polar ice sheets (Bell et al., 2018; Lenaerts et al., 2016; Trusel et al., 2015).

Vertical lake drainage caused by hydrofracturing occurs when water fills a crevasse in the ice sheet to where the water pressure exceeds the fracture strength of the ice (K. E. Alley et al., 2018). The crevasse may propagate through the full ice thickness to the bed, forming a moulin through which the lake drains (Das et al., 2008; McGrath et al., 2012). Rapid lake drainage has been suggested as a mechanism for the breakup of floating ice shelves (Banwell et al., 2019; T. A. Scambos et al., 2000), including the disintegration of the Larsen B IS (Figure 2.1) in 2002 (Banwell, MacAyeal, and Sergienko, 2013; Glasser and T. A. Scambos, 2008; T. Scambos, C. Hulbe, and Fahnestock, 2003). The breakup of an IS may lead to an increase in ice discharge from upstream glaciers (De Angelis and Skvarca, 2003), and an associated increase in their contribution to sea-level rise. Following the collapse of the Larsen B IS in 2002, the Hektoria, Green and Evans glaciers accelerated by up to 8 times their original speed (Rignot et al., 2004).

Meltwater, which enters cracks, crevasses and moulins on grounded ice, drains into the sub- or en-glacial environments (McGrath et al., 2012; van der Veen, 2007). In Greenland, rapid delivery of surface water to the bed has been found to reduce basal friction and temporarily increase ice flow velocities by up to an order of magnitude (Tedesco et al., 2013). It has been hypothesized that mechanisms similar to those observed in Greenland may also occur in East Antarctica (Langley et al., 2016). Indeed, a recent study has shown evidence of five glaciers on the Antarctic Peninsula (Drygalski, Hektoria, Jorum, Crane and Cayley) undergoing near-synchronous speed-up events in March 2017, November 2017 and March 2018 (Tuckett et al., 2019). This suggests the surface meltwater may have entered the subglacial hydrological system.

The conditions under which drainage occurs and indeed whether lakes can

cause hydrofracture, drain rapidly and affect IS stability on Antarctica remain unclear. Supraglacial hydrology may exert a larger effect on Antarctica's future evolution. For example, the UN Paris Agreement's limit on the rise in global temperatures of 1.5°C (`https://unfccc.int/sites/default/files/english_paris_agreement.pdf`) will likely cause the Antarctic Peninsula to experience irreversible, dramatic change to glacial, terrestrial and ocean systems (Siegert et al., 2019). Under this warming (1.5°C), ice shelves will experience a continued increase in meltwater production and meltwater will therefore become more extensive (Siegert et al., 2019). The impact of increased meltwater upon IS stability and ice dynamics lacks understanding. Therefore, mapping the distribution and evolution of the hydrological system from Earth observation has become a key priority of research.

2.2 Data & methods

Here, we describe the selection and pre-processing of S2 and L8 satellite imagery, the identification of candidate water pixels (using NDWI) and the approach used to mask cloud, rock, slush, blue-ice and shaded pixels. The steps involved in post-processing the data and separating SGL and SGC are outlined. The methods differ between sensors as L8 provides thermal information (band 10), whereas S2 does not. Thresholds on individual bands (or indices) are specific to the spectral properties of each sensor, and therefore require adjustment for each sensor (M. Moussavi et al., 2020).

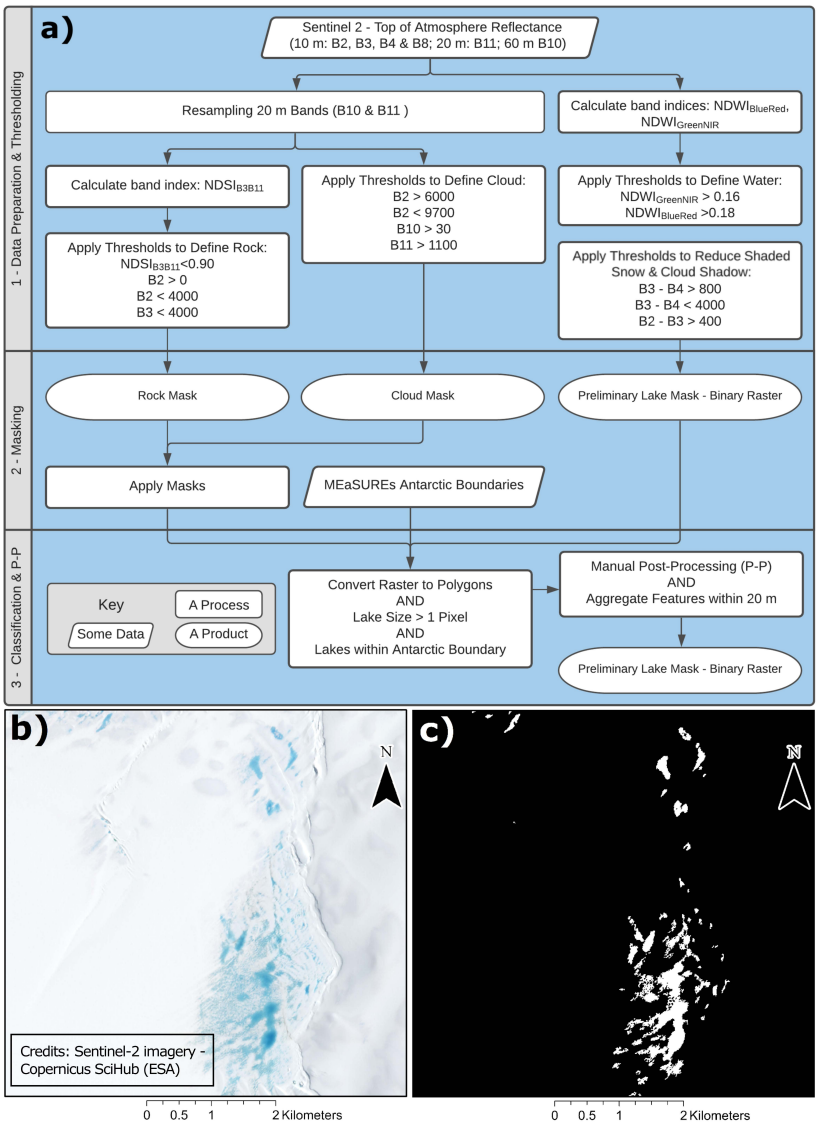


Figure 2.2: **S2 processing chain (a): The input files are the S2 Multi-Spectral Instrument (MSI) products (Red, Blue, Green (RGB) composite of tile T13CET_20170106 from 06 January 2017 on Pine Island Glacier (b)) and the output is the supraglacial lake and channel binary (surface water - not water) classification (c). Imagery was accessed from ESA's Copernicus Scihub,** `scihub.copernicus.eu`**.**

2.2.1 Satellite imagery

In this paper, we use 1682 S2 satellite images to map supraglacial hydrology in West Antarctica. For locations where no S2 data are available, we include 604 L8 images to supplement our dataset. We assess all available scenes with cloud cover below 10%, from 1st to 31st January 2017 on the WAIS. To maximise coverage on the Antarctic Peninsula, which typically experiences more cloudy conditions, we extend the period to 10th February 2017 and use scenes with cloud cover up to 40%.

S2 data are freely available as Top-of-Atmosphere (TOA) reflectance data from the Copernicus Open Access Hub: `https://scihub.copernicus.eu/`. S2 bands 2 (blue), 3 (green), 4 (red), and 8 (NIR) exist at a resolution of 10 m, the highest spatial resolution acquired by the sensor. In contrast, bands 1 (SWIR Cirrus) and 11 (SWIR) are acquired at a coarse resolution of 60 m and 20 m, respectively, and are therefore re-sampled to 10 m using nearest neighbour interpolation for consistency with red, green and blue (RGB) and NIR bands (Williamson et al., 2018). The S2 pixel values represent TOA reflectance units $\times$10,000 and are known as TOA reflectance integers (reflectance$\times 10^4$).

L8 data are freely available from the United States Geological Survey (USGS) Earth Resources Observation Science (EROS) Centre (`https://eros.usgs.gov`) and are provided as a Level-1 data product comprising quantized and calibrated scaled Digital Numbers (DN). Before use, we convert L8 images to TOA reflectance or brightness temperature values following the method of (Chander, Markham, and Helder, 2009). Besides conversion to TOA reflectance, the blue, green, red and NIR bands of L8 data are pan-sharpened using an Intensity Hue Saturation method (Rahmani et al., 2010). This increases the resolution from the native 30 m to 15 m, for comparability with S2, which has a native resolution of 10 m. The remaining L8 bands used, 6 (SWIR, 30 m) and 10 (thermal infra-red sensor, 100 m) are also

re-sampled, using nearest neighbour interpolation as with the S2 data.

2.2.1.1 Normalised Difference Water Index (NDWI) thresholding

Multi-spectral satellite imagery is commonly used to detect open water on ice sheet surfaces (Langley et al., 2016; A. A. Leeson et al., 2020; K. E. Miles et al., 2017; M. Moussavi et al., 2020; Stokes et al., 2019; Williamson et al., 2018). These methods exploit differences in the spectral signatures of open water and snow/ice/firn at optical frequencies. The NDWI performs well in identifying SGL in Antarctica, using either NIR and green bands (Equation 1.1), or blue and red bands (Equation 1.2 Morriss et al., 2013; M. S. Moussavi et al., 2016; Stokes et al., 2019; Williamson et al., 2017; Xu, 2006).

Supervised classification algorithms are in their infancy in the supraglacial hydrology field (M. Dirscherl et al., 2020; Halberstadt et al., 2020) and large-scale, continental studies require validation and testing for generalisation and transferability of the methods. The aim of this study was to produce a dataset to assist such studies and, consequently, NDWI thresholding was selected as our approach. Currently, NDWI thresholding methods are the standard approach to mapping supraglacial hydrology in Greenland and Antarctica (Morriss et al., 2013; M. Moussavi et al., 2020; M. S. Moussavi et al., 2016; Stokes et al., 2019; Williamson et al., 2017; Xu, 2006).

Open water features appear as a dark blue colour in optical satellite images because of the rapid attenuation of red light in water relative to blue light. NDWI ratios are, therefore, well suited to map lakes, as they exploit the properties of lakes which make them more easily distinguished from ice at short optical wavelengths (blue wavelengths), and from the snow at long optical wavelengths (red wavelengths) (Box and Ski, 2007; Y.-L. Liang et al., 2012; Morriss et al., 2013; M. Moussavi et al., 2020; Pope et al., 2016; Sneed and Hamilton, 2007; K. Yang and

L. C. Smith, 2013). However, identifying supraglacial lake and channel pixels using NDWI alone is insufficient, because slush, rocks, clouds and shadows can be spectrally similar to water (M. Moussavi et al., 2020). For this reason, they require additional processing steps to identify and mask these features in each image. Additional, manual post-processing steps described in Section 2.1.3 Post-processing, carried out by human experts, provide data which is of much higher quality than spectral thresholding alone. The processing chain used to map SGL and SGC in S2 and L8 imagery is shown in Figures 2.2 and 2.3.

2.2.1.2 Cloud, rock masking and elimination of slush, blue-ice and shaded pixels

Thresholds are applied to individual bands, spectral indices, and band combinations, such that we isolated water pixels based on multiple spectral properties. The first step in this process is to remove areas of exposed rock, which are often misclassified as water by the NDWI algorithm (Figures 2.2 and 2.3). We generate rock masks for S2 images by defining a threshold (<0.9) on a Normalised Difference Snow Index (NDSI) (Equation 2.1). The NDSI divides the difference in the green and SWIR bands by the sum of the bands. To remove snow and cloud from the rock mask, thresholds are applied to blue (<4000 reflectance $\times 10^4$) and green (<4000 reflectance $\times 10^4$) bands (M. Moussavi et al., 2020). Alongside rock, this mask is also used to remove areas of open ocean, which are found next to ice shelves.

$$NDSI = \frac{GreenBand - SWIRBand}{GreenBand + SWIRBand} \tag{2.1}$$

We performed rock and seawater masking for L8 images by applying a threshold (>650) to the ratio of the blue band and TIRS band (Equation 2.2). To remove snow from the rock mask, a threshold is applied to the blue band (<0.35 reflectance) (M. Moussavi et al., 2020).

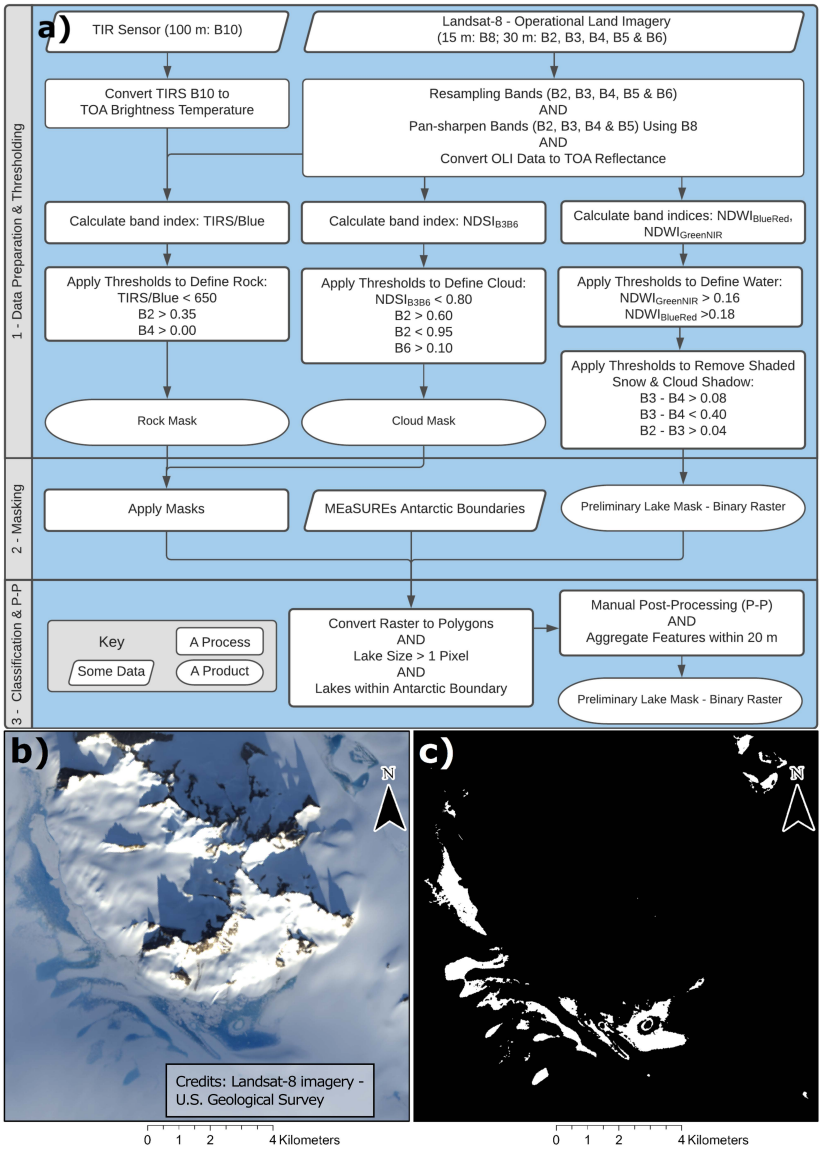


Figure 2.3: **L8 processing chain (a). The input files are the L8 Operational Land Imager (OLI) and TIRS products (RGB composite of tile LC08_L1GT_192129_20170118 from 18 January 2017 on Pine Island Glacier (b)) and the output is the final supraglacial lake and channel dataset (c), in raster and vector format. Imagery accessed from USGS,** `earthexplorer.usgs.gov`**.**

$$\frac{TIRS_1\ BrightnessTemperature}{BlueBand} \tag{2.2}$$

We generated cloud masks for S2 imagery by applying thresholds on the blue band of >6000 reflectance $\times 10^4$ and < 9700 reflectance $\times 10^4$ and thresholds of >1100 reflectance $\times 10^4$ and >30 reflectance $\times 10^4$ to SWIR and SWIR Cirrus, respectively (M. Moussavi et al., 2020) (Figure 2.2). We employ the cloud mask for L8 imagery using thresholds on the blue (>0.60 reflectance and <0.95 reflectance) and SWIR bands (>0.10 reflectance) and a threshold of <0.80 on the NDSI equation (M. Moussavi et al., 2020). For bands that have been resampled to increase spatial resolution (SWIR and SWIR Cirrus for S2, SWIR and Long-Wave Infra-Red (LWIR) for L8), unwanted edge effects are introduced around the edge of the tiles during the up-sampling process. Thresholds on the red band (>0 reflectance for L8) and blue band (>0 reflectance $\times 10^4$ for S2) are therefore used to remove those edge effects.

We extracted water pixels using thresholds on two NDWI calculations, Equations 1 and 2. The first step to calculate $NDWI_{GNIR}$ (Equation 1.1) sets a threshold of 0.16 (for both sensors), above which pixels are considered to have the potential to be water. As stated above, with this threshold alone, the output typically contains slush, blue-ice, shaded rock and cloud shadow pixels. Rock and cloud pixels are removed in their masking processes, respectively. To reduce misclassification of slush and blue-ice pixels, a threshold of 0.18 is further applied to $NDWI_{BR}$ (M. Moussavi et al., 2020). To highlight the difference between light attenuation properties in water and shaded snow surfaces, thresholds are applied to a combination of blue, green and red bands (M. Moussavi et al., 2020). We filtered shaded snow and cloud shadows using 800 (0.08) < green - red < 4000 (0.40) and blue - green > 400 (0.04) for S2 reflectance $\times 10^4$ (and L8 reflectance). Previous analysis on the distribution of pixel values from L8 and S2 tiles (Figures 2

and 4 in M. Moussavi et al., 2020) represent the different spectral properties of lakes, slush, snow, shaded snow, clouds, cloud shadows, sunlit rocks, and shaded rocks. This analysis was used to determine the thresholds in their approach. The thresholds we select for rock and cloud masking were adapted from their approach (M. Moussavi et al., 2020) and the source code from GitHub (Moussavi, Mahsa, 2019). Thresholds were selected to produce maximum lake delineation, with any additional false positives created removed in the manual post-processing stage. The analysis conducted to select the thresholds, the NDWI thresholding approach and the additional band filters is described in Appendix A

2.2.1.3 Post-processing

The processing chain outlined generates a binary raster of 'water' and 'not water' pixels. We subsequently converted groups of water pixels in the raster into polygons representing discrete lakes or channels. Features smaller than two pixels (200 m^2 in S2 and 450 m^2 in pan-sharpened L8 imagery) are removed from the datasets, as such features are considered to be below the detection limit of the sensors, and more likely to be areas of slush rather than open water (Pope et al., 2016). In addition, all features that are beyond the boundary of Antarctica (i.e., not on grounded ice or floating ice shelves) based on the MEaSUREs coastline of Antarctica (Mouginot and Irvine, 2017) are removed accordingly.

Despite the rock/cloud/shadow masking steps that are applied during image processing, areas of shadow, cloud, rock, crevassing and blue ice can still be misclassified as water, and thus erroneously converted into lake/channel polygons. We manually removed these 'false positives', regarding their appearance in true colour (RGB) composite images during post-processing. $\sim$50% of images required such post-processing step, mainly because of the presence of misclassified rock and/or shadow. Up to $\sim$39% (6708) of all 17186 polygons ($\sim$42% of area) delineated

automatically are false positives upon manual inspection and subsequently deleted. We identify 10,478 supraglacial features in this study. Most of all, false positives are linked to shaded rock, with either rock in shadow or snow cast into shadow by rock. A few densely populated crevassed regions contribute to ∼20% of misclassified features, but because of their small size (<10 pixels), this represents a much smaller percentage of the total misclassified area.

Although no dedicated checks are carried out to assess the number of false negatives, during the visual inspection for false positives, no large features (>50 pixels) were found to be excluded. Minor features (<5 pixels) may remain uncharted. However, their influence on the overall area mapped, and therefore the volume of surface water, is likely to be minimal. Finally, overlapping polygons from each sensor, location and time instance throughout January 2017, are dissolved and polygons within a distance of 20 m are aggregated - to provide a more continuous delineation of hydrological connectivity.

2.2.2 Accuracy assessment

We tested the fidelity of our method by comparing our results with 97, 184 and 105 (119, 135 and 46) manually delineated lakes and SGC from S2 (and L8) imagery on test sites crossing the grounding line on Amery, George VI (GVI) and Bach ice shelves respectively (Figure 2.1). These ice shelves vary in their glaciological and climatological characteristics, which result in a range of feature geometries and settings that are considered to represent the whole of Antarctica. In each of the three regions, we selected test regions that encompass extensive surface hydrological meltwater and host close to 100 individual supraglacial features of varying sizes. This resulted in test regions measuring 210 km^2 (for Amery IS) and 100 km^2 (for GVI IS and Bach IS).

Table 2.1: **Sensitivity (sen.), specificity (spec.) and accuracy (acc.) of the S2 and L8 methods for each of the test sites: Amery, George VI and Bach ice shelves and the mean values across sensors and sites for each.**

Test site	S2 sen.	L8 sen.	S2 spec.	L8 spec.	S2 acc.	L8 acc.	Mean sen.	Mean spec.	Mean acc.
Amery	0.93	0.964	0.979	0.992	0.971	0.987	0.947	0.986	0.979
George VI	0.833	0.614	0.995	0.999	0.985	0.982	0.724	0.997	0.984
Bach	0.797	0.75	0.998	0.999	0.992	0.993	0.774	0.999	0.993
Mean	0.853	0.776	0.991	0.997	0.983	0.987	0.815	0.994	0.985

Amery IS, the third-largest IS in Antarctica, is in East Antarctica. The chosen test site on Amery IS is well suited for automated processing due to clear spectral differences between surface water and ice pixels in the region. However, a small area of blue ice, which has a similar spectral signature to that of open water, is challenging to differentiate automatically. GVI, one of the largest ice shelves on AP, constrained between the western side of the AP and Alexander Island, loses most of its mass to melt rather than calving (Roberts et al., 2008). GVI IS has two ice fronts, one situated around 500 km further north and so experiences different climatic conditions (Cook and D. G. Vaughan, 2010). The test site, situated around the middle of the IS, crosses the grounding line of Alexander Island and contains rock and shaded pixels, presenting a more difficult task for the method. Bach IS is on the coast of Alexander Island, to the east of the AP. Although to date, it has shown relative stability in an area where other ice shelves (particularly Wilkins IS) have undergone major collapse (Humbert and Braun, 2008; T. Scambos et al., 2009), Bach IS could be the next IS under threat of break-up (Cook and D. G. Vaughan, 2010). The test region on the Bach IS offers a contrasting stress regime (unconfined vs confined flow) to that of the GVI IS, which has the potential to create different lake geometry (Cook and D. G. Vaughan, 2010; T. A. Scambos et al., 2000).

We manually delineated lakes and channels in each test area using true colour

(RGB) composites of S2 and L8 data. Three separate 'expert' users delineated each area (with expertise in remote sensing of supraglacial hydrology). We combined the three manual inventories to form a high fidelity reference dataset of the lake and channel features, where only pixels that are unanimously assigned as 'water' (i.e., identified as water by all three users) are included. To assess the performance of our method, we calculated confusion matrices that compared Manual (Man) and the final Automated (Auto) datasets (following post-processing) on a per-pixel basis. From the confusion matrices, sensitivity, specificity and accuracy have been derived (Equations 2.3, 2.4 and 2.5).

The sensitivity, or true positive rate, is the number of True Positive (or Man Water: Auto Water in a confusion matrix) predictions divided by the number of manually identified water pixels in the test data. It is a measure of how well we correctly identified surface water pixels.

$$Sensitivity = \frac{Man.\,Water : Auto.\,Water}{Total\ Man.\,Water} \tag{2.3}$$

The specificity, or true negative rate, is the number of True Negative predictions (or Man Not Water: Auto Not Water) divided by the number of manually identified not-water pixels in the test data. It is a measure of how well all other non-water pixels are identified.

$$Specificity = \frac{Man.\,Not\ Water : Auto.\,Not\ Water}{Total\ Man.\,Not\ Water} \tag{2.4}$$

We calculated the accuracy (ACC) by dividing the sum of True Positive and True Negative predictions by the total number of pixels. It gives a quantitative assessment for the accuracy of all pixels, both water and not water.

$$ACC = \frac{(Man.\,Water : Auto.\,Water) + (Man.\,Not\ Water : Auto.\,Not\ Water)}{All\ Pixels} \tag{2.5}$$

We computed sensitivity, specificity and accuracy values for each test site and sensor, Table 2.1. Across both sensors, sensitivity ranges between 61.4% and 96.4%, specificity between 97.9% and 99.9% and accuracy between 97.1% and 99.3%. On average, S2 yields a higher sensitivity (85.3% versus 77.6%) than L8, while values for specificity and accuracy have higher averaged values for L8 than S2.

The large range in sensitivity is likely because of shallow lakes (deemed so by manual users) being classified as ice by the NDWI threshold, especially on GVI and Bach ice shelves, where there are many shallow lakes. In contrast, the range in specificity (i.e., how well non-water pixels are identified) is smaller, because the analysis was carried out after manual post-processing, and misclassified pixels were already removed. This suggests that, for applications where identifying shallow lakes is important, we may further improve the sensitivity by incorporating additional manual checks for false negatives into the NDWI thresholding approach.

2.3 Results & discussion

2.3.1 Distribution of supraglacial lakes and streams in WAIS

We used 1682 S2 and 604 L8 scenes to map 10,478 individual supraglacial features in West Antarctica, including the Antarctic Peninsula (Figure 2.4). The dataset comprises 10,223 SGL and 255 channels.

We found SGL in expected regions, on and around the grounding zone of the Antarctic Peninsula ice shelves including Larsen C (∼130 lakes), Larsen D (∼250), GVI (∼5,550), Wilkins (∼1450) and Bach (∼950). We also discovered lakes on grounded ice close to where the remnants of Larsen A (∼10) and B (∼150) ice shelves are located. Sulzberger IS (∼290), Pine Island Glacier (∼360), Riiser-Larsen (∼240) and around the Trans-Antarctic Mountains/Ross IS on Darwin (∼270) and

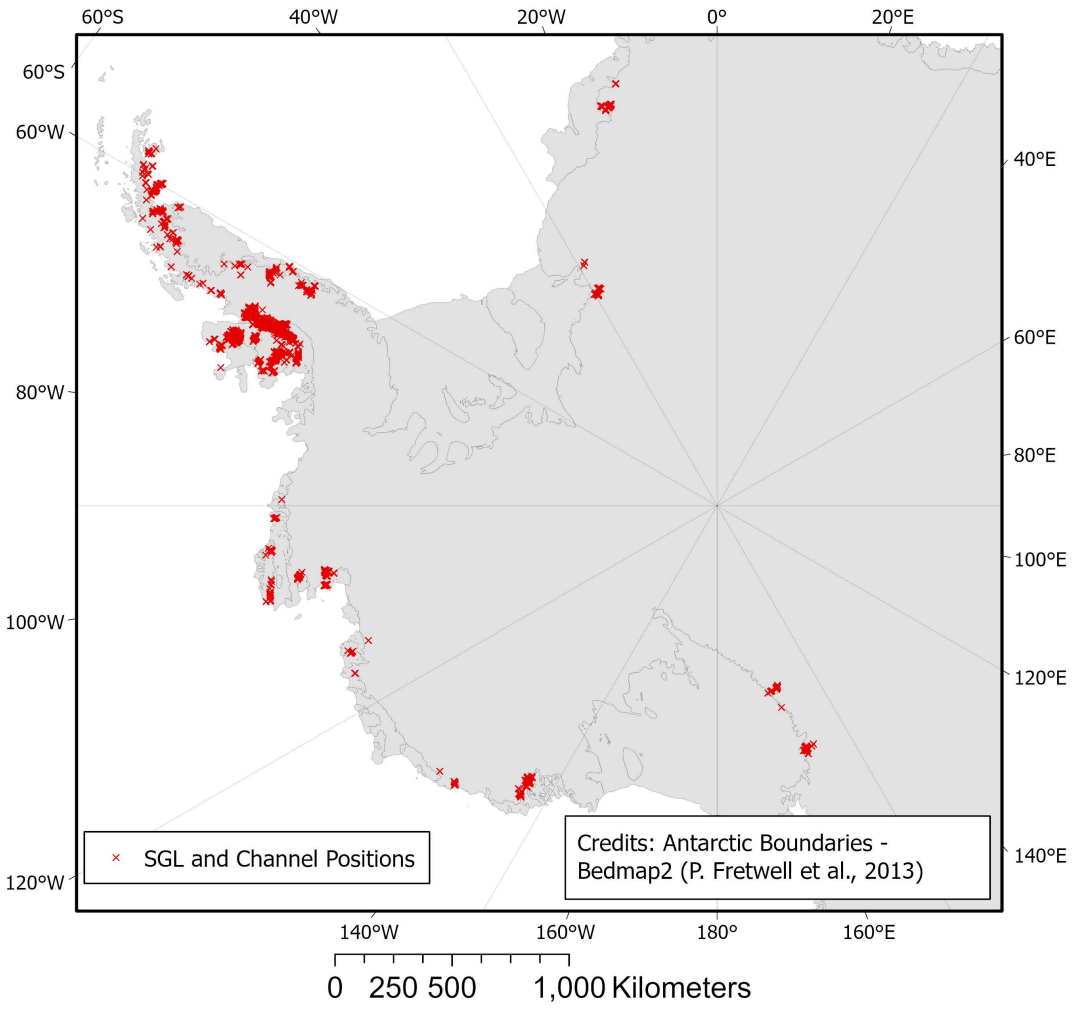


Figure 2.4: **Location of the 10,478 SGL and SGC on the WAIS and AP represented by the red crosses as mapped in S2 and L8 imagery from January and February 2017. Antarctic boundaries are according to Bedmap2 (Fretwell et al., 2013).**

Nimrod (∼90) glaciers, are identified to host SGL activity in this study and others (Kingslake et al., 2017). Studies identifying lakes in the Ford Ranges region (on Hull Glacier (∼45) and Nickerson IS (∼80)) and in Amundsen Sea region (on Dotson (∼35), Abbot (∼125) and Cosgrove (∼45) ice shelves) were published for the first time recently, (J. F. Arthur et al., 2020a; M. Dirscherl et al., 2020), and we confirm the

occurrence of lakes here during January 2017. We identified 255 supraglacial channels on or around the margin of Larsen C, Larsen D, remnants of Larsen B, GVI, Bach, Wilkins, Riiser-Larsen, Dotson and Sulzberger ice shelves and near to the Hull and Pine Island Glaciers. We identified supraglacial meltwater for the first time on the Getz IS, with one lake crossing the grounding line, while a further four border the IS (Figure 2.5). It has been suggested that increased surface melt on the Getz IS will lead to collapse unless active surface drainage can mitigate the effect of surface loading by exporting water to the ocean (Bell et al., 2018).

The proportion of area covered by meltwater in a localised region (Figure 2.6) was calculated using the cumulative lake and channel area in a hexagonal bin. Each bin measured 100 m between parallel sides of the hexagon, while any feature within a search radius of 5 km (longest feature: $\sim$4.7 km) contributed to the proportion of the bin in question. The proportions range from 0 (where no lakes are within 5 km of a bin) to 0.089 km^2 of meltwater area per 1 km^2, with the highest density regions on the Peninsula (GVI, Wilkins and Bach ice shelves), Ford Ranges, Trans-Antarctic mountains and Pine Island Glacier. GVI, which measures $\sim$24,000 km^2 (Cook and D. G. Vaughan, 2010), has a total meltwater area of 29.4 km^2, and a percentage cover of supraglacial meltwater across the IS of 0.12%. Wilkins (meltwater area: 14.0 km^2, total area: $\sim$11,000 km^2 Cook and D. G. Vaughan, 2010) and Bach (meltwater area: 13.0 km^2, total area: 4,500 km^2 Cook and D. G. Vaughan, 2010) ice shelves have maximum percentage cover of 0.13% and 0.29% respectively. Areas with a low proportion of area covered, hosting just a few lakes include Getz IS, the western margin of the Filchner-Ronne IS, Hull Glacier and on James Ross Island off the coast of the Northern AP.

We have assessed area distribution for SGL and SGC (Figure 2.7). The total area covered by lakes (114.7 km^2) and channels (4.7 km^2) was found to be 119.4 km^2. The proportion of features on grounded ice (GI), floating IS and crossing the grounding

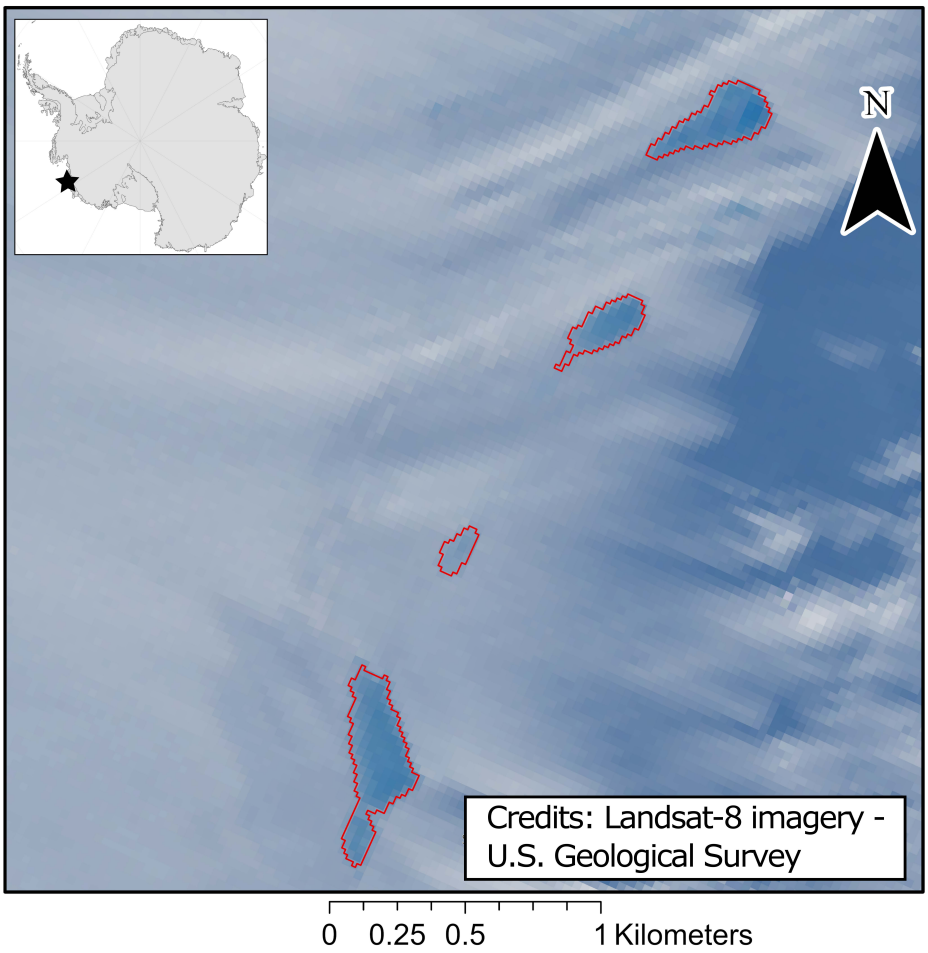


Figure 2.5: **Lakes identified for the first time in the Western Amundsen Sea Sector of West Antarctica, with the largest, southernmost lake of the four crossing the grounding line of the Getz IS. The image is a true colour composite of L8 satellite imagery, Tile LC08_L1GT_166131_20170112 from 12 January 2017. The red outlines are the lake polygons resulting from our NDWI threshold approach. Inset: Getz IS and lake locations in Antarctica. Imagery was accessed from USGS,** `earthexplorer.usgs.gov`.

line (GL) are computed (Figure 2.8). Distribution of glaciologically important parameters (Figure 2.9b-f,h) for the 10,478 supraglacial features including, distance

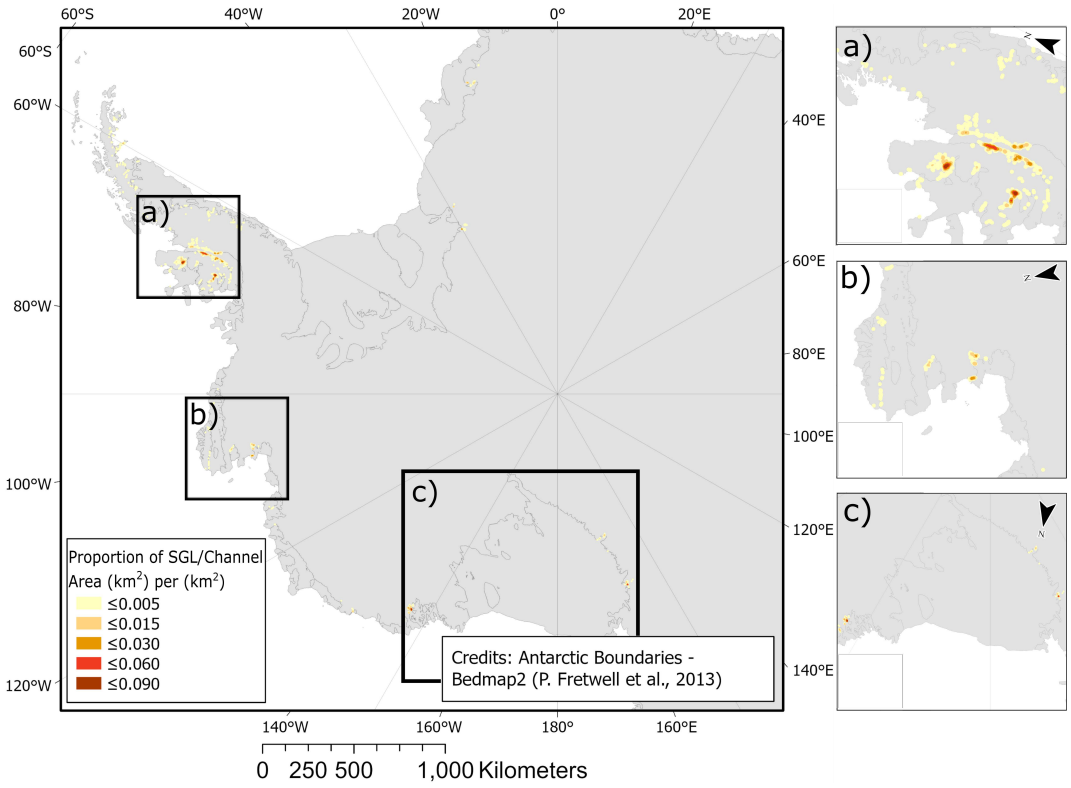


Figure 2.6: **The proportion of lake and channel area covered by meltwater per km^2 in each region on WAIS and AP, where surface water is identified. Inset: high cover regions on a) AP (GVI, Wilkins and Bach ice shelves), b) Amundsen Sea region (Pine Island Glacier) and c) Ford Ranges (Sulzberger IS) and Trans-Antarctic mountains (Ross IS, Darwin and Nimrod glaciers). Antarctic boundaries according to Bedmap2 (Fretwell et al., 2013).**

to the grounding line, exposed bedrock and coastline were calculated. Ice surface elevation, slope and velocity were observed for each feature and their distribution plotted. Finally, the distribution of meltwater volume was calculated (Figure 2.9a).

The largest lake identified (~2.9 km^2) intersects the grounding line of the Sulzberger IS, although most lakes and channels are an order of magnitude smaller. For example, we found 8700 (83%) of the features to have an area of less than 0.1

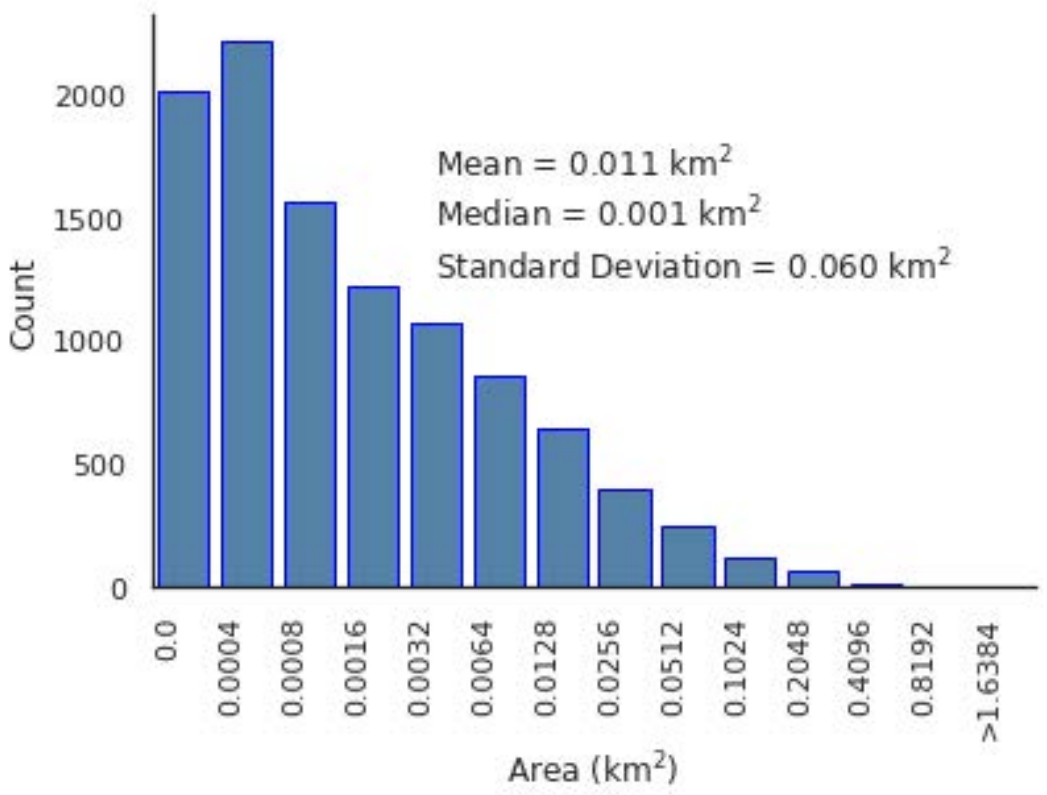


Figure 2.7: **Distribution of SGL and channel area (km^2) on WAIS and AP. Note: bin sizes double from left to right. Values for mean, median and standard deviation for the distribution are included in the figure.**

km^2 (Figure 2.7). Lakes make up 96.1% of total feature area and 97.6% of all features (Figure 2.9g). More than half (54.9%) of the total open water area was entirely on floating ice, with 27.3% on grounded ice entirely (Figure 2.8). Among the features on grounded ice, the lake found the farthest inland of the grounding line is on the Antarctic Peninsula, 47.2 km from the GVI IS. In terms of floating ice, the lake found farthest from the grounding line is also on GVI at a distance of 12.6 km. Over half of the open water area (56.4%) is within 1 km of the grounding line according to Bedmap2 (Fretwell et al., 2013), with 17.8% of total open water area intersecting

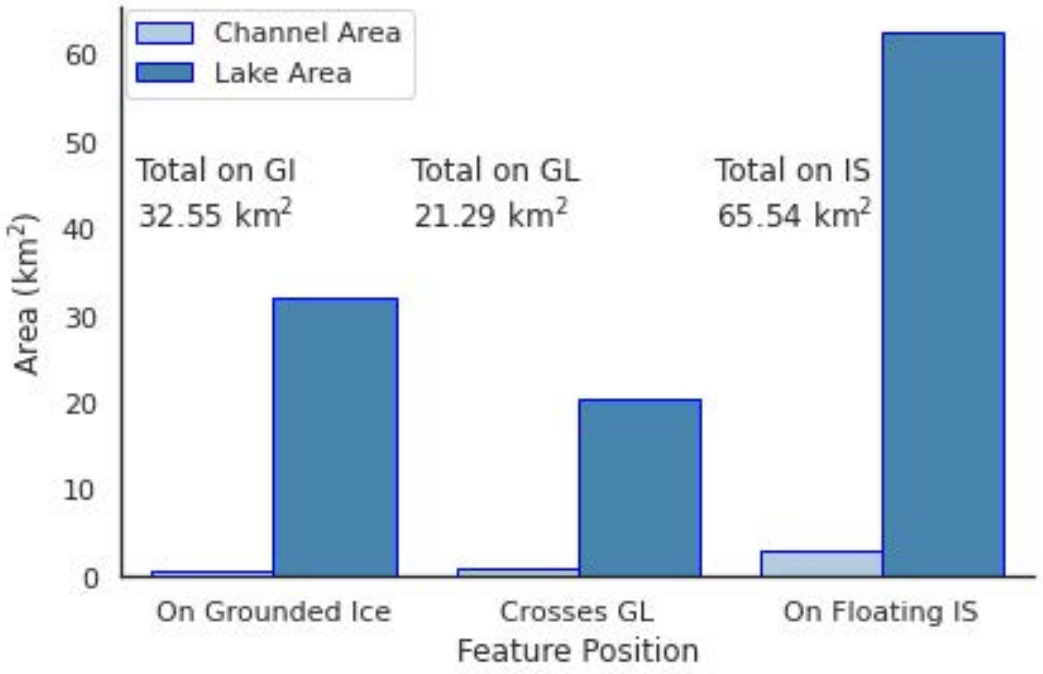


Figure 2.8: **Area split between channels and lakes completely on Grounded Ice (GI), crossing the Grounding Line (GL) and completely on floating ice shelves IS, across the WAIS and AP.**

the grounding line (Figure 2.9b).

Exposed bedrock has a lower albedo than snow or ice, which can increase the absorption of incoming solar energy, leading to higher rates of melting within the local area. We find that 78.1% of the total feature area (and 80.1% of all features) exists within 10 km of exposed bedrock (Figure 2.9c). Lakes are also found at substantial distances inland (Figure 2.9d), including over 504 km away from the closest coastline. Most of the open water, however, (64.9% of features and 63.1% of area) was found within 100 km of the coast, according to MEaSUREs data for the coastline of Antarctica (Mouginot and Irvine, 2017; Rignot et al., 2013).

We found most features (80.8%, representing 77.5% of the total feature area) at

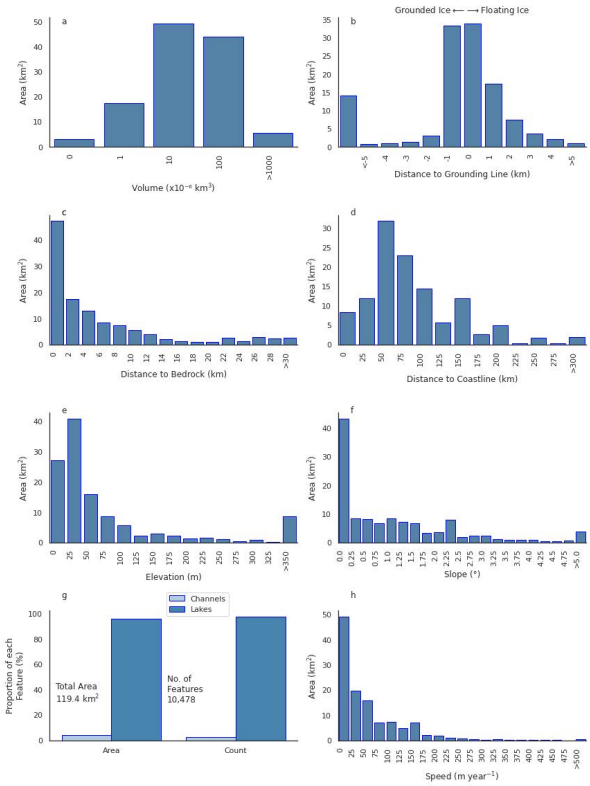


Figure 2.9: **Area distributions of supraglacial lakes and channels on the WAIS and AP according to various glaciological variables. (a) individual feature volumes from the area-volume scaling relationship (Equation 2.6); (b) distance of each feature to the grounding line (negative values indicate lake/channel positions further inland from the grounding line); (c) distance of each feature to nearest exposed bedrock; (d) distance of each feature to the ice margin/coastline; (e) elevation at the centroid of each feature; (f) the surface slope at the centroid of each feature; (g) area and frequency split between channels and lakes on the WAIS and AP; (h) ice-flow speed for each feature. Imagery was accessed from ESA's Copernicus Scihub,** `scihub.copernicus.eu` **and USGS,** `earthexplorer.usgs.gov`**.**

low elevations (Figure 2.9e), i.e. between 0 and 100 m a.s.l, while 87 lakes/channels (or 0.4% of area) were at elevations greater than 1000 m a.s.l., with two (in the

mountain region around 40 km East of GVI IS) as high as 1306 m a.s.l. Most lakes/channels (57.9%) occur on surface slopes $<1^{\circ}$ (Figure 2.9f). This accounts for 55.7% of the total area.

To estimate the ice flow velocity at the geometric centroid of each feature, we extracted the ice surface velocity from the MEaSUREs InSAR-Based Antarctica Ice Velocity Map (Mouginot, Scheuchl, and Rignot, 2012; Rignot, Mouginot, and Scheuchl, 2011; Rignot and Irvine, 2017). Ice flow velocities in lake-covered regions ranged from $\sim$0 to $>$1357 m/year, however 57.8% of the total feature area (and around 47% of the total features) was on ice flowing slower than 50 m/year (Figure 2.9h).

To estimate the volume of water contained within each feature, we use an area-volume (A-V, Equation 2.6) scaling relationship from literature (Stokes et al., 2019). Based on this relationship, the total volume of meltwater stored in SGL and streams is estimated to be 0.085 km^3 across the entire WAIS and AP.

$$V = 7.16 \times 10^{-4} A \tag{2.6}$$

Due to proportionality between area and volume, the feature containing the maximum volume of water ($\sim$0.002 km^3) is the lake on the Sulzberger IS, identified as the largest by surface area. 86.9% ($>$9000 features) of all lakes/channels have volume between 0 km^3 and 0.0001 km^3, while this range accounts for only 17.2% of total area. Conversely, 41.4% of the total lake/channel area is represented by just 144 features that have a volume greater than 0.0001 km^3.

2.3.2 Lake vs channel features

The increased spatial resolution offered by the current generation of optical satellite sensors, such as S2, makes mapping supraglacial rivers and channels possible.

Here, in contrast to previous studies in Antarctica, we distinguish between lakes and channels using a K-Means clustering approach (D. Arthur and Vassilvitskii, 2007), combining six shape index metrics. The first, a standard Area-Perimeter Ratio (A:P), (Equation 2.7) divides the total area of a feature (A) by the length of its perimeter (P).

$$A : P = \frac{A}{P} \tag{2.7}$$

Second, we use the Iso-Primetric Quotient (IPQ), (W. Li, Goodchild, and Church, 2013), i.e., the ratio of the area of the feature to the area of a circle whose circumference, C, is equal to the perimeter. It is also known as the Polsby-Popper score when it is used to quantify the degree of gerrymandering of political districts (Polsby and Popper, 1991).

$$IPQ = \frac{A}{\pi(\frac{C_P}{2\pi})^2} = \frac{4\pi A}{P^2} \tag{2.8}$$

Providing spatial analysis of complex geographical features can be characterized by the fractal dimension. The Fractal dimension index (*Fractal*) (Y. Chen, 2020) reflects shape complexity across a range of spatial scales. Therefore, it overcomes one of the major limitations of the straight perimeter-area ratio as a measure of shape complexity. Depending on the number of vertices in a polygon, the Fractal dimension index can be a variety of logarithmic ratios (Y. Chen, 2020) (Equation 2.9).

$$Fractal = \frac{2log(\frac{P}{4})}{log(A)} \tag{2.9}$$

Another metric is the ratio of the feature area to the area of a Minimum Bounding Circle (MBC) (A_{MBC}), which is needed to enclose the feature (Equation 2.10). This ratio is known as the Reock score (Reock, 1961).

$$Reock = \frac{A}{A_{\text{MBC}}} \tag{2.10}$$

To measure compactness of the feature (i.e. how neatly the area fits within the perimeter, the most compact shape is a circle) the Schwartzberg score (Inc, 2010) can be calculated (Equation 2.11). It is the ratio of the perimeter of the feature to the circumference of a circle, C_A, whose area is equal to the area of the feature.

$$Schwartzberg = \frac{1}{\frac{P}{C_A}} = \frac{C_A}{P} = \frac{2\pi\sqrt{\frac{A}{\pi}}}{P} \tag{2.11}$$

The final metric, a Width-Length ratio (W:L) (Equation 2.12) is calculated as the ratio of the width (W_{MBR}) to the length (L_{MBR}) of the Minimum Bounding Rectangle (MBR), which surrounds the feature. The MBR is the smallest (by width) required to enclose the full area of the feature.

$$W : L = \frac{W_{\text{MBR}}}{L_{\text{MBR}}} \tag{2.12}$$

The shape indices (Equations 2.7-2.12) were computed for every polygon in the final dataset (Table 2.2). Unsupervised K-Means Clustering (D. Arthur and Vassilvitskii, 2007) was carried out in 6-dimensional space, using each of the six shape indices through the Multivariate Clustering tool on ArcGIS Pro Version 2.5.2. K-Means algorithms identify a starting point (seed) from among the supraglacial features, to grow each cluster. We randomly selected the first seed, while we chose subsequent seeds by directing the selection to seeds farthest in data space from the existing seeds. Small lakes, below 500 m^2 in area, introduced noise to the classification and were labelled as lakes before clustering.

This resulted in 20 distinct clusters which were manually determined to be feature types, of varying shapes and size, lakes or channels (Figure 2.10). Samples from 6 of the 20 clusters are shown in Figure 2.10, with the corresponding value for

each shape index in Table 2.2. As expected, ribbon lakes are similar to channels in most metrics, as both take long, narrow forms. However, the A:P (Equation 2.7) value differs vastly between channels and ribbon SGL. The values displayed for A:P, Fractal, Reock and W:L (Equations 2.7, 2.9, 2.10 and 2.12) demonstrate clear differences between channels and all other lake classes. While IPQ and Schwartzberg (Equations 2.8 and 2.11) are useful in delineating standard, smaller lakes from channels. Through this method, we identify 10,223 lakes and 255 channels to be present during January 2017 on the WAIS and AP (Figure 2.10).

***Table 2.2:* Value of each individual shape index (Equations 2.7-2.12) for the feature type defined in Figure 2.10.**

Shape index	Large SGL	Complex SGL	Ringed SGL	Standard SGL (mean)	Ribbon SGL	Channel
A:P	41.18	28.02	22.03	15.2	42.15	11.9
IPQ	0.11	0.02	0.09	0.4	0.08	0.12
Fractal	1.44	1.35	1.39	1.42	1.42	1.34
Reock	0.31	0.11	0.25	0.38	0.05	0.08
Schwartzberg	0.33	0.16	0.3	0.62	0.28	0.33
W:L	0.54	0.34	0.4	0.5	0.08	0.15

The values reported for accuracy, sensitivity and specificity (Table 2.1, Section 2.2) are for the thresholding approach, which consists of all water pixels, including channels and lakes. Although it would be valuable to provide validation metrics for the classification of water into channels and lakes, due to the lack of an objective definition as to what lakes and channels are it is not possible to compute accuracy, sensitivity, or specificity metrics at present. Channels and lakes are defined from within the classification of surface water, based solely upon their shape. To concretely define channels, would require auxiliary data, such as water flow and topography at instances in close temporal proximity to the satellite imagery. The aim of our channel and lake discrimination is therefore not to provide a measure or

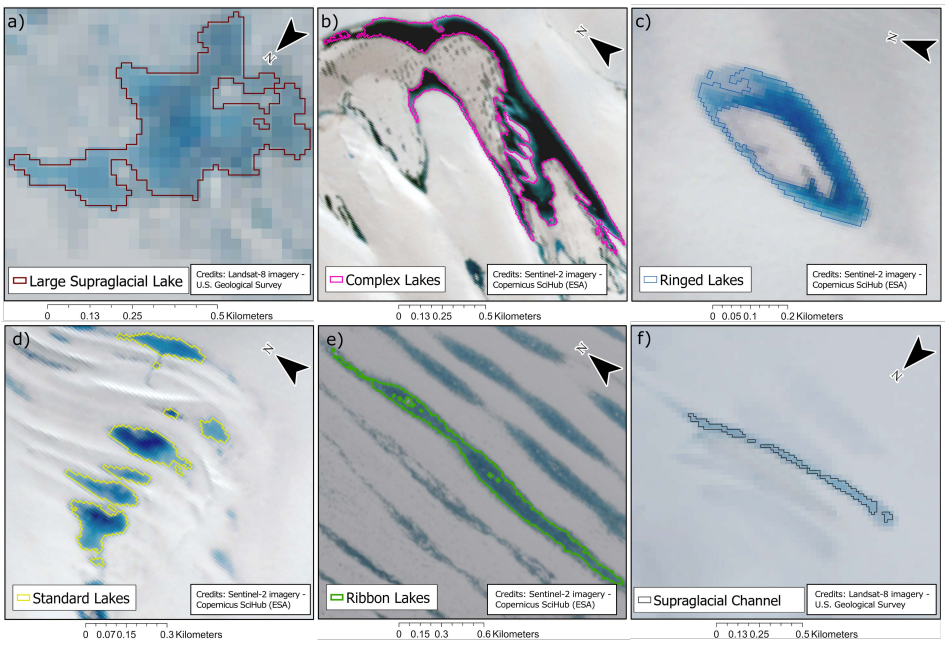


Figure 2.10: **Outlines of supraglacial features over RGB composites from S2 and L8 imagery. These outlines demonstrate 6 of the distinct clusters from the K-Means approach. a) A large SGL covering 0.20 km^2 on Hull Glacier (Figure 2.1); b) complex SGL with area 0.35 km^2 on Bach IS (Figure 2.1); c) ring lake with area 0.05 km^2 on Bach IS; d) 11 'standard' SGL on Bach IS with areas ranging from 300 m^2 to 0.02 km^2; e) ribbon lake on GVI IS (Figure 2.1) which spans 2.6 km and covers 0.19 km^2; and f) discontinuous supraglacial channel spanning 1.3 km and covering 0.04 km^2 near Hull Glacier. RGB composites formed from L8 tile LC08_L1GT_022114_20170111 (a,f) from 11 January 2017 and S2 tiles T18DXF_20170129 (b,c,d) from 29 January 2017 and T19DEB_20170103 (e) from 3 January 2017.**

definition of each, but rather it is an indicator that should be viewed more as a guide to the relative split between them.

2.3.3 Comparison to supraglacial features in East Antarctica

In combination with a previous study (Stokes et al., 2019), our study provides the first continent-wide assessment of Antarctic SGL. We find that, in the austral summer of 2017, the Antarctic ice sheet hosted approximately 76,000 supraglacial features comprising ∼10,500 (119.4 km^2) identified in this study together with ∼65,500 (1383.5 km^2) previously identified in East Antarctica (Figure 2.11). We estimate 1502.9 km^2 meltwater area and volume totalling 1.08 km^3 (Equation 2.6) across the entire Antarctic Ice Sheet during the month of January 2017. To ensure complete coverage, we define our WAIS longitudinal boundaries such that they cover all areas not mapped by (Stokes et al., 2019). This results in the Antarctic coastline (measuring ∼35,500 km Fretwell et al., 2013) being split approximately equally between WAIS (plus AP) and EAIS. The largest lake recorded within the EAIS dataset (on Amery IS) measures 71.5 km^2, 25 times larger than the largest WAIS lake, which is on Sulzberger IS (∼2.9 km^2). Amery IS has the highest density of supraglacial lake activity on EAIS with ∼893.3 km^2 total meltwater coverage. Amery IS measures 62,620 km^2 (Foley et al., 2013), meaning the maximum percentage coverage of supraglacial meltwater on the IS is 1.43%, while most densely populated regions in WAIS and AP, GVI, Wilkins and Bach ice shelves are between a factor of five and ten times less (with maximum percentage coverage of 0.12%, 0.13% and 0.29% respectively).

Finally, it is important to note that there are two differences between our approach and that of Stokes et al. (2019), which may cause contrasting classifications. The first is a difference in the method used to classify water pixels. Our study attempted to classify both lakes and channels using a dual NDWI approach, while Stokes et al. (2019) focused on SGL alone. The second source of difference is because of the selection of data used. Where several images are

available for specific regions, Stokes et al. (2019) sampled the image closest to the peak melt-season, i.e., mid-January, to provide a snapshot of SGL activity. Stokes et al. (2019) report around 6% of the total coastline was not mapped in their study, due largely to the cloud in the scenes. Conversely, our method quantified the maximum extent of SGL and SGC throughout the entire month to combat the effects of cloud cover and therefore was based upon a compilation of all available imagery from 1st to 31st January 2017 (and up to 10th February 2017 over the AP).

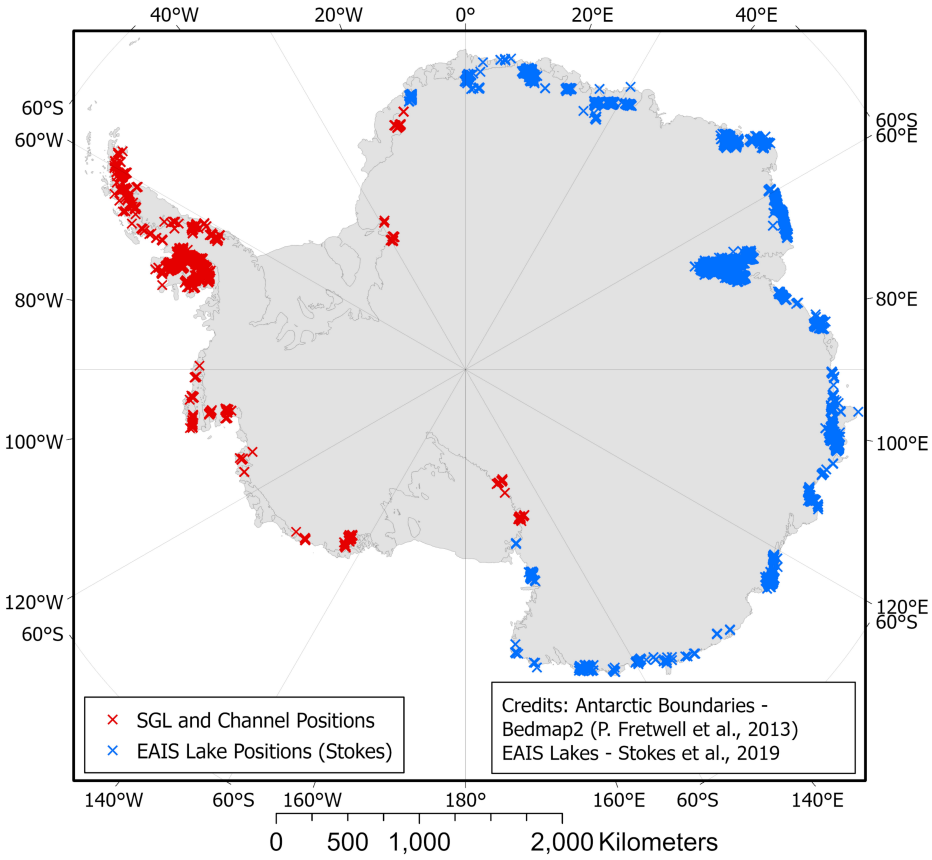


Figure 2.11: **The location of the 10,478 SGL and channels on the WAIS and AP (red crosses) and 65,459 SGL (blue crosses; mapped by Stokes et al., 2019) in January and February 2017. Antarctic boundaries are according to Bedmap2 (Fretwell et al., 2013).**

2.4 Conclusion

We have mapped, for the first time, the full extent of supraglacial hydrology on the Antarctic Peninsula and West Antarctic Ice Sheet during the 2017 melt-season using a Dual-NDWI thresholding approach. We identify 10,223 SGL and 255 supraglacial channels (10,478 features in total), which occupy a total area of 119 km^2 (114.7 km^2 lakes, 4.7 km^2 channels). For the first time, SGL have been identified on and around the margin of the Getz IS, while a significant number of hydrological features are identified on GVI, Wilkins and Bach ice shelves on the Antarctic Peninsula, Sulzberger IS in the Ford Ranges, and Pine Island Glacier in the Amundsen Sea region.

This new inventory provides a baseline Earth System dataset which, in combination with the work of (Stokes et al., 2019), represents the first continent-wide assessment of the supraglacial hydrology of Antarctica. With the operating schedules of the S2 and L8 satellites, optical data are now routinely available at weekly sampling, meaning that it is now possible to expand this study to monitor lake dynamics in near to real-time. This will allow for a better understanding of the evolution and dynamics of SGL and SGC, and how they might change in response to a warming climate. Such approaches would require advanced levels of automation, because of the scale of data required. Importantly, our study provides a high fidelity dataset that can train, calibrate, and validate such approaches.

2.5 Data & code availability

The dataset described within this study has many potential applications. As NDWI thresholds are the traditional approach to mapping SGL activity on ice sheets, the

results of this large-scale study provide a clear picture of the maximum melt extent in January of the 2017 melt season. Because of the scale of the dataset (across the WAIS), the results provide a baseline for future monitoring of supraglacial hydrology and could be used to assess regional climate model simulations of surface melting and run-off. Supervised ML algorithms require labelled data to train the algorithms. The lake and channel dataset described here will be valuable as training data for pixel-based or object-based approaches in ML, such as RF classification (M. Dirscherl et al., 2020). Others can use the dataset produced in this study to assist approaches that utilise other types of satellite data, for example, those that exploit SAR imagery but that require a priori lake distribution (A. A. Leeson et al., 2020; K. E. Miles et al., 2017).

Alongside the final map of meltwater extent, the dataset contains meltwater polygons for each sensor (S2 and L8, alongside the respective source sensor data), which form the final map and are useful for ML processes. The data's usage for training, validation or independent testing is flexible to the user's choice, providing the data are used alongside imagery from each sensor independently. It can be used entirely for training/testing or, if a user prefers, subsetted to provide independent train and test data. The final map of meltwater extent is not to be used for ML and as such does not contain the predictor data.

The code used to produce the supraglacial hydrology features is written in Python and can be accessed on Zenodo and GitHub (D. Corr, 2021).

The mapped supraglacial lake and channel polygons are available on Zenodo (`https://doi.org/10.5281/zenodo.5642755`, D. Corr et al., 2021) as digital Geographic Information Systems (GIS), shapefiles (.shp), Keyhole Markup language Zipped (.kmz) and GIS GeoJSON files. The datasets consist of the final lake and channel polygon maps for both sensors combined (i.e. our final maximum extent map of supraglacial hydrology), plus polygons for each sensor, L8 (17,571

individual polygons) and S2 (23,389 individual polygons). In addition, predictor data for each sensor (i.e., the data tiles containing all bands for S2 and L8) are provided for each of the polygons.

Additional L8 and S2 imagery are freely available at `earthexplorer.usgs.gov` and `scihub.copernicus.eu` respectively. Scripts for downloading the data were extracted from GitHub (Hagolle, Olivier, 2014) and (Hagolle, Olivier, 2015), however, with changes to data structure on both repositories, these scripts may no longer be effective. Alternatively, imagery is available to download from Google Cloud Storage using Python scripting (Nunes, Vasco, 2016).

Chapter 3

Supraglacial hydrology coverage has increased on the Greenland ice sheet over the last decade

Diarmuid Corr[1], Amber Leeson[1,2], Malcolm McMillan[1,3], Ce Zhang[1,2,4], Emily Glen[1], Jennifer Maddalena[1], and Laura Melling[1].

[1]Centre of Excellence in Environmental Data Science, Lancaster Environment Centre, Lancaster University, Lancaster.
[2]Data Science Institute, Lancaster University, Lancaster.
[3]UK Centre for Polar Observation and Modelling, Lancaster University, Lancaster.
[4]UK Center for Ecology & Hydrology, Library Avenue, Lancaster.

Correspondence:
Diarmuid Corr (dcorr103@gmail.com, d.corr@lancaster.ac.uk)

The following work is prepared as a manuscript to be submitted for publication as: "Supraglacial hydrology coverage has increased on the Greenland ice sheet over the last decade".

This chapter outlines the work carried out as part of the ESA funded project 4DGreenland, in which we evaluated supraglacial hydrology on the Greenland Ice Sheet on a decadal scale. This forms the first Greenland-wide assessment to help quantify the mass balance of the Greenland Ice Sheet and its contribution to global sea level rise. I apply a RF classification and radiative transfer model to satellite images, mapping the extent and depth of hydrology on the ice sheet.

DC developed the code, carried out the main body of work, and drafted this paper. AL, MM, JM, and CZ provided supervision and contributed extensively to the science, technical details, and structure of this paper. EG and LM conducted data processing, contributed to the creation of training data, and contributed to methodological development. All authors contributed to the manuscript text.

Abstract

Supraglacial hydrology (SGH) describes the study of the complex system of interconnected slush, lakes, and channels on the surface of a glacier or ice sheet. These features dictate the storage and transport of meltwater atop the ice sheet system; meltwater which has a significant impact on global sea levels. In this novel study, we quantify the extent and volume of supraglacial meltwater throughout the entire Greenland Ice Sheet, exploring meltwater evolution over the past decade, at both high spatial (10-30 m) and temporal (monthly) resolution. We adapt and validate a random forest (RF) algorithm to delineate supraglacial meltwater using 144000 Sentinel-2 and Landsat-8 images, acquired between May 1st and September

30th for each year between 2014 and 2022. A radiative transfer model is used to estimate the volume of surface water present on the Ice Sheet in monthly snapshots. We find a persistent strong seasonal cycle in SGH behaviour, with meltwater generally present in May, (extent and volume) peaking in July/August, and surface water beginning to refreeze mostly in September. Throughout the melt season, meltwater drains atop the ice sheet through channels and by moulins or cracks to the sub- and englacial environments. Using this dataset, we determine that there has been a positive trend in the extent and volume of SGH features over the past decade, and although interannual variability is high, this suggests that temperature increases linked to climate change are leading to greater supraglacial hydrology on the Greenland Ice Sheet. Our analysis reveals a significant increase in total meltwater over drainage basins in the north, east, and south of the ice sheet; while basins across the entire ice sheet contain meltwater in more regions than at the beginning of the decade. We observe significant seasonal trends in the mean firn air content and mean ice slab content in many of the Zwally drainage basins that experience a significant increase in supraglacial hydrology, implying that a reduction in firn air content and an increase in ice slab content could be a potential source of increase in SGH activity.

3.1 Introduction

Mass loss from the Greenland (GrIS) and Antarctic (AIS) ice sheets is predicted to be the dominant contribution to global sea level rise in coming decades (Fox-Kemper et al., 2021). According to models and observations, the ice sheets are undergoing rapid changes in response to increases in temperatures linked to climate change, which are projected to continue for at least several decades, even if global temperatures stabilise (Bamber et al., 2018; Lenaerts et al., 2015, 2019;

Otosaka et al., 2022; Shepherd et al., 2012, 2018, 2020; B. Smith et al., 2020; Trusel et al., 2018). Land ice locked in the ice sheets of Greenland (7.4 m equivalent sea level rise) and Antarctica (58.3 m equivalent sea level rise) has the potential to increase global sea level by 65.7 m and is the largest potential source of future sea level rise (Fox-Kemper et al., 2021; Morlighem et al., 2017; D. Vaughan et al., 2013). The total mass of ice on the GrIS, in particular, has decreased considerably in the last few decades (Madsen et al., 2022; Moon et al., 2020; Otosaka et al., 2022; Shepherd et al., 2020; Slater et al., 2021). Almost 60% of the GrIS mass loss is attributed to the difference between snow accumulation and surface melt, and the rest is associated with dynamic mass loss, as a result of accelerated ice flow towards the ocean (M. R. van den Broeke et al., 2016).

The supraglacial hydrological network of Greenland is a complex system of interconnected slush, lakes, and rivers/streams on the surface of the ice sheet. The network fluctuates, forming throughout the summer melt period and typically reaching maximum capacity in July/August (Otto, Holmes, and Kirchner, 2022; K. Yang et al., 2021). Meltwater is stored in supraglacial lakes (SGLs) in topographic depressions on the surface of the ice sheet. Some lakes form early in the melt season and persist for months or over winter (Koenig et al., 2015; Lampkin et al., 2020; Wendleder et al., 2021; K. Yang et al., 2021), the presence of which decreases the albedo of the ice sheet and increases the absorption of solar radiation (Leidman et al., 2021). Many lakes drain laterally into supraglacial rivers that transport meltwater across the surface over timescales of weeks to months (L. C. Smith et al., 2015; Tedesco et al., 2013). Other meltwater enters cracks and crevasses and drains into sub- or englacial environments through hydrofracturing (McGrath et al., 2012; van der Veen, 2007) or vertical drainage (K. E. Alley et al., 2018) over timescales from hours to days (Das et al., 2008). The injection of liquid water into the englacial ice sheet system leads to the warming of the ice through cryohydrologic warming

(T. Phillips, Rajaram, and Steffen, 2010). Cryohydrologic warming is theorised to reduce ice viscosity and thus contribute to faster ice flow (T. Phillips, Rajaram, and Steffen, 2010; T. Phillips et al., 2013). Vertical SGL drainage occurs when the water pressure exceeds the fracture strength of the ice (K. E. Alley et al., 2018), creating moulins when crevasses propagate completely through the ice sheet (Das et al., 2008; McGrath et al., 2012) that have sufficient capacity to drain lakes over hours or days (Das et al., 2008). Rapid injection of water to the bed acts to modulate the flow of grounded ice by overloading the capacity of the subglacial system, increasing the subglacial water pressure, and decreasing friction, resulting in increased basal sliding (Tedesco et al., 2013). Meltwater that is not stored in lakes on, in, or under the GrIS (Lampkin et al., 2020; K. Yang et al., 2021), refreezes at the end of the melt season (Koenig et al., 2015; Lampkin et al., 2020) or is transported to the ocean as runoff, through rivers and waterfalls (L. C. Smith et al., 2015, 2021) directly contributing to sea level rise and a loss of the mass balance of the GrIS.

Supraglacial hydrology features (SGHF) can be distinguished by the eye in optical satellite imagery, due to their distinctive blue colour against the surrounding white, or grey, ice. However, given that there are many tens of thousands of these features (D. Corr et al., 2022; Stokes et al., 2019; K. Yang et al., 2021), and they appear in many thousands of satellite images, automated approaches to mapping these features are commonly used. The standard approach to mapping SGHF combines spectral thresholding (typically the Normalised Difference Water Index, NDWI) with extensive manual post-processing (D. Corr et al., 2022; Fitzpatrick et al., 2014; M. Moussavi et al., 2020; Stokes et al., 2019; Williamson et al., 2018; K. Yang and L. C. Smith, 2013). Manual post-processing is required to remove areas of shadow, cloud, rock, crevassing, and blue ice that are spectrally similar to meltwater and sometimes misclassified as water (D. Corr et al., 2022). However, given the volume of satellite data now available, methods that

require manual intervention are not feasible for repeat monitoring of surface hydrology at a continental scale. Instead, machine learning (ML) algorithms such as random forests (RF), Convolutional Neural Networks (M. Dirscherl et al., 2021) and k-means clustering (Halberstadt et al., 2020) offer alternative methods to tackle this problem, but have been underutilised until now. In particular, the implementation of ML algorithms has shown promise in limited area studies on the GrIS, AIS, and alpine glaciers (Dell et al., 2022; M. Dirscherl et al., 2020; Hu et al., 2022; Wangchuk and Bolch, 2020). Here, we use 138963 Sentinel-2 (S2) and 5708 Landsat-8 (L8) satellite images to map SGHF on the surface of the Greenland Ice Sheet, employing a random forest algorithm, trained and validated on representative samples that account for the spatial and temporal variation of the state of the GrIS. Using a radiative transfer approach (Banwell et al., 2014; Pope et al., 2016; Sneed and Hamilton, 2007; Tedesco and Steiner, 2011; Williamson et al., 2018), we calculate the decadal evolution in the volume of surface meltwater of Greenland.

Most studies dedicated to mapping SGHF on the GrIS have aimed to develop our process-level understanding of the seasonal evolution of supraglacial lakes and have focused on one or more study sites associated with specific glaciological basins (Y.-L. Liang et al., 2012; McMillan et al., 2007; Sundal et al., 2009; Williamson et al., 2018). More widespread studies that consider a larger area have also produced interesting findings in terms of both seasonal evolution and the drainage behaviour of lakes (Selmes, Murray, and James, 2011) and in terms of interannual and long-term evolution in their spatial coverage (Howat et al., 2013; Igneczi et al., 2016). The only known systematic map of supraglacial lakes on an interannual timescale, covering the entire ice sheet, is available as yearly snapshots for the period 2016-2018 Hu et al., 2022. Our study explores the evolution of meltwater at high spatial (10-30 m) and temporal (monthly) resolution for an entire decade, quantifying the extent and volume of supraglacial meltwater across the entire

Greenland Ice Sheet. This analysis is important to understand how the GrIS responds to climate change and what the implications are for future mass imbalance on the ice sheet scale.

3.2 Data & methods

Here, we describe the data and workflow (Figure 4.3) to map the SGHF and estimate volume, on a continental scale. We detail the input data, the RF algorithm, the generation of training data, the validation, and the depth calculations for volume estimates.

3.2.1 Data

We map SGHF on the GrIS in 138963 scenes acquired by Sentinel-2 between 2017 and 2022, and in 5708 scenes acquired by Landsat-8 between 2014 and 2016. We omit 350777 S2 (72% of all scenes) and 6418 L8 (53% of all scenes) scenes with cloud cover greater than 30% and solar elevation angles lower than 20°. A subset of the Sentinel-2 (91) and Landsat-8 (71) scenes is used in the development and training of the final algorithm (Table B.1).

3.2.1.1 Sentinel-2

S2 data are freely available as Top-of-Atmosphere (TOA) reflectance data from the Copernicus Open Access Hub: `https://scihub.copernicus.eu/` and from Google Cloud Storage. S2 bands 2 (blue), 3 (green), 4 (red), and 8 (near infrared - NIR) exist at a resolution of 10 m, the highest spatial resolution acquired by the sensor. Bands 5, 6, 7, 8A (vegetation red edge 1-4), 11 (short-wave infrared, ($SWIR_1$)) and 12 ($SWIR_2$) are acquired at a coarser resolution of 20 m. Bands 1 (coastal aerosol), 9

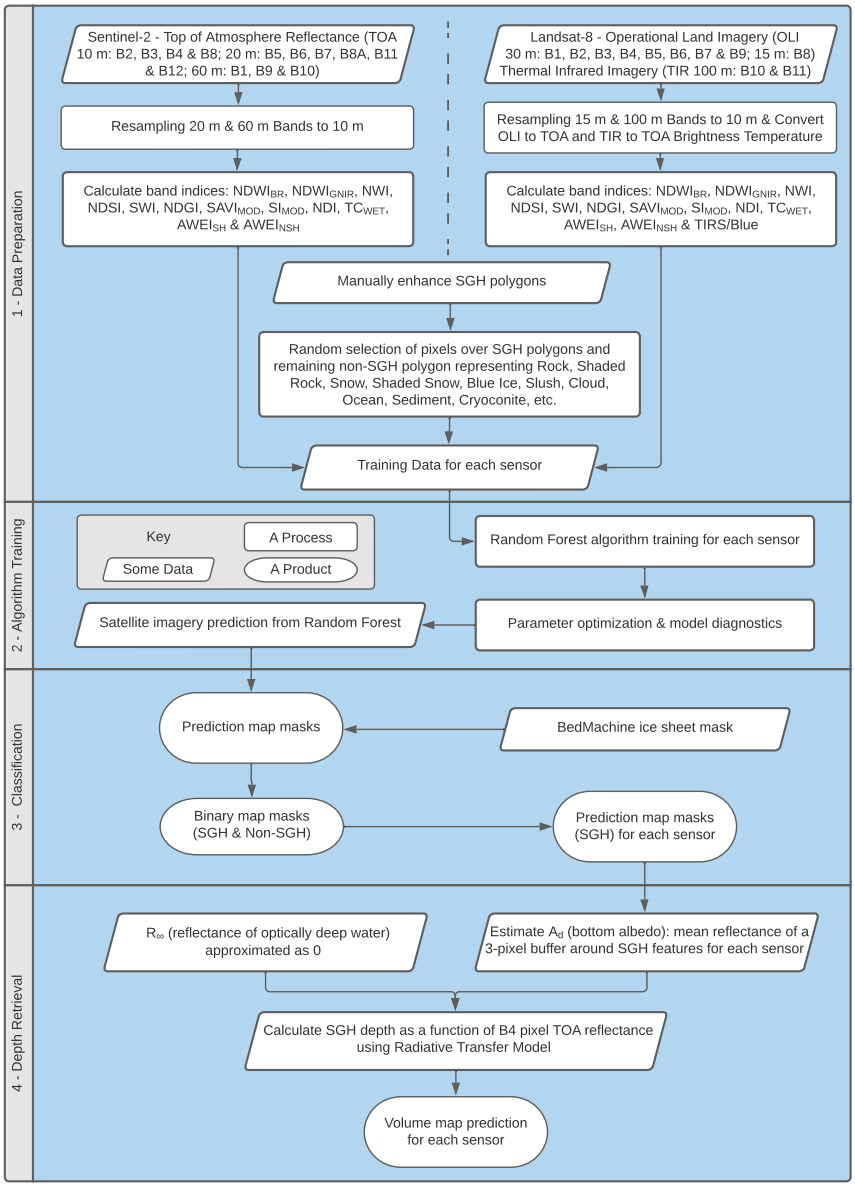


Figure 3.1: **S2 and L8 processing chains which result in SGHF extent and volume maps.**

(water vapour), and 10 ($SWIR_{cirrus}$) exist at 60 m. Bands with coarser resolution than 10 m are resampled to 10 m using nearest-neighbour interpolation for

consistency with the red, green, and blue (RGB) and NIR bands (D. Corr et al., 2022; Williamson et al., 2018). S2 pixel values represent TOA reflectance units $\times$10,000 and are known as TOA reflectance integers (reflectance$\times 10^4$). The Sentinel-2 mission currently comprises two satellites, Sentinel-2A and Sentinel-2B. Each satellite has a revisit period of 10 days; however, since they are phased 180°from each other, an on-the-ground repeat cycle of 5 days is possible depending on the cloudiness and solar elevation angle of a scene.

3.2.1.2 Landsat-8

L8 data are freely available from the United States Geological Survey (USGS) Earth Resources Observation Science (EROS) Centre (`https://eros.usgs.gov`) and Google Cloud Storage. Level-1 L8 data, comprising quantised and calibrated scaled digital numbers (DN), are converted to TOA reflectance or brightness temperature values (Chander, Markham, and Helder, 2009). L8 bands 1 (coastal aerosol), 2 (blue), 3 (green), 4 (red), 5 (NIR), 6 ($SWIR_1$)), 7 ($SWIR_2$), and 9 ($SWIR_{cirrus}$) are used in their native resolution of 30 m. The bands 8 (panchromatic, 30 m), 10 (thermal infrared ($TIRS_1$), 100 m resolution) and 11 ($TIRS_2$, 100 m resolution) are resampled to 30 m using nearest-neighbour interpolation as with the S2 data. Landsat-8 has a revisit time of 16 days, more than three times that of the Sentinel-2 constellation. This results in a lower temporal sampling of the GrIS using L8 imagery.

3.2.1.3 Band indices

Band indices are assembled from the S2 and L8 TOA bands (Table 3.1) and combined with the individual spectral bands to form a composite tiff (25 bands). Indices used here include $NDWI_{BlueRed}$ (a), $NDWI_{GreenNIR}$ (b), New Water Index (NWI, (c) Feng, 2009), Normalized Difference Snow Index (NDSI, (d) named in places as Modified NDWI, Hall, Riggs, and Salomonson, 1995; Xu, 2006), Soil/Water Index (SWI, (e)

M. Dirscherl et al., 2020), Normalized Difference Glacier Index (NDGI, (f) Keshri, Shukla, and Gupta, 2009), modified Soil-Adjusted Vegetation Index ($SAVI_{mod}$, (g) Huete, 1988), modified Shadow Index (SI_{mod}, (h) H. Li et al., 2016), Tasselled Cap for wetness (TC_{wet}, (i) Kauth and G. Thomas, 1976; Schwatke, Scherer, and Dettmering, 2019), Automated Water Extraction Index including shadow ($AWEI_{sh}$, (j)) or with shaded area removal ($AWEI_{nsh}$, (k) Feyisa et al., 2014), Normalized Difference Index (NDI, (l) M. Dirscherl et al., 2020), and a combination of the thermal infrared sensor (TIRS) and blue band (m) in L8 (M. Moussavi et al., 2020).

Table 3.1: **Indices calculated from the S2 and L8 sensors, used in the random forest algorithm.**

Index Name		Equation	Terrestrial Feature(s)	Sensor(s)	Reference
$NDWI_{BlueRed}$ (a)	=	$\frac{Blue - Red}{Blue + Red}$	Surface Water	S2 & L8	M. Moussavi et al., 2020
$NDWI_{GreenNIR}$ (b)	=	$\frac{Green - NIR}{Green + NIR}$	Surface Water	S2 & L8	D. Corr et al., 2022
NWI (c)	=	$\frac{Blue - (NIR+SWIR_1+SWIR_2)}{Blue + (NIR+SWIR_1+SWIR_2)}$	Surface Water, Shadow	S2 & L8	Feng, 2009
NDSI (d)	=	$\frac{Green - SWIR_1}{Green + SWIR_1}$	Rock, Sediment, Ice	S2 & L8	Hall, Riggs, and Salomonson, 1995; Xu, 2006
SWI (e)	=	$\frac{Blue - SWIR_1}{Blue + SWIR_1}$	Rock, Sediment, Ice	S2 & L8	M. Dirscherl et al., 2020
NDGI (f)	=	$\frac{Green - Red}{Green + Red}$	Rock, Sediment, Ice	S2 & L8	Keshri, Shukla, and Gupta, 2009
$SAVI_{mod}$ (g)	=	$2 \times \frac{Green - NIR}{1 + Green + NIR}$	Rock, Sediment, Ice, Shadow	S2 & L8	Huete, 1988
SI_{mod} (h)	=	$\frac{Blue - NIR}{Blue + NIR}$	Shaded Ice/Snow, Shadow	S2 & L8	H. Li et al., 2016
TC_{wet} (i)	=	$(0.1509 \times Blue) + (0.1973 \times Green)$ $+ (0.3279 \times Red) + (0.3406 \times NIR)$ $-(0.7112 \times SWIR_1) - (0.4572 \times SWIR_2)$	Surface Water	S2 & L8	Kauth and G. Thomas, 1976 Schwatke, Scherer, and Dettmering, 2019
$AWEI_{sh}$ (j)	=	$Blue + (2.5 \times Green)$ $-(1.5 \times (NIR + SWIR_1)) - (2.5 \times SWIR_2)$	Surface Water, Shadow	S2 & L8	Feyisa et al., 2014
$AWEI_{nsh}$ (k)	=	$4 \times (Green - SWIR_1)$ $- (0.25 \times NIR) - (2.75 \times SWIR_2)$	Surface Water	S2 & L8	Feyisa et al., 2014
NDI (l)	=	$\frac{Green - Blue}{Green + Blue}$	All Features	S2 & L8	M. Dirscherl et al., 2020
TIRS_Blue (m)	=	$\frac{TIRS_1}{Blue}$	Cloud	L8	M. Moussavi et al., 2020

3.2.1.4 Training data

Machine Learning algorithms require comprehensive, accurate training and validation data sets. Due to the lack of in situ data, we generated our training and validation data from satellite imagery. Specifically, we use images that 1) cover a range of glaciological and climatological environments (i.e., acquired in the Watson River catchment, west Greenland, northwest Greenland, east Greenland, and the North East Greenland Ice Stream (NEGIS) basin) and stretch from around the ice sheet margin to regions more than 200 km inland (Figure 3.2), 2) cover a broad temporal range, that is, are sampled at peak meltwater (July-August) and for periods with less meltwater (May and September) across multiple melt seasons, and 3) within which a range of surface types are visible. Training data are extracted at points that are labelled as surface water or non-surface water. The labels of the surface water training data are determined by manually enhancing the maps generated using NDWI algorithms (D. Corr et al., 2022) and an approach used to enhance the delineation of supraglacial channels (K. Yang et al., 2019a) from true-colour images, resulting in polygons of areas of surface water. Non-surface water training data is created by identifying pixel types that are spectrally similar to surface water, including slush, blue ice, rock, snow/ice, cloud, cloud shadows, sediment, and cryoconite, and manually creating polygons of each type. We then randomly subsample pixels within surface water and non-surface water polygons to reduce user bias in pixel identification. Around 5000 (2500 water if available, 2500 non-water) pixels are randomly selected within each of the 91 S2 and 71 L8 scenes that comprised the training dataset.

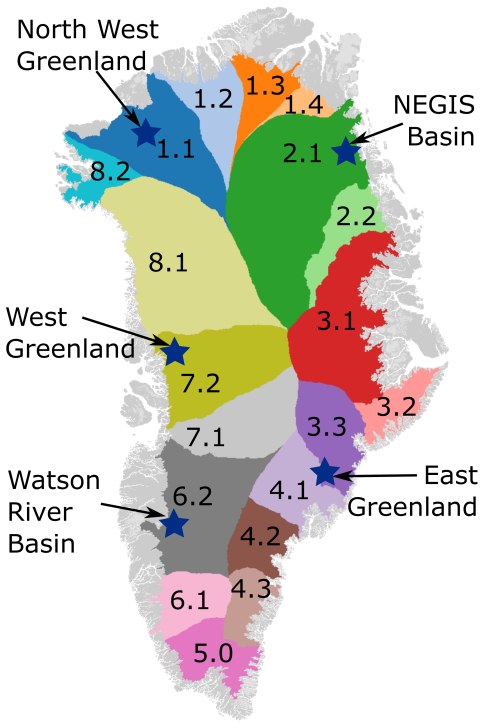


Figure 3.2: **Drainage basins on the GrIS (Zwally et al., 2012). Locations marked with a star are used in the selection and creation of training data for Landsat-8 and Sentinel-2 algorithms. The ice sheet and island boundaries are according to BedMachine v4 (Morlighem et al., 2017). The GrIS is divided into nineteen drainage basins (Zwally et al., 2012), (Figure 3.2). The 19 basins are divided into 8 separate regions (1 = North, 2 = Northeast, 3 = East, 4 = Southeast, 5 = South, 6 = Southwest, 7 = West, 8 = Northwest).**

3.2.2 Classification of supraglacial hydrology using a random forest algorithm

Random forest (RF) is a supervised learning algorithm that requires labelled training data (Breiman, 2001). It is made up of a group (forest) of unrelated decision trees. After a given number of trees are generated, the most popular classification is

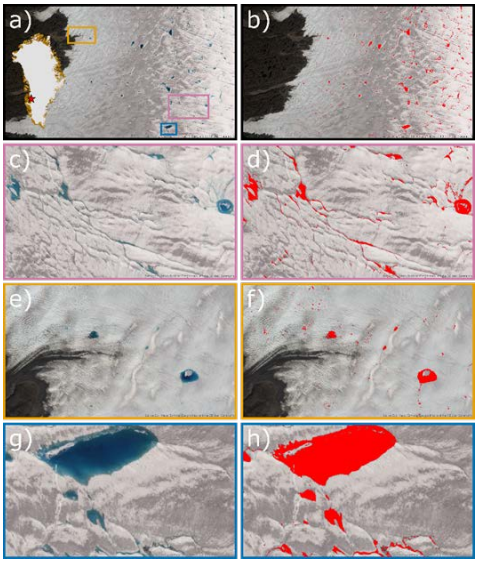


Figure 3.3: **Examples of our random forest classification of supraglacial hydrology on the Watson River Region of the Greenland Ice Sheet from Sentinel-2 satellite imagery (tiles: T22WEA_20180710 and T22WEV_20180710). True colour images of the supraglacial system are shown in figures a (regions of each of the coloured outlines), c (supraglacial lakes interlinked by supraglacial channels, pink), e (supraglacial lakes near the ice sheet margin, orange), and g (a large supraglacial lake, blue), while b, d, f, and h are the same images with the regions classified as surface water overlaid in red.**

determined. RF resists overfitting, which can be a major source of error in ML processes, is efficient, overcomes missing data, and is flexible depending on the application (Breiman, 2001). For large data sets, such as the S2 and L8 archives used in this study, the main memory requirement is the storage of the data itself. More details on RF algorithms are presented in the Appendix.

Due to differences in the resolution and reflectance values of each sensor, the random forest algorithms are trained independently (for S2 and L8), using the Scikit-learn RandomForestClassifier module (Pedregosa et al., 2018), with

hyperparameter optimisation conducted for each case. Hyperparameter optimisation is designed to determine optimum values for the number of trees in the forest (40-S2 and 250-L8), the maximum depth of a tree (10 and 10), the minimum number of samples required to split an internal node (53 and 45), the minimum number of samples required at a leaf node (23 and 15), the maximum number of features to consider when applying the splits (no maximum in either case), and whether bootstrap methods are used when building trees (yes for both S2 and L8). For each hyperparameter, we optimise the parameters using an AUC-ROC (Area Under Curve-Receiver Operating Characteristics) curve approach (Song, 2015), cross-validating with 3 folds to ensure rigour.

Once trained, the RF algorithm is applied to unseen data, to produce a binary pixel-by-pixel classification (water, non-water) of supraglacial hydrology (Figure 3.3), for each scene in the dataset. Features smaller than two pixels (200 m^2 in S2 and 1800 m^2 in L8 images) are removed from the datasets, as such features are considered below the detection limit of the sensors and are more likely to be areas of slush or noise than open water (Pope et al., 2016). The classified images are then mosaicked to form monthly snapshots (May through September) of the SGH coverage (Section 2.5). Manual post-processing is carried out over each mosaic to remove obvious misclassifications, including some remaining areas of cloud cover and shaded regions around crevasses and fjords. 6-8% of the final mosaics were deemed misclassifications and removed from the data set by manual post-processing.

3.2.3 Depth retrieval using radiative transfer model

SGHF depths are calculated pixel by pixel using a radiative transfer model (RTM, Equation 1.3) (Banwell et al., 2014; Pope et al., 2016; Sneed and Hamilton, 2007;

Tedesco and Steiner, 2011; Williamson et al., 2018). Specifically, depths are estimated from the red and green band reflectance values, using the same S2 and L8 images that were used to map the extent of the lake.

3.2.4 Validation of the algorithms

As overall accuracy is not an ideal metric to measure the performance of an algorithm containing imbalanced data, we calculate the precision, recall, and F1 score of the confusion matrices to determine the accuracy and reliability of the algorithms. The F1 score (F score or F measure) is the harmonic mean of precision and recall; that is, it conveys the balance between the error caused by false positives (precision) and the error caused by false negatives (recall), taking into account a poor performance of the minority class. Overall accuracy is reported for completeness.

Table 3.2: **Validation metrics for the cross-validation of RF algorithm applied to Landsat-8 and Sentinel-2 training data sets. Tests are repeated three times for each sensor, with different randomly selected data from the Watson River catchment, west Greenland, northwest Greenland, east Greenland, and NEGIS on each occasion, to reduce bias in the results.**

Sensor Test No.	Precision	Recall	**F1-Score**	**Overall Accuracy [%]**
Landsat-8 Test 1	0.90	0.99	0.94	99.0
Landsat-8 Test 2	0.91	0.98	0.94	99.0
Landsat-8 Test 3	0.88	0.97	0.92	98.6
Sentinel-2 Test 1	0.87	0.99	0.93	97.0
Sentinel-2 Test 2	0.87	0.98	0.92	96.9
Sentinel-2 Test 3	0.87	0.98	0.92	96.9

An 80:20 train-test split is applied to the training/testing data set, and since they are imbalanced (more non-SGHF pixels than SGHF pixels), we perform stratified random sampling to reduce bias. Stratified random sampling selects approximately the same percentage of samples of each target class, ensuring that the minority class is represented in the training process. The tests, carried out on data from the Watson River catchment, west Greenland, northwest Greenland, east Greenland, and NEGIS (Table B.1), are repeated three times with different randomly selected data on each occasion (Table 3.2). The metrics (precision, recall, F1 score, and overall accuracy) are similar in all tests. Although the precision, and therefore F1 score, suggest better performance for L8 over S2, the S2 sensor is attempting to resolve features at a much higher resolution and therefore smaller water bodies with more subjective boundaries and a more challenging classification. However, the precision is consistent with 0.87 for S2 and varies between 0.88 and 0.91 for the L8 tests. The recall varies from 0.97 to 0.99 across the six tests, which we note outperforms the recall reported for the standard spectral thresholding approach, which reported mean recall values of 0.82 (D. Corr et al., 2022) on the Antarctic Ice Sheet. The F1 score, which is marginally higher on average for L8, ranges from 0.92 to 0.94. The general accuracies of 99% (L8) and 97% (S2) demonstrates the discrepancy between the general accuracy and a metric that punishes poor performance of the minority class. The individual confusion matrices for each test are presented in the Appendix (Tables B.3 and B.4).

The performance of the RTM with each band was evaluated against lake depth measured by ICESat-2 transects over four lakes in the Watson region (L. Melling et al., 2023). The transects covered 400 data points, and the sum of the depths was calculated by integrating the depth values over the transect length. Depths were calculated for the three methods: ICESat-2, RTM green band and RTM red band (Table 3.3). For the red band, lake depths were found to be underestimated by 33.1%,

while the green band was overestimated by 69. 7%. As such, we use the red band in our study to calculate lake volume, but note that our estimates probably constitute a lower bound.

Table 3.3: **Comparison of summed depths calculated by the green and red band RTM with depth derived from ICESat-2.**

ICESat-2 Depth [m^2]	Green RTM Depth [m^2]	Red RTM Depth [m^2]	% Diff. Green	% Diff. Red
1224.74	2078.15	819.88	+69.7	-33.1

3.2.4.1 Comparison of results between Landsat-8 & Sentinel-2

Since S2 data is only available from 2017, it is useful to be able to use L8 to extend the record back in time (to 2014). However, with a resolution of 30 m, the L8 sensor is unable to resolve streams and lakes smaller than 1800 m^2, while S2 can resolve those as small as 200 m^2 (Figure 3.4). To estimate the effect this has on the final mosaicked data sets, we have conducted a comparison of the estimated extents and volumes from each sensor for the 2017 melt season and compare the results. In each melt season, S2 identifies a much higher lake coverage than L8 (Table 3.4). The greatest percentage differences occur as the melt season progresses, suggesting that the lower spatial resolution of the L8 sensor may affect the classification of drained or refrozen lakes. As greater temporal sampling results in an increased probability that peak meltwater extent (and/or volume) is captured, we assume that L8's lower temporal sampling has a negative impact on the reliability of our assessment. We calculate the mean percentage differences for the area and volume estimates.

$$\% \text{ Difference} = 100 \times \frac{|\text{S2 Value - L8 Value}|}{\left(\frac{\text{S2 Value + L8 Value}}{2}\right)} \tag{3.1}$$

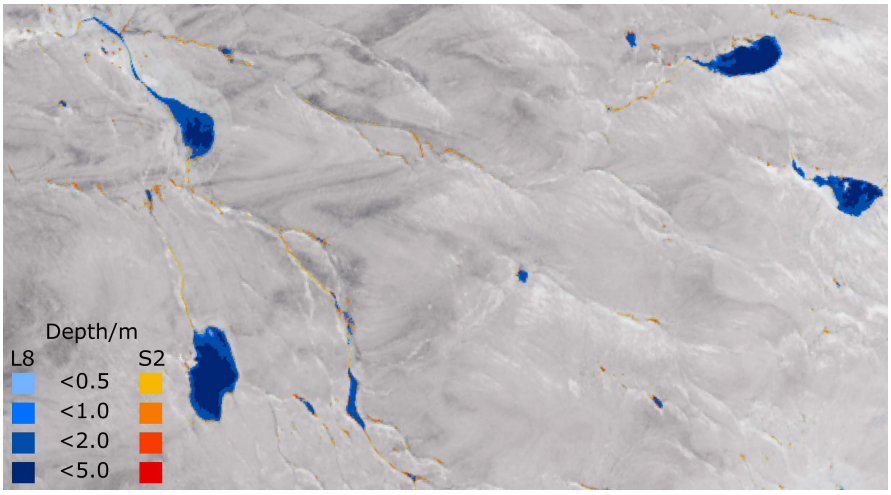


Figure 3.4: **Depth estimates from S2 (T22WEV_20200712T202317 - reds) and L8 (LC08_L1TP_008013_20200712_20200722_01_T1 - blues) sensors on an RGB composite of the same L8 tile in the Watson region on 12 July 2020.**

Assuming that the S2 data are more reliable, we calculate the area and volume for each sensor and use the difference to 'correct' the L8 estimates between 2014 and 2016. Therefore, we find that L8 underestimates area by 78.5% and volume by 50.2% when compared to S2, and therefore we scale the L8 estimates to area and volume by these values, respectively.

Table 3.4: **Area and volume comparisons results obtained by Sentinel-2 and Landsat-8 sensors for each month of the 2017 melt season.**

Date	S2 Area [km^2]	S2 Volume [km^3]	L8 Area [km^2]	L8 Volume [km^3]	% Difference Area	% Difference Volume
May 2017	3983.4	1.62	3188.0	1.17	22.2	32.4
June 2017	7127.2	3.68	3655.0	2.32	64.4	45.5
July 2017	29320.1	13.09	10092.0	7.03	97.6	60.2
August 2017	17932.9	9.31	5981.9	4.77	99.9	64.6
September 2017	2398.6	1.26	732.5	0.77	106.4	48.4
Mean	-	-	-	-	78.5	50.2

3.2.4.2 Transferability of the RF algorithm

To assess the spatial and temporal transferability of the algorithm, 40 scenes in total are assessed (Table B.1). This included data from 2018 (May 26, July 10, August 19, September 13) and 2019 (May 29, July 8, August 4, September 1) in the Watson study area (Figure 3.2) and data from 2019 (1 June, July 13, August 13, and September 3) in the NEGIS study area (Figure 3.2). In each case, the tests were conducted by training the RF algorithm on a subset of the data; for example, trained on all 2018 data and tested on the unseen 2019 data or trained on all data from May, July, and August and trained on the unseen September data.

The spatial transferability assessment evaluated the performance of the RF algorithm when trained on data from the Watson River and tested on unseen data from the NEGIS. NEGIS is a prominent ice flow feature, draining 12% of the ice sheet interior through three large outlet glaciers in northeast Greenland (Larsen et al., 2018). Although believed to be historically stable, since 2006 it has undergone marked thinning and retreat (Khan et al., 2014, 2022), with evidence of an inland expansion and increases in the total area of SGLs during this period (Turton et al., 2021). Compared to other basins, the Watson River basin in southwest Greenland drains a much smaller area of the ice sheet (0.7%, Lindbäck et al., 2015; van As et al., 2018) and hosts a vast interconnected system of supraglacial lakes and channels (A. A. Leeson et al., 2015; K. Yang et al., 2021). The southwest region typically hosts the highest number of SGLs (Hu et al., 2022).

Our validation exercise shows that the methods are transferable in space and in time (Table 3.5) and that this transferability extends similarly to the high (2019) and low (2018) melt seasons. The algorithm performs best when trained on data from May, August, and September and tested on unseen data from July. The algorithm's performance drops in May and September. In these months, the boundary between

open water and slush is less well defined, leading to difficulty in defining what is open water and what is slush. These results with consistent and high-performing F1 scores demonstrate that the method is a valid approach. For the individual confusion matrices for each test, see the Appendix (Tables B.5:B.7). The method can classify supraglacial hydrology effectively when tested on unseen data in both space and time. This is reassuring for an operational setup and indicates that the method is reliable for roll-out on the ice sheet scale for future melt seasons.

Table 3.5: **Yearly, seasonal and spatial transferability metrics (precision, recall, F1 score and accuracy) for the algorithm tested on Sentinel-2 data from 2018 and 2019 in the Watson River and NEGIS regions of Greenland.**

Transferability Test	Precision	Recall	**F1-Score**	**Overall Accuracy [%]**
Yearly-2018	0.90	0.99	0.94	95.9
Yearly-2019	0.97	0.93	0.95	96.1
Seasonal-May	0.84	0.95	0.89	95.2
Seasonal-Jul	0.96	0.99	0.98	97.6
Seasonal-Aug	0.98	0.91	0.95	94.9
Seasonal-Sep	0.88	0.87	0.87	96.0
Seasonal-Sep 2018 & May 2019 Omitted	0.97	0.93	0.95	97.6
Spatial-NEGIS	0.97	0.94	0.96	96.6

3.2.5 Mosaicking & analysing the data sets

Next, we organise our 144,671 classified scenes into monthly (May-September) snapshots and 120 km by 120 km tiles covering the entire ice sheet (1495 for each melt season, Figure 3.5). For S2, each tile contains 144 million ($12,000 \times 12,000$) cells at 10 m resolution, whereas for L8, each tile contains 16 million ($4,000 \times 4,000$) cells at 30 m resolution.

For each tile and month on the grid, each classified image that overlaps a cell on that tile is stacked. If any cell is considered to contain surface water in any classified image, it is assigned a value of 1. Cells that are not considered to contain surface water in any classified image are assigned a value of 0. Therefore, we consider our resulting estimate to represent a maximum extent in each month, although we note that few (¡5) observations of each cell of the grid in each month. Due to temporal sampling and the impact of clouds, the true maximum extent may be higher in some cases; however, our results are a best-case approximation. We repeat the same procedure to generate annual maximum extent estimates. The same procedure is adopted to organise our SGL volume data.

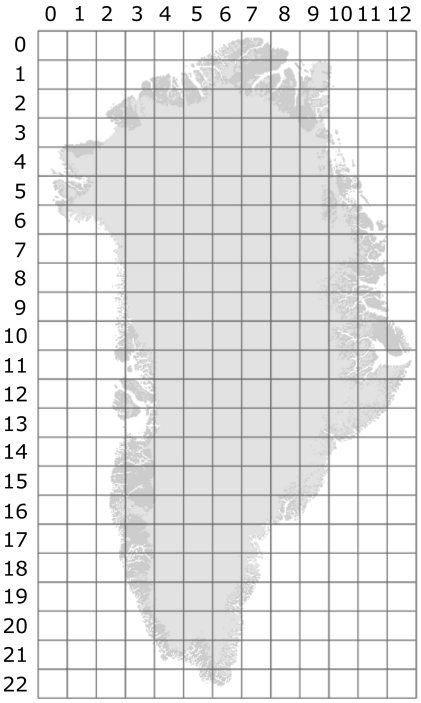


Figure 3.5: **The numbered grid which is used to divide the GrIS dataset into consistent areas measuring 120** km **by 120** km **and containing 144 million (**$12,000 \times 12,000$**) cells at 10** m **resolution for S2 data, and 16 million (**$4,000 \times 4,000$**) cells at 30** m **resolution for L8 data.**

To quantify changes in hydrological activity on the ice sheet throughout our time series, we use Pearson's r to signify the sign and significance of the decadal trend for the extent and volume of SGHF for each month (May through September), while the magnitude of the trend is obtained from a least-squares fit. The r values (and related P-values) are calculated for the monthly extents and volumes of each basin over the years (Tables B.9-B.11). The sign of r determines the direction of the trend; for example, positive numbers suggest that SGH is increasing over time. The trend is determined to be statistically significant if the corresponding P-value is less than 0.05.

3.3 Results & discussion

By applying an RF algorithm to Landsat-8 and Sentinel-2 sensor images, we map the extent and evolution of SGHF throughout the GrIS at monthly resolution from 2014 to 2022. This study evaluates supraglacial hydrological features throughout Greenland, on a decadal scale, for the first time.

3.3.1 Evolution of SGH over the Greenland Ice Sheet

To examine the evolution of SGH over the GrIS we analyse the spatial and temporal elements on the ice sheet scale. First, we examine the distribution of SGHF in space (Figure 3.6). Next, we examine changes in the monthly extent and volume of SGHF in the Zwally drainage basins over time (Figures 3.8b, B.1, B.2). We then explore the drivers of increased supraglacial hydrology (Figures 3.9a, 3.9b), before investigating the elevation at which SGH occurs (Figure 3.10).

We explore the spatial changes in the hydrological extent for each month over the years 2014 to 2022. We calculate the cumulative area of supraglacial hydrology

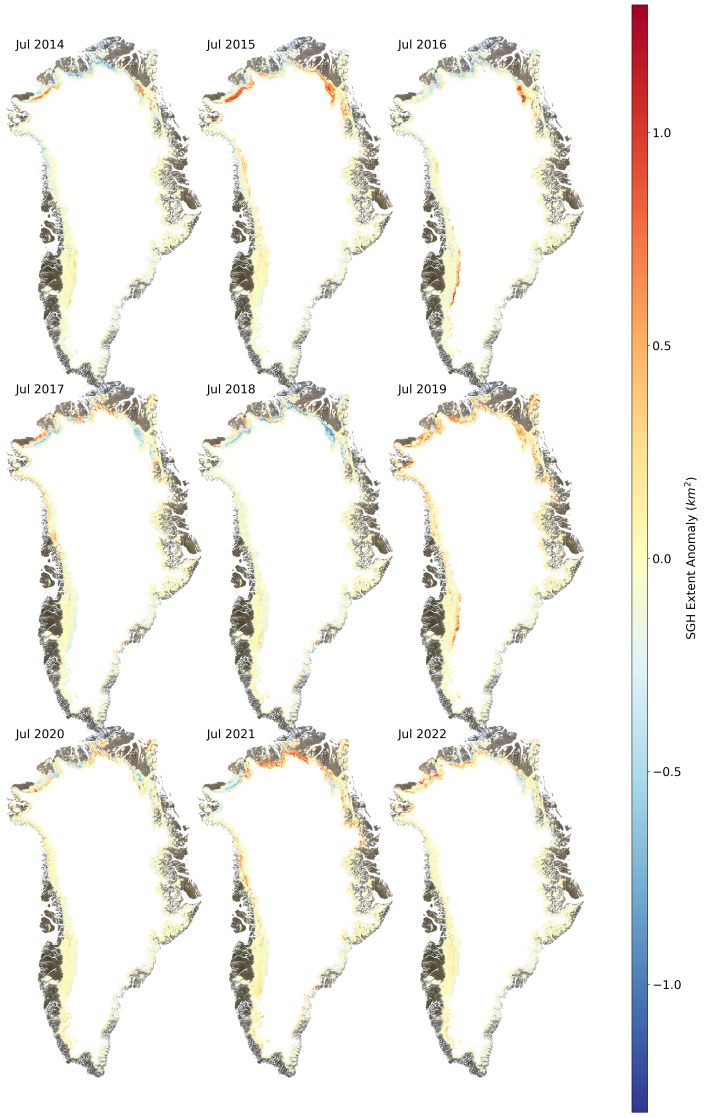


Figure 3.6: **The supraglacial hydrology extent anomaly for July compared to the mean extent in 1200 m grid cells on the Greenland Ice Sheet. The anomaly was calculated, taking the difference of the mean cell value of the 9 years in each cell from the extent value for each year. Figures for the other months, May, June, August, and September, are included in the Appendix (Figures B.3a-d).**

within 1200 m resolution cells on a monthly basis and compute the monthly means for each cell over the period 2014-2022. The anomaly for each cell is determined by subtracting the mean area of that cell in the given month from the cumulative area in that month (Figures 3.6, B.3). The extent of supraglacial hydrology varies spatially. Positive anomalies, where the extent is greater than the mean, indicate increased surface meltwater occurrence. Negative anomalies, where the extent is lower than the mean, indicate a decrease in supraglacial meltwater. This analysis reveals the complexities in evaluating trends on the ice sheet (or storage basin) scale and suggests that variations in the extent of meltwater from the high and low melt seasons may reduce the statistical significance of trends; for example, although we observe positive anomalies in the extent of the SGHF in July from 2019-2022 over much of the northern region, positive anomalies in 2015 and negative anomalies in 2018 may reduce the significance of the trend (Figure 3.6). We present the July data here because it is typically the maximum lake extent.

Almost all SGHF occurs around the margin of the ice sheet - we see very few (<0.1%) SGHF more than 150 km from the margin in all nine years studied here. The north, northeast, northwest, and southwest regions consistently contain the largest volume and extent of surface water. The east and southeast coasts have the least amount of meltwater in almost all months and years. This is likely because these areas are characterised by mountainous regions of higher altitude and slope, as well as small ablation zones due to accumulation and firn depths (B. P. Y. Noël, 2019; Turton et al., 2021). These conditions are not favourable for lake formation. We find that the extent of supraglacial water peaked in July 2019 with approximately 3.5% of the ice sheet area, while the volume peaked in 2021 with approximately 20 km^3 (Figure 3.7). This suggests that 2019 contained more shallow features, while 2021 contained deeper lakes and channels. For their respective peaks, we observed approximately 42% of lakes by area and 40% by volume in the

north, 15% and 10% in the northeast, 7% and 8% in the east, 1% and 2% in the southeast, 1% and 2% in the south, 13% and 20% in the southwest, 6% and 5% in the west, and 15% and 11% in the northwest. It is typically observed that surface meltwater on the Greenland Ice Sheet is present in May, reaches its peak in July, and starts to refreeze in September. This is in agreement with previous research (K. E. Alley et al., 2018; Das et al., 2008; McGrath et al., 2012; Otto, Holmes, and Kirchner, 2022; L. C. Smith et al., 2015; Tedesco et al., 2013; van der Veen, 2007; K. Yang et al., 2021), which has also found that lakes drain laterally across the surface and vertically to sub- or englacial environments throughout the melt season.

To examine trends in supraglacial extent and volume at different points in the melt season, we divided our analysis into 19 drainage basins (Zwally et al., 2012). At the ice sheet level, we observe statistically significant increases in the meltwater area in four of the nineteen basins and statistically significant increases in the volume of meltwater in seven of the nineteen. Geographically, these basins are located in the north, northeast, east, southwest, and south (Figures B.1a, B.1b). The signals from area and volume exhibit distinct patterns, which is intriguing given their direct correlation; for example, our research indicates that the total monthly SGHF area (km^2) is directly proportional to the total monthly SGHF volume (km^3) on the Greenland Ice Sheet (SGHF Area = 2530.7 $\times$ SGHF Volume, Figure B.5).

Although storage basin 6.2 in southwest Greenland dominates the maximum meltwater extent and almost every monthly extent, interestingly, we do not see a significant trend in the extent of SGHF or volume estimates in the time series (Figures B.1a, B.1b). Many studies have been conducted in west Greenland due to the high levels of supraglacial meltwater, however, our results emphasize the importance of continental studies as this region does not show any statistically significant trends. A similar phenomenon is observed throughout the west and

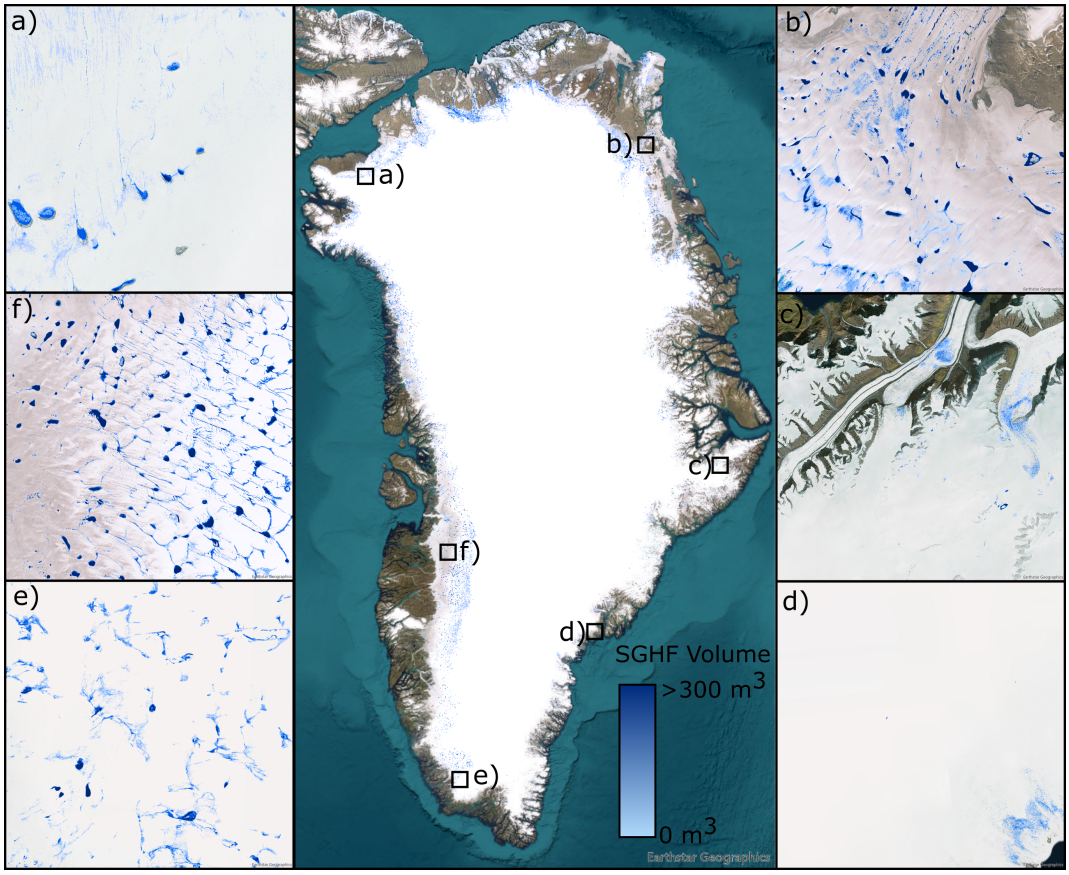
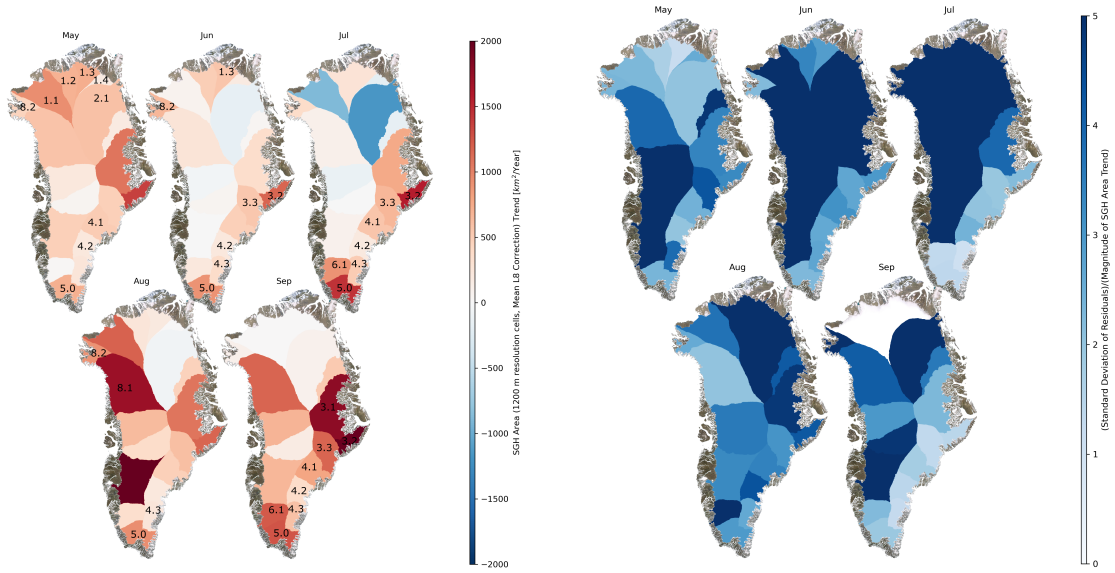


Figure 3.7: **Supraglacial meltwater volumes on the Greenland Ice Sheet during July 2021, and detailed views across the ice sheet in the northwest (a), northeast (b), east (c), southeast (d), south (e), west (f).**

northwest basins. However, we identify statistically significant trends in SGH activity over basins in the north (May: Basin 1.3-increasing area; Basin 1.4-increasing volume), northeast (September 2.2-volume), east (September: 3.2-area; 3.2, 3.4-volume), southeast (July, September: 4.3-area; June, July, September: 4.3-volume), south (July, September: 5.0-area; June, July: 5.0-volume), and southwest (July, September: 6.1-volume) throughout the decade (Figures B.1a, B.1b).

In regions where supraglacial water accumulates in lakes, there exists the

potential for this water to reach the glacier bed. Spatially, these areas delineate regions of the ice sheet that exhibit susceptibility to dynamic perturbations. Consequently, our interest extends beyond mere trends in meltwater extent and volume; we are equally concerned with variations in areas susceptible to such environmental forces. To assess this, we initially categorise ice sheet areas proximal to a lake within a given month. Subsequently, we analyse trends in this particular metric. To facilitate this process, we resample our SGHF extent dataset into a coarser scale product (1200 m x 1200 m). In this resampled dataset, cells receive a value of 1 if they encompass any portion of an SGHF and 0 if they do not. This better represents the area of the ice sheet that might be affected by SGHF features than a simple sum of their area. Examining these data, we find that many regions of the ice sheet experience statistically significant positive trends in SGHF affected area, especially the south and southeast, which have seen a significant increase in SGHF affected area in all months of the year. Additionally, we observe an increase in the area affected by SGHF in the north of the ice sheet in May, the west of the ice sheet in August, and the east of the ice sheet in September (Figure 3.8a). The southwest and northeast of the ice sheet both host abundant populations of SGHF; however, we find no statistically significant increase in SGHF affected area here. We explore this further by calculating the ratio of standard deviation of the residuals on the SGFH affected area time series with its trend, i.e. to provide a measure of how variability in the time series affects the statistical significance of its trend (Figure 3.8b, B.2a, B.2b). We observe that for basins with no statistically significant trend in the area covered by the SGFH, the ratio is very high, suggesting that seasonal and interannual variability is large enough to conceal long-term trends. For these basins with high temporal variability, a longer observational record is needed to confirm the significance of the trends.



(a) **Trend in the SGHF occurrence.** *(b)* **Ratio for SGHF occurrence.**

Figure 3.8: **Trends for: (a) SGHF occurrence in the Zwally storage basins over the decade from 2014-2022, calculated for the products available at resampled 1200 m resolution and with the sensor correction applied to Landsat-8 era data (2014-2016) and (b) the ratio of the standard deviation of the residuals of the trend to the magnitude of the trend for the SGHF occurrence in the Zwally storage basins. In most cases, where the basins (in a) show no significant trend, this ratio is many times that of the basins with a significant trend. Labelled basins are those where we find statistically significant trends for the SGHF occurrence. In September, very little to no meltwater is present in the northern regions, resulting in a zero value for the standard deviation and a white colour in the plot.**

3.3.1.1 Drivers of increased supraglacial hydrology

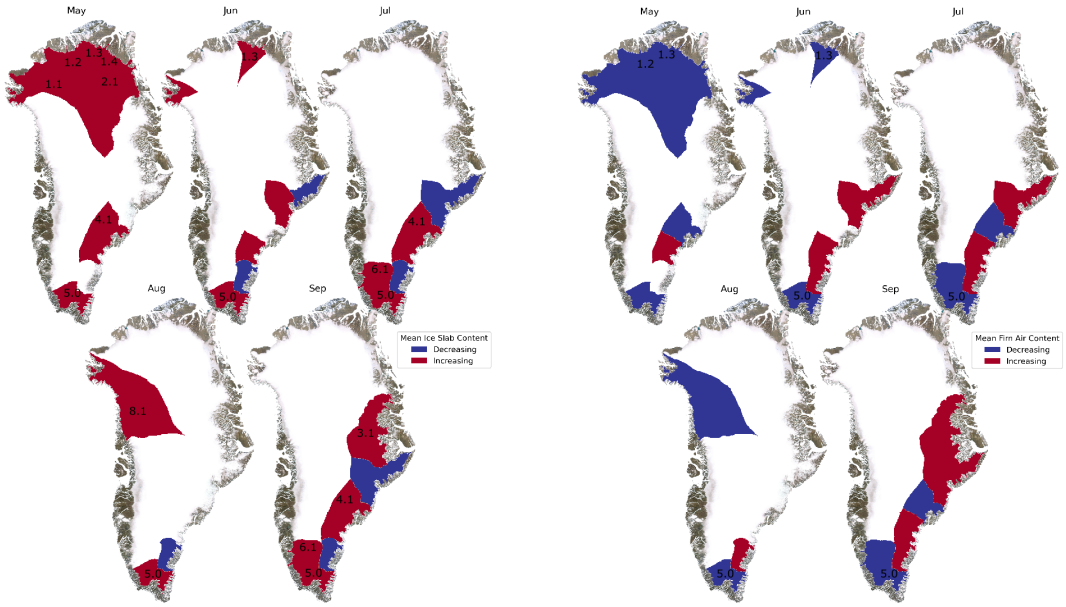
To investigate the drivers of the increase in SGHF we observe on the GrIS, we examined climate and ice surface property data from the IMAU-FDM v1.2G (Institute for Marine and Atmospheric Research Utrecht-firn densification model,

Brils et al., 2022) and RACMO2.3p2 (regional climate model, B. Noël et al., 2018). We calculate the average firn air content, average ice slab content, average firn temperature at 10 m depth, and average irreducible water content for the years 2014-2020 for each basin (e.g. Figure 3.2) from IMAU-FDM v1.2G. We also calculate the average runoff, average precipitation, average snowmelt, average snowfall, and average air temperature 2 m above the ice surface for the years 2014-2022 for each Zwally basin from RACMO2.3p2.

Firn air content, the vertically integrated volume of air contained within the firn, in the upper layers of snow on the ice sheet, is essential to control the storage and release of meltwater. Compaction of the firn reduces the permeability of the snowpack, making it more difficult for the meltwater to percolate through the firn (Pfeffer, Meier, and Illangasekare, 1991; Vandecrux et al., 2019). This leads to more water remaining on the surface, resulting in an increase in the formation of supraglacial lakes and channels (Pfeffer, Meier, and Illangasekare, 1991; Vandecrux et al., 2019). Changes in the ice slab content can affect the drainage of stored meltwater within the ice sheet. Ice slabs are dense layers or regions of ice that are distinct from the surrounding ice. These slabs are typically characterised by their relatively low permeability, high density, and reduced air content compared to the firn that forms the upper layer of the ice sheet. If the ice slab becomes more impermeable due to changes in its physical properties, it can act as a barrier, preventing downward percolation of the meltwater (Culberg, Schroeder, and Chu, 2021; MacFerrin et al., 2019). The irreducible water content is the fraction of the pore space within the firn that remains filled with water even under high pressure, and changes in this parameter can influence the way water is stored and released, influencing supraglacial hydrology (Brils et al., 2022; Covi, Hock, and Reijmer, 2023). Runoff, which is the volume of water that flows across the ice surface, is closely related to changes in meltwater production, snowmelt, and rainfall. An

increase in runoff can be caused by higher temperatures and increased melt, resulting in more extensive supraglacial water networks (Gantayat et al., 2023; Y. Li et al., 2022; K. Yang et al., 2019b). Precipitation, snowmelt, and snowfall are all directly related to the input of water into the ice sheet, and variations in these factors can increase or decrease the accumulation and ablation of the surface of the ice sheet (McIlhattan et al., 2020; B. Noël et al., 2015). For example, a decrease in snowfall can lead to reduced albedo and increased surface melt, resulting in a more extensive supraglacial hydrology (McIlhattan et al., 2020). Lastly, changes in temperature can affect the overall energy balance of the ice sheet. Higher air temperatures lead to increased surface melting, which contributes to the formation of supraglacial water bodies (Hanna et al., 2021). Subsurface firn temperature variations can also affect the rate of refreezing, which in turn affects the availability of liquid water on the surface (Brils et al., 2022; Culberg, Schroeder, and Chu, 2021).

We observe significant seasonal trends in the mean firn air content and mean ice slab content in the Zwally drainage basins in the north, south, and east, which experience a significant increase in supraglacial hydrology (Figures 3.9a, 3.9b). This suggests that these ice slabs render the near-surface of the ice sheet impermeable, inhibiting the downward percolation of the meltwater, and thus increasing the ponding at the surface (Culberg, Schroeder, and Chu, 2021; MacFerrin et al., 2019; Pfeffer, Meier, and Illangasekare, 1991; Vandecrux et al., 2019). Although the IMAU-FDM v1.2G model and our estimates of the occurrence of SGHF do not cover the same time period, this analysis implies that a decrease in the firn air content and an increase in the ice slab content could be a potential cause of the increase in the occurrence of SGH (Figures B.1a, B.1b, 3.8a). We examine trends for other variables, such as firn temperature at 10 m depth, irreducible water content, runoff, precipitation, snowmelt, snowfall, and air temperature 2 m above the ice surface, but no significant consistent trends were observed in the basins with



(a) **Significant trends in the mean ice slab content.** *(b)* **Significant trends in the mean firn air content.**

Figure 3.9: **Trends for the mean ice slab content and mean firn air content for the years 2014-2020 calculated from IMAU-FDM (Firn Densification Model) v1.2G (Brils et al., 2022), where we find a significant trend for the occurrence of SGHF (Figure 3.8a) across the years 2014-2022 for each Zwally basin. Labelled basins are those where we find statistically significant trends in the IMAU-FDM data.**

increasing SGH (Figures B.3a-B.3b, B.4a-B.4e). However, similar to the SGHF occurrence trends, variations from the low and high melt seasons may reduce the significance of the trends analysed. Longer-term studies are needed to assess in detail the drivers of increasing SGH.

3.3.1.2 Meltwater trends according to elevation

To gain a better understanding of the areas that contribute to an increase in supraglacial hydrology, we investigated the elevation at which SGH occurs. We identified a considerable increase in SGH occurrence, typically at elevations below 1500 m, for example, in the south and southeast in June, south and east in July, northwest in August, and south, west, and southwest in September. The decreasing SGHF trends for the elevation bands in the east and south regions during July and August explain some of the mismatching trends in the ice slab content analysis. However, supraglacial hydrology has been significantly increasing in some regions with elevations above 2000 m, such as the northeast in May and the south and west in September. This analysis implies that supraglacial hydrology is moving further inland, to higher elevations in some parts of the ice sheet (Figure 3.10).

3.4 Conclusion

In this study, we present a novel analysis of supraglacial water extent and volume that spans a decade throughout the Greenland Ice Sheet. We developed and rigorously validated a random forest algorithm designed to classify SGHF using the optical sensors Sentinel-2 and Landsat-8. This RF-based approach allowed us to delineate hydrological features at an unprecedented resolution and scale compared to conventional methods. Our findings reveal robust seasonal patterns in SGHF behaviour, characterised by the presence of meltwater in May, peaking during July/August, and subsequent refreezing or drainage in September each year.

In addition, we observed variability in both the extent and volume of SGHF on the Greenland Ice Sheet year by year. In particular, between 2014 and 2022, there was a statistically significant increase in the extent, volume, and occurrence of SGHF in numerous basins, particularly those located in the northern, eastern, and

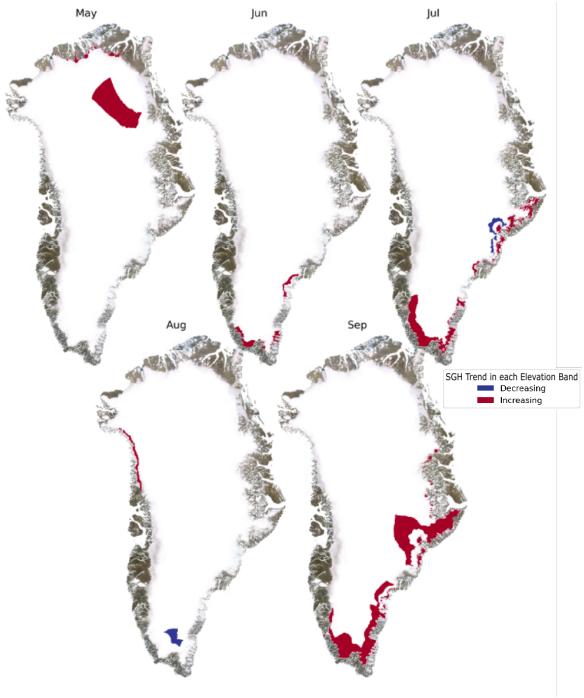


Figure 3.10: **Significantly changing trends for the occurrence of SGHF at the elevation bands (0-500m, 500-1000m, 1000-1500m, 1500-2000m, 2000-2500m, >2500m) according to ArcticDEM (Porter et al., 2023), where a significant trend was found for the occurrence of SGHF (Figure 3.8a) over the years 2014-2022 for each Zwally basin.**

southern regions of the ice sheet. This notable increase is predominantly attributed to increased SGHF activity at elevations below 1500 metres, although some higher-altitude regions also exhibited a marked increase.

The study also revealed a significant decrease in mean firn air content and a noteworthy increase in mean ice slab content in the basins that experienced a substantial increase in the SGHF. These findings suggest that alterations in these specific variables may drive the observed increase in SGHF. Although the duration of our current time series has been sufficient to detect significant trends across

multiple basins, the complexity of areas characterised by pronounced seasonal and interannual variability underscores the necessity for sustained observations to establish an extended time series effectively.

3.5 Data & code availability

The mapped supraglacial hydrology polygons and rasters are available on Zenodo (`https://doi.org/10.5281/zenodo.7573885`, D. Corr, 2023c) as digital GIS shapefiles (.shp) and GeoTIFF (.tif). Data consist of monthly and yearly extents and depths of supraglacial hydrology in Greenland for 2014 through 2022. L8 and S2 imagery are freely available at `earthexplorer.usgs.gov` and `scihub.copernicus.eu` respectively. Scripts to download data were extracted from GitHub (Hagolle, Olivier, 2014) and (Hagolle, Olivier, 2015), however, with changes to the data structure on both repositories, these scripts may no longer be effective. Alternatively, imagery is available to download from Google Cloud Storage using Python scripting (Nunes, Vasco, 2016).

The code used to produce the supraglacial hydrology features is written in Python and can be accessed on Zenodo and GitHub (D. Corr, 2023a). The authors express their gratitude for the use of the High End Computing (HEC) Cluster at Lancaster University in conducting this study.

Chapter 4

A new method for probabilistic prediction of supraglacial lakes on the southwest Greenland Ice Sheet

Diarmuid Corr[1], Israel Martínez-Hernández[2], Amber Leeson[1,3], Malcolm McMillan[1,4].

[1]Centre of Excellence in Environmental Data Science, Lancaster Environment Centre, Lancaster University, Lancaster.

[2]Department of Mathematics and Statistics, Lancaster University, Lancaster.

[3]Data Science Institute, Lancaster University, Lancaster.

[4]UK Centre for Polar Observation and Modelling, Lancaster University, Lancaster.

Correspondence:

Diarmuid Corr (dcorr103@gmail.com, d.corr@lancaster.ac.uk)

The following work is prepared as a manuscript to be submitted as: "A new method for probabilistic prediction of supraglacial lakes on the southwest Greenland Ice Sheet".

This chapter presents the research conducted to use optical satellite imagery to predict the probability of supraglacial meltwater in a given lake image and the uncertainty associated with them at the pixel level.

DC developed the code, carried out the main body of work, and drafted this paper. AL, MM, and IMH provided supervision and contributed extensively to the science, technical details, and structure of this paper. All authors contributed to the manuscript text.

Abstract

Greenland's supraglacial lakes are pools of liquid meltwater that form in depressions on the ice sheet during the summer months. These lakes are a sign of increased melting, which affects the storage and movement of meltwater on the ice sheet and can have an impact on the dynamics of the ice sheet, leading to an increase in sea levels when they drain. In this study, we employ a Bayesian statistical framework known as Integrated Nested Laplace Approximation with Stochastic Partial Differential Equation (INLA-SPDE) to predict the probability of supraglacial water presence at the pixel level within satellite imagery. Our approach uses multinomial logistic regression to discern one of three lake border conditions (well defined, blurred, and without a clear border) within these images. By leveraging a Matérn covariance function constructed on an SPDE mesh, we capture the spatial relationship between pixels depending on the border condition. The model is able to recognise the connections between random spatial effects

related to the border conditions, the predictor variables (Sentinel-2-based indices and individual reflectance bands), and the manually generated water/non-water labels. This enables us to compute probabilistic estimates for each pixel in a lake image and to measure the uncertainty associated with it pixel-by-pixel.

We explore the effect of thresholds on our probabilistic estimates to create a binary classification of surface water and compare the effectiveness of this classification to that of traditional static thresholding, convolutional neural network (CNN), and random forest (RF) algorithms. Our INLA-SPDE approach 1) significantly exceeds the performance of the static thresholding approaches, 2) produces results comparable to those of the CNN and RF algorithms, which were trained and tested on the same input data and unseen test data, and 3) provides the added value of a probabilistic estimate of the presence of water with an associated uncertainty. Our discoveries demonstrate the effectiveness of this Bayesian framework, offering a promising path for reliable prediction and classification of SGLs from optical satellite imagery. This approach is especially beneficial in cases where it is hard to distinguish the borders of a lake. It provides a more precise and adjustable prediction than a strict binary classification. This makes it a useful tool for analysing supraglacial lakes on a large scale and over a period of time, as these lakes can vary significantly in both space and time.

4.1 Introduction

In the last decade, observations have revealed rapid and significant changes in the Greenland Ice Sheet (GrIS) in response to increases in temperatures linked to global warming (Otosaka et al., 2022; Shepherd et al., 2018, 2020; B. Smith et al., 2020; Trusel et al., 2018). Models have predicted that these effects will continue to be prominent for several decades, even if global temperatures stabilise (Box et al., 2022; Goelzer et al., 2020; Lenaerts et al., 2019). The GrIS has suffered substantial ice mass loss in previous decades (Madsen et al., 2022; Moon et al., 2020; Slater et al., 2021), 60% of which is credited to the difference between snow accumulation and surface melt, the rest to dynamic mass loss (M. R. van den Broeke et al., 2016).

When surface meltwater accumulates in depressions on the surface of the ice sheet, supraglacial lakes (SGL), which vary in size from tens of metres to tens of kilometres (Kingslake et al., 2017; Selmes, Murray, and James, 2011; Stokes et al., 2019) and depths of centimetres to 12 m (J. F. Arthur et al., 2020a; Box et al., 2012), are formed (Echelmeyer, Clarke, and Harrison, 1991). SGLs play a significant role in ice sheet dynamics and mass balance. They increase surface albedo, which can lead to increased melting because white surfaces (such as snow and ice) reflect higher fractions of incoming radiation than darker surfaces (Bell et al., 2018; Leidman et al., 2021; Lenaerts et al., 2017). Meltwater that enters the englacial system, through cracks or crevasses on the surface, can transfer latent heat to the surrounding environment when it freezes during winter in a process called cryohydrologic warming (T. Phillips, Rajaram, and Steffen, 2010; T. Phillips et al., 2013). If enough meltwater enters cracks or crevasses, fracture growth can increase (Lai et al., 2020; T. Scambos et al., 2009) and open a connection between the ice surface and the bed, in a process known as hydrofracture. SGLs have been observed to drain rapidly through 1000 m thick ice through hydrofracture (Banwell,

MacAyeal, and Sergienko, 2013; Das et al., 2008; Doyle et al., 2013; Stevens et al., 2015; Tedesco et al., 2013); or slowly through channels and overtopping their banks (Hoffman et al., 2011; Tedesco et al., 2013). Rapid delivery of meltwater to the basal environment has been found to reduce basal friction and temporarily increase ice flow velocities by up to an order of magnitude (Tedesco et al., 2013) and impacts frontal melting of marine terminating glaciers (Tuckett et al., 2019). Monitoring the behaviour and volume of SGLs helps quantify their impact on total ice sheet mass loss and, consequently, on their contribution to global rise in sea level (Enderlin et al., 2014; M. van den Broeke et al., 2009).

SGLs around the GrIS margin have been extensively studied. Maps of SGLs are available interannually, covering the entire ice sheet, for the periods 2014-2022 (D. e. a. Corr, n.d., In Prep) and 2016-2018 (Hu et al., 2022). Additionally, smaller-scale studies have been conducted to investigate lakes in different glaciological conditions (A. A. Leeson et al., 2012, 2015; A. A. Leeson et al., 2013; Y.-L. Liang et al., 2012; McMillan et al., 2007; Sundal et al., 2009; Williamson et al., 2018). These studies have focused mainly on understanding the seasonal evolution of SGLs and applying algorithms to study sites associated with specific glaciological basins. Other research has revealed trends in the seasonal evolution and drainage behaviour of lakes (Selmes, Murray, and James, 2011) and in terms of the interannual and long-term evolution of their spatial coverage (Howat et al., 2013; Igneczi et al., 2016). The traditional approach to mapping SGL features combines spectral thresholding (usually the Normalised Difference Water Index, NDWI) with extensive manual post-processing (D. Corr et al., 2022; Fitzpatrick et al., 2014; M. Moussavi et al., 2020; Stokes et al., 2019; Williamson et al., 2018; K. Yang and L. C. Smith, 2013). Manual intervention is necessary to remove shadows, clouds, rocks, crevasses, and blue ice that are spectrally similar to meltwater and are sometimes misclassified as water (D. Corr et al., 2022). However,

with the large amount of satellite data now available, methods that require manual input are not suitable for repeated monitoring of surface hydrology on a continental scale. Given that there are tens of thousands of these features (D. Corr et al., 2022; Stokes et al., 2019; K. Yang et al., 2021) and appear in many thousands of satellite images, automated mapping techniques are often used. Machine learning (ML) algorithms such as random forests (RF, D. e. a. Corr, n.d.; M. Dirscherl et al., 2020; Hu et al., 2022, In Prep), convolutional neural networks (CNN, M. Dirscherl et al., 2021; Yuan et al., 2020), and k-means clustering (Halberstadt et al., 2020) offer potential solutions to this problem and are now widely used.

These techniques offer a way of classifying surface features; however, they provide only a binary classification with a limited estimate of uncertainty. A few studies have attempted to measure the level of uncertainty by assigning a percentage to the total area of supraglacial hydrology throughout the ice sheet (e.g., Stokes et al., 2019). This is based on the assumption that the total area of lakes identified from the models is likely to be similar to manual digitisation on the ice sheet scale. Other research has applied uncertainty values on a lake-by-lake basis, for example, an uncertainty of half-pixel width (125 m) to the circumference of each lake (Y.-L. Liang et al., 2012). No research has yet evaluated uncertainty at the pixel level. This is especially relevant for lakes with indistinct borders, where the most uncertainty is found in the prediction. As a result of their distinct blue hue compared to the white or grey ice surrounding them, many supraglacial lakes are easily identifiable in optical satellite images. Their distinct borders make them suitable for tracking changes in size, depth, and distribution over time using conventional methods, which can provide information on glacier melting patterns and hydrological processes, for example, hydrofracture-driven drainage (D. Corr et al., 2022; Fitzpatrick et al., 2014; M. Moussavi et al., 2020; Stokes et al., 2019; Williamson et al., 2018; K. Yang and L. C. Smith, 2013). However, others with

poorly defined or unclear borders are more difficult to identify. Supraglacial lakes with blurred borders have less defined edges compared to easily discernible lakes. The borders of these lakes may be partially obscured by the surrounding ice, snow, or slush (Dell et al., 2022; Machguth, Tedstone, and Mattea, 2023; Sneed and Hamilton, 2007). This can make it difficult to accurately delineate the extent of the lake from static thresholding and traditional methods (D. Corr et al., 2022; Fitzpatrick et al., 2014; M. Moussavi et al., 2020; Stokes et al., 2019; Williamson et al., 2018; K. Yang and L. C. Smith, 2013). How well defined the edges or borders of SGLs are depends on temperature, precipitation, elevation, undulations in the surface of the ice sheet. SGLs without a clear border are characterised by a lack of discernible edges that separate lake water from the glacier's surface. These lakes may appear as areas of standing water or slush that blend into ice or snow.

To address this issue, we use a Bayesian statistical approach (Integrated Nested Laplace Approximation with Stochastic Partial Differential Equation, INLA-SPDE) to predict the probability that each pixel in a given image is supraglacial water (Lindgren, Rue, and Lindström, 2011; Martins et al., 2013; Rue, Martino, and Chopin, 2009; Simpson, Lindgren, and Rue, 2012a). We combine the benefits of INLA's efficient Bayesian inference with SPDE's capacity to model spatial dependencies. We can use spatial random effects to acknowledge spatial relationships, such as lake pixels are more likely to be near other lake pixels, and lakes with clear borders will have more distinction at their edges. This will help us identify water features in satellite imagery, while taking into account the spatial correlations present in SGL features (C. Yang et al., 2023). Our study provides probabilistic predictions of surface water at the pixel level and the associated standard deviations, which represent the uncertainty in the prediction. A threshold is then used to classify the pixels into two categories according to the prediction. The threshold for this method is set to a probability of 0.50, which is the point at

which it is more likely that a pixel is water than not. This threshold can be adjusted within the model to ensure a certain level of certainty.

4.2 Data & methods

Here, we describe the data and workflow (Figure 4.3) to predict SGLs from Sentinel-2 images. In Section 2.1 we provide an overview of the glaciological and climatological context of the study area. Next, in Section 2.2 we outline the definition of three SGL border conditions (well defined, blurred, and without a clear border). In Section 2.3 the source bands and calculated band indices from Sentinel-2 data are outlined. Finally, in Section 2.4 we describe the model framework, including the logistic regression algorithm to determine the border conditions, the INLA-SPDE approach to predict the probability of SGL, with the associated mesh selection and mixed-effects model.

4.2.1 Study area & lake border conditions

The western part of Greenland is an important region for studying the interactions between glacial systems, hydrological networks, and changing climatic conditions, for example, the committed SLR from 2000 to 2019 was four times greater in here than in the east (Box et al., 2022). As the global climate continues to warm, the region has experienced increased supraglacial meltwater (Hu et al., 2022), altering the hydrological regime and posing challenges to local ecosystems and communities (Fitzpatrick et al., 2014; Pitcher and L. C. Smith, 2019; L. C. Smith et al., 2015; L. C. Smith et al., 2017; K. Yang and L. C. Smith, 2016; K. Yang et al., 2021). We categorise SGLs, by eye, into three types based on the clarity of their edge pixels: well defined, blurred, and without a clear border (Figure 4.2).

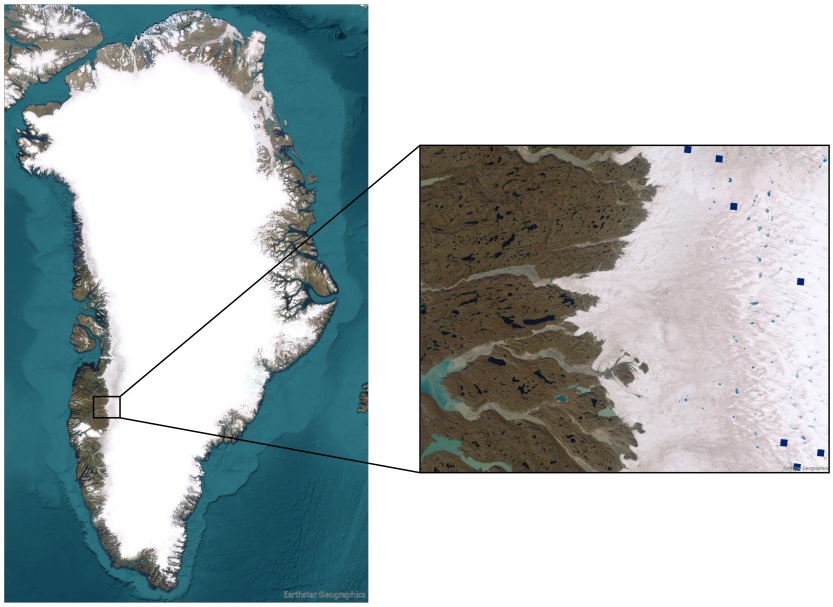


Figure 4.1: **Locations of the seven study sites on the southwest Greenland Ice Sheet. Basemap source: Esri, Maxar, Earthstar Geographics, and the GIS User Community.**

4.2.2 Training & testing data

We compile our data set using Sentinel-2 satellite imagery acquired during multiple melt seasons (2017-2022), throughout the melt season (June, July, and August), at seven sites that host supraglacial lakes in southwest Greenland (66°46'13.3917"-67°32'41.8113" N, 48°32'9.0654"-48°40'43.4618" W; Figure 4.1), and divide into training and testing sets. The training set consists of SGL images that vary in terms of melt season, month, research site, and lake border conditions (Figure 4.2, Table C.1). The test set is not seen by the algorithm, and the lake border conditions and binary water/non-water pixels are not specified. Our methodology uses INLA-SPDE on Sentinel-2 optical satellite imagery, allowing the incorporation of spatial random effects to capture spatial dependencies and identify features with differing borders within the imagery.

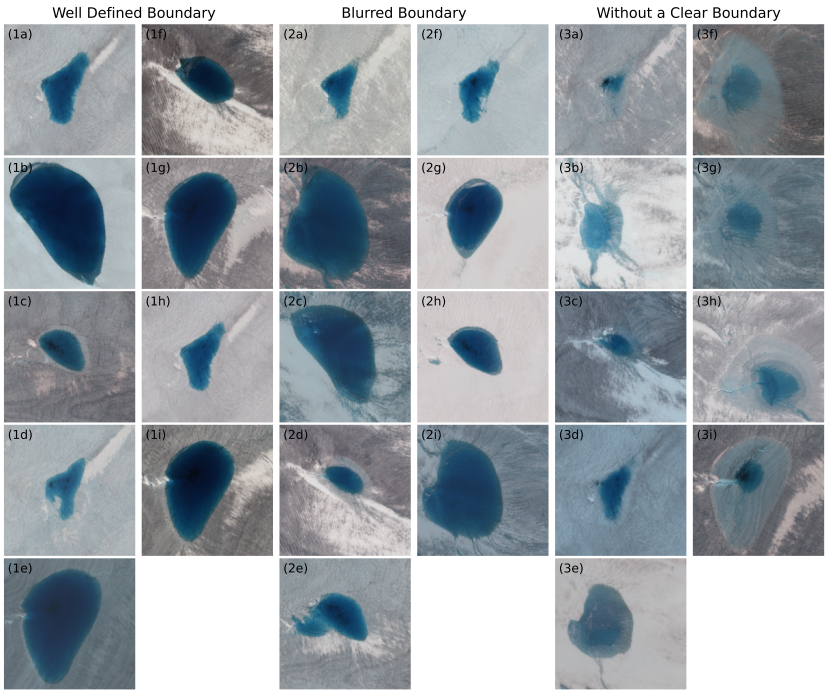


Figure 4.2: **Training data acquired during multiple melt seasons (2017-2022), throughout the melt season (June, July, and August), at seven sites that host supraglacial lakes in southwest Greenland for the three types of border conditions (well defined (1a-i), blurred (2a-i), and without a clear border (3a-i)). The geographic locations and date of acquisition are provided in Table C.1. The true-colour Sentinel-2 composites depict the lakes that vary in terms of melt season, month, research site, and lake border condition.**

The labels for the training data are determined by manually digitising the SGLs in true-colour composites of Sentinel-2 imagery. These labels, which are represented as a binary raster of 1s (water) and 0s (non-water), are combined with a combination of individual spectral bands and indices from Sentinel-2 imagery. Although manual digitisation is a reliable way to map lake areas, there is still uncertainty associated with the approach (A. A. Leeson et al., 2013).

4.2.2.1 Sentinel-2 bands & indices

The Sentinel-2 bands and the indices derived from them constitute the input data used to train and validate the algorithm. Sentinel-2 bands are freely available as top of atmosphere (TOA) reflectance data. These data can be accessed from the Copernicus Open Access Hub (`https://scihub.copernicus.eu/`) and Google Cloud Storage (Google Cloud Storage). The images have a resolution of 10 m, the highest spatial resolution acquired by the sensor, and contain bands 2 (blue), 3 (green), 4 (red), and 8 (near infrared - NIR). Bands 5, 6, 7, 8A (vegetation red edge 1-4), 11 (short-wave infrared, ($SWIR_1$)) and 12 ($SWIR_2$) are acquired at a coarser resolution of 20 m, while bands 1 (coastal aerosol), 9 (water vapour) and 10 ($SWIR_{cirrus}$) exist at 60 m. All bands with coarser resolution than 10 m are resampled to 10 m using nearest-neighbour interpolation for consistency with the red, green, and blue (RGB) and NIR bands (D. Corr et al., 2022; Williamson et al., 2018). The pixel values of the S2 images represent TOA reflectance units $\times 10000$ and are known as TOA reflectance integers (reflectance$\times 10^4$).

Band indices derived from optical satellite imagery are numerical values or combinations of pixel values in various spectral bands that are used to obtain particular information (terrestrial features) regarding the surface of the ice sheet. Band indices are combined with individual spectral bands to create composite multiband tiffs (25 bands). The indices are $NDWI_{BlueRed}$ (a), $NDWI_{GreenNIR}$ (b), New Water Index (NWI, c), Normalised Difference Snow Index (NDSI, d - also known as Modified NDWI), Soil/Water Index (SWI, e), Normalised Difference Glacier Index (NDGI, f), modified Soil Adjusted Vegetation Index ($SAVI_{mod}$, g), modified Shadow Index (SI_{mod}, h), Tasselled Cap for wetness (TC_{wet}, i), Automated Water Extraction Index with shadow ($AWEI_{sh}$, j) or without shaded area removal ($AWEI_{nsh}$, k), and Normalised Difference Index (NDI, l). These indices are derived from the S2 TOA

bands (Table 4.1).

Table 4.1: **Indices calculated from the Sentinel-2 sensor, used in the random forest algorithm.**

Index Name		Equation	Terrestrial Feature(s)	Sensor(s)	Reference
$NDWI_{BlueRed}$ (a)	=	$\frac{Blue - Red}{Blue + Red}$	Surface Water	S2 & L8	M. Moussavi et al., 2020
$NDWI_{GreenNIR}$ (b)	=	$\frac{Green - NIR}{Green + NIR}$	Surface Water	S2 & L8	D. Corr et al., 2022
NWI (c)	=	$\frac{Blue - (NIR+SWIR_1+SWIR_2)}{Blue + (NIR+SWIR_1+SWIR_2)}$	Surface Water, Shadow	S2 & L8	Feng, 2009
NDSI (d)	=	$\frac{Green - SWIR_1}{Green + SWIR_1}$	Rock, Sediment, Ice	S2 & L8	Hall, Riggs, and Salomonson, 1995; Xu, 2006
SWI (e)	=	$\frac{Blue - SWIR_1}{Blue + SWIR_1}$	Rock, Sediment, Ice	S2 & L8	M. Dirscherl et al., 2020
NDGI (f)	=	$\frac{Green - Red}{Green + Red}$	Rock, Sediment, Ice	S2 & L8	Keshri, Shukla, and Gupta, 2009
$SAVI_{mod}$ (g)	=	$2 \times \frac{Green - NIR}{1 + Green + NIR}$	Rock, Sediment, Ice, Shadow	S2 & L8	Huete, 1988
SI_{mod} (h)	=	$\frac{Blue - NIR}{Blue + NIR}$	Shaded Ice/Snow, Shadow	S2 & L8	H. Li et al., 2016
TC_{wet} (i)	=	$(0.1509 \times Blue) + (0.1973 \times Green)$ $+ (0.3279 \times Red) + (0.3406 \times NIR)$ $-(0.7112 \times SWIR_1) - (0.4572 \times SWIR_2)$	Surface Water	S2 & L8	Kauth and G. Thomas, 1976 Schwatke, Scherer, and Dettmering, 2019
$AWEI_{sh}$ (j)	=	$Blue + (2.5 \times Green)$ $-(1.5 \times (NIR + SWIR_1)) - (2.5 \times SWIR_2)$	Surface Water, Shadow	S2 & L8	Feyisa et al., 2014
$AWEI_{nsh}$ (k)	=	$4 \times (Green - SWIR_1)$ $- (0.25 \times NIR) - (2.75 \times SWIR_2)$	Surface Water	S2 & L8	Feyisa et al., 2014
NDI (l)	=	$\frac{Green - Blue}{Green + Blue}$	All Features	S2 & L8	M. Dirscherl et al., 2020

We use a standard scaler to modify our training data so that each band/index has a mean of 0 and a standard deviation of 1. Standard scaling can improve the model's convergence and overall performance, and is therefore a crucial step in building data pipelines and recommended in the application of Machine Learning.

4.2.3 Model framework

Application of the INLA-SPDE model comprises three key stages: (1) data preparation, (2) training and validation, and (3) prediction and classification (Figure

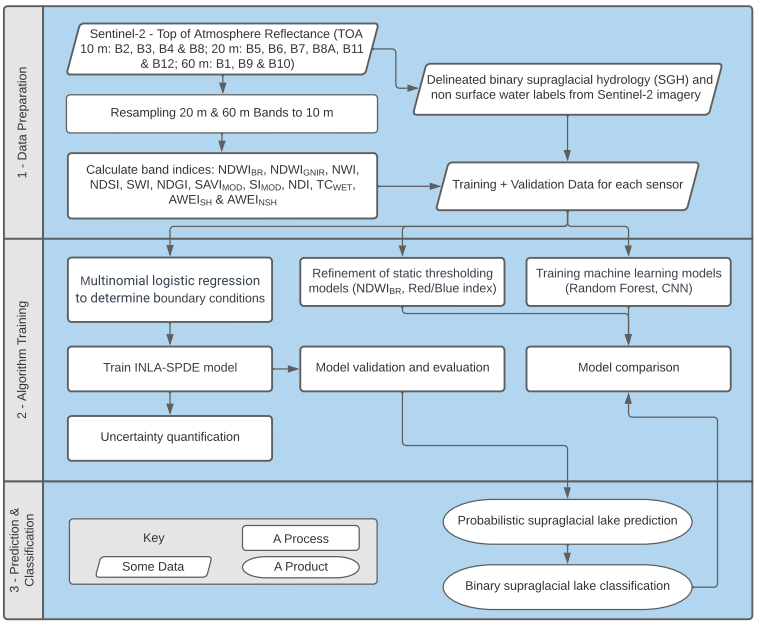


Figure 4.3: **Data pipeline for the preparation, training, evaluation, validation, comparison, and prediction of SGLs in the workflow.**

4.3). In this section, we describe our method for (2) training and evaluation. First, a multinomial logistic regression algorithm is used to determine border conditions. The INLA-SPDE model is trained for each border condition, using the same mesh for all three. This means that any image can be evaluated without needing to know the lake border condition. By including the mesh in the analysis, the model can take into account the intricate spatial characteristics between pixels. The mesh allows for modelling of spatial dependencies and interactions, which can describe the proximity, patterns, and presence of meltwater between pixels. A Matérn covariance model is constructed on the SPDE mesh to capture spatial dependencies by capturing the influence of nearby pixels on each other. Our training data consist of the Sentinel-2 bands and indices described in Section 2.3. Each multi-band input image is 192×192 pixels and depicts an SGL with one of the three border

conditions.

4.2.3.1 Determination of border conditions

To discern the appropriate border condition in the training of the INLA-SPDE model, a multinomial logistic regression (Böhning, 1992; Ramadhan, Astri Novianty, and Casi Setianingsih, 2017) model is trained on the extracted features from each band in the GeoTIFF composite (Equation 4.1). The extracted features include calculated Shannon entropy, mean reflectance value, standard deviation of reflectance values (SD), fractal dimensionality, skewness, kurtosis, smoothness, and contrast values (Ahmadzadeh et al., 2017; Banda and Angryk, 2010) on each index (Table 4.1) and Sentinel-2 reflectance band. The model takes the values for the features extracted from the satellite imagery as input (and learns the relationships between the extracted features and the border condition) and predicts the probability of each border condition for each image. The border condition with the highest probability is assigned to each image. The predicted border condition determines the selection of the INLA-SPDE model. Each category in the response variable (Border Condition) represents a possible condition (defined, blurred, or without a clear border). The linear predictor ($f(k,i)$, Equation 4.1) is used to predict the probability that an observation i (that is, a given image) has the result k (that is, is a given border condition). In multinomial logistic regression, a softmax function ($softmax(x)$) is used to convert the linear combination into probabilities.

$$f(k,i) = \beta_{0k} + \sum_{i=1}^{8} \beta_{ik} x_i \tag{4.1}$$

$$softmax(f(k)) = \frac{e^{f(k)}}{\sum_{i=1}^{3} e_i^{f(k)}} \tag{4.2}$$

Where: $f(k,i)$ represents the linear predictor for each border condition, $k = 1, 2, 3$

and the predictor variable $i = 1...8$. β_{0k} represents the intercept for each border condition. β_{ik} represents the estimated coefficients for each border condition and predictor variable. x_i represent the eight predictor variables, Shannon entropy, mean reflectance value, standard deviation of reflectance values, fractal dimensionality, skewness, kurtosis, smoothness, and contrast. $softmax(f(k))$ represents the softmax function for the linear predictor $f(k)$.

4.2.3.2 INLA-SPDE model

Bayesian inference is a technique for making informed predictions in the face of uncertainty. It uses prior knowledge of a problem to create a prior probability distribution, which is then combined with the observed data associated with the parameters to form a likelihood. This likelihood is a measure of how well the model explains the data. The posterior probability distribution, $P(\boldsymbol{\theta}|\boldsymbol{y})$, is then calculated using Bayes' theorem, which combines the prior probability distribution, $P(\boldsymbol{y}|\boldsymbol{\theta})$, with the likelihood, $P(\boldsymbol{\theta})$, and divides by a normalisation constant, $P(\boldsymbol{y})$, for the data $\boldsymbol{y}$ given parameters $\boldsymbol{\theta}$. This posterior distribution reflects the updated understanding of the parameters after considering the data and provides the basis for informed predictions.

$$P(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\boldsymbol{y})} \tag{4.3}$$

INLA is a useful tool for performing Bayesian inference focused on estimating the posterior marginals of the model parameters (Rue, Martino, and Chopin, 2009). Posterior marginals capture the uncertainty, measure the accuracy of the prediction, and allow the formation of credible intervals to quantify the likelihood of meltwater being present. We outline the steps necessary to reproduce our results here, but for full details of the methods, we refer the reader to (Gómez-Rubio, 2020; Krainski et al., 2019). Instead of estimating a difficult to acquire highly multivariate joint

posterior distribution, $P(\boldsymbol{\theta}|\boldsymbol{y})$ the model focuses on obtaining approximations for univariate posterior distributions, $P(\boldsymbol{\theta}_i|\boldsymbol{y})$. Although, multivariate posteriors have the advantage of being able to model complex relationships and correlations between multiple variables. Univariate posteriors have the benefit of being simpler to analyse and interpret, more computationally efficient, and less likely to overfit than multivariate posteriors. Therefore, INLA-SPDE provides a good balance between computational feasibility and informative inference, making it a useful tool for Bayesian analysis. The initial step is to create a spatial mesh that encompasses the lake region to capture spatial influences. Subsequently, a mixed-effects model is fitted, with parameters that vary for each lake border condition. This mixed-effects model combines spatial effects and fixed effects. Fixed effects capture the correlation between the data obtained from satellite imagery bands (predictor variables) and the likelihood of a lake's existence (response variable). Meanwhile, spatial effects take into account local discrepancies and spatial autocorrelation in the data, for instance, water pixels are more likely to be situated near other water pixels. All effects are determined by the input satellite data and are used to generate a linear predictor for each pixel that can be used to calculate a probabilistic prediction.

4.2.3.2.1 Mesh selection using stochastic partial differential equations

A spatial mesh is a grid of nodes that covers the area where observations are made. In this study, we use a consistent mesh on an image of consistent dimensions (192 $\times$ 192 pixels), which reduces much of the complexity of selecting a non-regular grid; therefore, the only variable in our mesh generation is the number of vertices in the grid (Figure 4.4).

INLA uses Delaunay triangulation, a geometric process which divides a set of

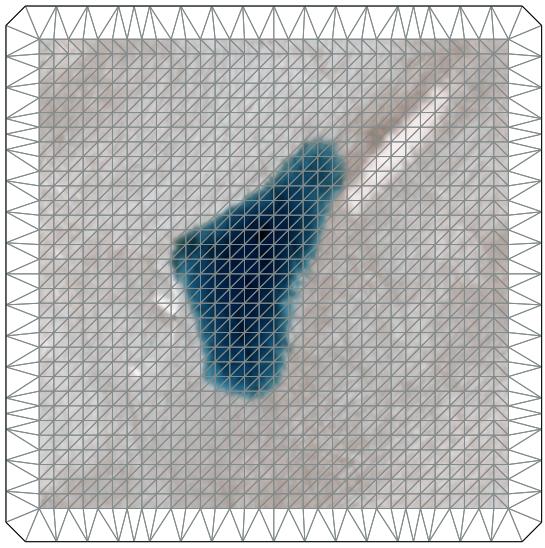


Figure 4.4: **The constructed mesh network for INLA-SPDE model overlaid on an example 192×192 pixel image depicting an SGL with a well defined border. The mesh is consistent for all models.**

points into non-overlapping triangles, to construct an SPDE model that explains how the values of a spatial field vary across space. This projection of the continuous space onto a discrete space (the mesh) is composed of small triangles that approximate the Matérn Gaussian field (Lindgren, Rue, and Lindström, 2011). The mesh is a necessary component for spatial modelling with INLA, as it is combined with spatial data to construct statistical models that capture spatial dependencies and make predictions or estimates within the given spatial domain. The mesh is designed to create the most regular small triangles possible while balancing computational cost and modelling accuracy. After creating the mesh, the spatial correlation structure of the SPDE is determined using the Matérn covariance function (Simpson, Lindgren, and Rue, 2012a,b). The spatial correlation structure can be used to determine how the presence or absence of lakes in one pixel is

associated with the presence or absence of lakes in adjacent pixels. The Matérn covariance function determines the correlation between the presence (or absence) of a lake in different pixels. The variance, smoothness, and range parameters of the Matérn covariance function are estimated from the observed data and can be used for spatial prediction. The SPDE model describes spatial random effects in a generalised linear mixed-effects binomial model (Lindgren, Rue, and Lindström, 2011; Martins et al., 2013; Rue, Martino, and Chopin, 2009; Simpson, Lindgren, and Rue, 2012a). Spatial random effects account for localised deviations from the fixed effects, which helps to discern pixels at the lake edge depending on the border condition.

4.2.3.2.2 Mixed-effects model

In hierarchical Bayesian models, such as INLA-SPDE, a combination of fixed, random, and spatial effects is used to account for various sources of variability in the data. In our model, random and spatial effects are incorporated into the SPDE model. Fixed effects represent the systematic part of the model and capture the overall trend or relationship between the response variable and the predictor variables. Here, the fixed effects are coefficients associated with the covariates, that is, the reflectance values of the 13 Sentinel-2 bands and the 12 band indices (Table 4.1).

The response variable (Equation 4.4), which takes a different form depending on the lake border condition defined by multinomial logistic regression (Equation 4.1), is modelled using a binomial family and the logit link function.

$$Labels \sim -1 + intercept + \sum_{j=1}^{25} X_j + f(SpatialEffects, model = spde) \tag{4.4}$$

Where: *Labels* represents the binary output corresponding to water/non-water. -1 means that no additional intercept is included in the model, as we have explicitly defined the *intercept*. *intercept* represents the explicitly defined intercept and is assigned a different value depending on the lake border condition. X_j represents the 25 predictor variables defined by the Sentinel-2 band indices and the source reflectance bands ($NDWI_{BlueRed}$, $NDWI_{GreenNIR}$, NWI, NDSI, SWI, NDGI, $SAVI_{mod}$, SI_{mod}, TC_{wet}, $AWEI_{sh}$, $AWEI_{nsh}$, NDI, Aerosol, Blue, Green, Red, $VNIR_1$, $VNIR_2$, $VNIR_3$, NIR, $VNIR_4$, Water Vapour, $SWIR_Cirrus$, $SWIR_1$, $SWIR_2$). $f(SpatialEffects, model = spde)$ represents the spatial random effects modelled using the SPDE approach and takes a different form depending on the lake border condition identified as defined in Section 2.4.1.

The linear predictor (η_i) is used to calculate the log-odds of the response variable for each model (Equation 4.5).

$$\eta_i = \beta_0 + \sum_{j=1}^{25} \beta_j X_{ij} + f(SpatialEffects_i, model = spde) \tag{4.5}$$

$$\rho_i = \frac{1}{1 + e^{-\eta_i}} \tag{4.6}$$

Where: η_i represents the linear predictor for each pixel, $i = 1...n$ and $n = 192 \times 192 = 36864$. β_0 is the estimated coefficient of *intercept*. β_j represents the estimated coefficients for the predictor variables, $j = 1...25$. X_{ij} represents the predictor variable for each Sentinel-2 band ($j = 1...25$) and pixel within the image ($i = 1...36864$). $f(SpatialEffects_i, model = spde)$ represents the spatial random effects for each pixel modelled using the SPDE approach, and is assigned different values depending on the lake border condition. ρ_i represents the probability that a given pixel (i) contains water.

The logarithmic odds are then transformed into probabilities using the logit link function (Equation 4.6) to model the probability (ρ) that a given pixel is water.

Therefore, the output of the model is a probabilistic prediction that each pixel in the (192×192) image contains supraglacial meltwater. A binary map of water/non-water pixels can be obtained by selecting an appropriate threshold.

4.2.4 Other classification approaches

We evaluate the precision of our algorithm by computing the binary cross-entropy (Good, 1952), which quantifies the dissimilarity between predicted probabilities and actual binary labels. However, comparing the performance with other models is difficult, as all models in the literature provide a binary classification of SGLs only, reporting F1 scores or overall accuracy. In order to perform this comparison, we apply a threshold to our probabilistic prediction to generate a binary lake/not-lake classification. F1 scores are calculated against the true labels for all images and compared for each approach: our INLA-SPDE approach, two static thresholds, CNN, and RF.

The static thresholding approaches, $NDWI_{GNIR}$ and $NDWI_{BR}$, which are the accepted method in the literature (D. Corr et al., 2022; Fitzpatrick et al., 2014; M. Moussavi et al., 2020; Stokes et al., 2019; Williamson et al., 2018; K. Yang and L. C. Smith, 2013), are optimised using the same training data as the INLA-SPDE model (Figure 4.2, Table 4.1) and an AUC-ROC curve approach (area under curve-receiver operating characteristics, Song, 2015) to determine the optimal threshold ($NDWI_{GNIR}$: 0.51; $NDWI_{BR}$: 0.04 for normalised data).

CNN and RF algorithms were trained on the same training data used in training the INLA-SPDE model and tested on the same unseen test data. Hyperparameter optimisation with cross-validation was carried out on the RF algorithm to ensure that the model achieved a good fit, with hyperparameters for the number of estimators/trees (200), maximum number of levels in each decision tree (10),

maximum number of features considered for splitting a node (square root of total features), minimum number of data points placed in a node before the node is split (5), and the minimum number of data points allowed in a leaf node (1) considered. The CNN architecture, based on an approach in the literature (Yuan et al., 2020), is made up of three convolutional layers, followed by a dense layer and an output layer for prediction. Max-pooling layers are used to reduce the feature maps after each convolutional layer. A rectified linear unit (ReLU) activation function is applied within the convolutional layers, with padding, to keep the image size constant. We have chosen a pool size of 3 and stride length of 1, and the He Normal kernel initializer, with the Adam optimiser, binary cross-entropy loss, and accuracy metric used in the model compilation. The final output layer produces binary classification predictions using a sigmoid activation function. We assess the performance and optimisation of our CNN by assessing loss and accuracy training curves. Our CNN is performing well as the loss is decreasing and the accuracy is increasing over the epochs, and the validation loss and accuracy are tracking the training curves, which shows that the learning is successful.

4.3 Results & discussion

To begin, we analyse the performance of probabilistic predictions using binary cross-entropy and providing uncertainty quantification. We then apply a threshold to the prediction to get a binary classification and compare the results with other classification techniques. Finally, we evaluate the impact of the chosen threshold on binary classification.

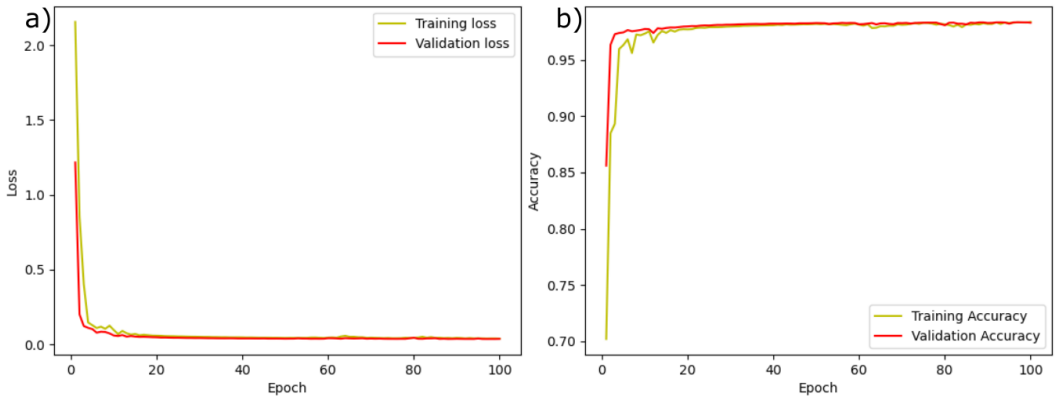


Figure 4.5: **Training and validation learning curves showing optimisation (loss, a) and performance (accuracy, b) of the CNN algorithm trained on our input data and validated using unseen test data.**

4.3.1 Validation & uncertainty quantification

Binary cross-entropy is a model metric that measures how well the predicted probabilities of a model align with the true binary labels, 0 or 1, and assigns a score that penalises the probabilities based on how close or far they are from the expected value (Table 4.2). We evaluate the precision of our algorithm without separating the border conditions by computing the binary cross-entropy, so that the model can be executed without the need to assign labels to the lake border conditions. A sample weight is used to adjust the binary cross-entropy calculation for each image to reduce the bias that can be caused by an unequal proportion of water/non-water pixels.

A binary cross-entropy value of 0 indicates perfect probabilities, and our calculated value is 0.06 for all values and the mean. This suggests that the model's predicted probabilities are quite close to the true binary labels, making relatively accurate predictions. The binary cross-entropy score used to evaluate the success of a classification problem varies, but it is generally agreed that scores below 0.20 are

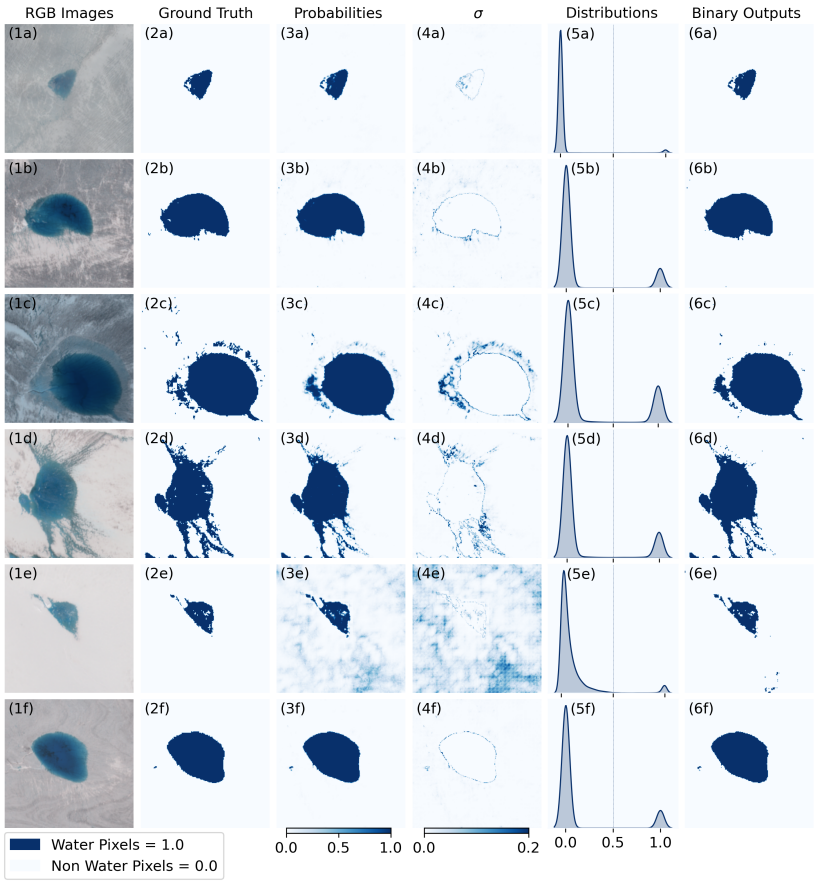


Figure 4.6: **I) Plots of the INLA-SPDE model outputs for first 6 unseen test images that contain SGLs in southwest Greenland, representing the three types of lake border condition. 1: RGB images; 2: True labels of binary water (dark blue) / non-water (pale blue); 3: Probabilistic predictions of our INLA-SPDE model between 0 (pale blue, non-water) and 1 (dark blue, water); 4: standard deviation (note different scale to predictions), σ of the probabilistic prediction (pale blue represent lower values, dark blue higher); 5: plotted distributions of the probabilistic prediction revealing the bimodal peaks associated with water and not-water; 6: binary outputs obtained from applying a threshold of $\rho_i \geq$0.5 on the probabilistic predictions.**

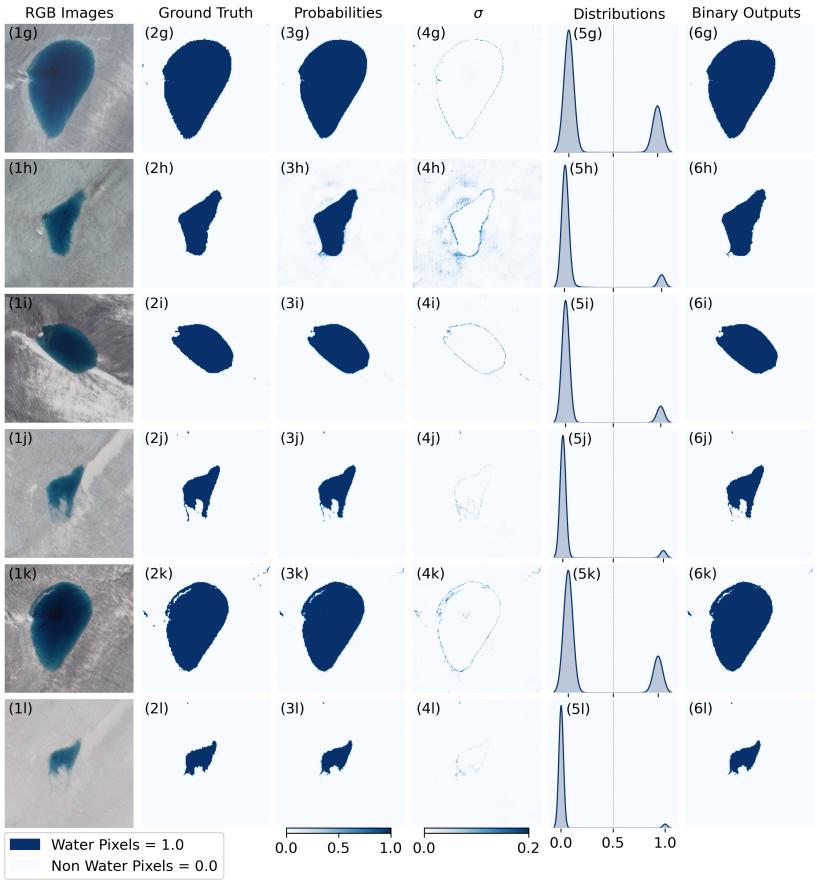


Figure 4.6: **II) Plots of the INLA-SPDE model outputs for the remaining 6 unseen test images that contain SGLs in southwest Greenland, representing the three types of lake border condition. 1: RGB images; 2: True labels of binary water (dark blue) / non-water (pale blue); 3: Probabilistic predictions of our INLA-SPDE model between 0 (pale blue, non-water) and 1 (dark blue, water); 4: standard deviation (note different scale to predictions), σ of the probabilistic prediction (pale blue represents lower values, dark blue higher); 5: plotted distributions of the probabilistic prediction revealing bimodal peaks; 6: binary outputs obtained from applying a threshold of $\rho_i \geq$0.5 on the probabilistic predictions.**

acceptable, scores below 0.05 are good, and scores below 0.02 are great (Brownlee, 2019). The performances for all images are deemed acceptable, while 7 of 12 perform well.

Although we evaluate the effectiveness of the multinomial logistic regression model through the binary cross-entropy metric for the final algorithm, we also provide a comprehensive assessment by computing this metric individually for each image with different border conditions applied. Our findings reveal that the predicted border conditions, i.e., the conditions predicted by multinomial regression, outperform other condition models in 7 of 12 images (Table 4.2), demonstrating their superior predictive capacity. The optimal border conditions agree with the border conditions we label in 8 of 12. Furthermore, even in cases where the model performance is not optimal, the binary cross-entropy values remain consistently close to the optimal score, differing by no more than 0.02 (Table 4.2). The results of this analysis indicate that our approach is the most accurate and consistent, although a single model could be used with only a slight decrease in performance. This highlights the strength and reliability of our approach when tested on different images.

This novel Bayesian model offers insight into the uncertainty of the predictions of supraglacial lakes at the pixel level, something that has not been done before. This is quantified by the standard deviation per pixel (σ, Figures 4.6 (4a-l)), which is narrower for more certain predictions and wider for more uncertain ones. In the majority of cases, the prediction is uncertain near the lake borders, although the uncertainty areas are smaller for those with a more distinct border (Figures 4.6 (4b, f, g, i, j, k, l)). For example, the average standard deviation across the images was 0.001 for those with a well-defined border, 0.005 for those with a blurred border, and 0.007 for lakes without a clear border, as indicated by multinomial regression. In other images, the borders are less well defined and the prediction is more

Table 4.2: **Validation of the INLA-SPDE approach with border conditions predicted by the multinomial regression (Predicted), border conditions we label (Labelled), binary cross-entropy (BCE) for well defined, blurred, and without clear borders, and the optimal border conditions calculated from BCE values (Optimal). The mean and standard deviation of each column are calculated for the heading in each column according to the border condition in each row.**

Value	Predicted	Labelled	Well Defined BCE	Blurred BCE	Without Clear BCE	Optimal
Image a	Without Clear	Without Clear	0.07	0.14	0.08	Well Defined
Image b	Blurred	Blurred	0.02	0.02	0.06	Blurred
Image c	Blurred	Blurred	0.16	0.13	0.12	Without Clear
Image d	Blurred	Without Clear	0.21	0.17	0.15	Without Clear
Image e	Without Clear	Without Clear	0.09	0.10	0.11	Well Defined
Image f	Well Defined	Well Defined	0.03	0.03	0.03	Well Defined
Image g	Well Defined	Well Defined	0.03	0.03	0.03	Well Defined
Image h	Blurred	Blurred	0.01	0.01	0.10	Blurred
Image i	Well Defined	Well Defined	0.01	0.02	0.04	Well Defined
Image j	Without Clear	Without Clear	0.03	0.03	0.03	Without Clear
Image k	Blurred	Blurred	0.02	0.01	0.90	Blurred
Image l	Without Clear	Without Clear	0.09	0.11	0.11	Well Defined
Mean BCE	0.06	0.06	0.06	0.07	0.15	0.06
BCE σ	0.05	0.05	0.06	0.07	0.23	0.05

uncertain (Figure 4.6 (4a, c, d)). In some images, the prediction is uncertain in regions where the RGB images appear to show no surface water, which could be caused by cryoconite or other land features with similar spectral characteristics (Figures 4.6 (4e, h)), emphasising the complexity of predicting SGL and underlining the need for uncertainty estimates. Combining the uncertainty with the probabilities gives a complete picture of the likelihood that any given image contains supraglacial water.

4.3.2 Comparison with other methods

In the literature, all models offer a binary classification of SGLs, making it difficult to compare their performance to our approach. Therefore, we use a threshold to create a binary classification of lake/non-lake from our probabilistic prediction (Figures 4.6 (6a-l). We then calculate F1 scores for all images and compare the results of our INLA-SPDE approach with those of four others. Probabilistic predictions reveal bimodal peaks around 0 and 1 4.6 (5a-l)), and so we set a threshold of $\rho_i \geq 0.5$, i.e., the value at which it is more likely that a pixel is water than non-water (Figures 4.6 (6a-l)).

We compare five algorithms: INLA-SPDE, $NDWI_{GNIR}$, $NDWI_{BR}$, CNN, and RF, to classify supraglacial lakes from Sentinel-2 images. We evaluate the accuracy, robustness, and suitability of each of the algorithms. In certain cases, static thresholding approaches can lead to a decrease in accuracy, such as Image a for $NDWI_{GNIR}$ and Images b and h for $NDWI_{BR}$ (Figure 4.7 (4a, 5b, 5h)). The standard deviation in F1 scores for all images is 6.5 and 11.5 times higher for the static thresholding approaches than for the INLA-SPDE approach (Table 4.3). In the literature, the CNN (M. Dirscherl et al., 2021; Yuan et al., 2020) and RF (D. e. a. Corr, n.d.; M. Dirscherl et al., 2020; Hu et al., 2022, In Prep.) approaches have been more effective and reliable than static thresholding. In this study, we report F1 scores of 0.97 and 0.98 for CNN and RF, respectively (Table 4.3). The INLA-SPDE model, with an F1 score of 0.97 for all images, performs similarly to the RF and CNN models (Figure 4.7); however, it is the only model that provides a probabilistic prediction (Table 4.3). By selecting the threshold in our approach, the user can decide between a more cautious estimation of the lake area (that is, a high threshold) and a more generous estimation (that is, a lower threshold), based on probabilistic predictions. Although other approaches have the option of selecting

thresholds, the thresholds are not based on probabilistic predictions. Therefore, if a threshold is chosen that is not ideal, it requires more manual processing. Compared to other methods, our technique relies on probability to determine thresholds, meaning that we can decide when a pixel is more likely to be water than not.

Table 4.3: **The F1 scores for each of the test images for the INLA-SPDE, $NDWI_{GNIR}$, $NDWI_{BR}$, CNN, and RF models were calculated. F1 score combines precision and recall representing the balance between a model's ability to correctly identify positive instances and its tendency to avoid false positives.**

Value	INLA-SPDE > 0.5	$NDWI_{GNIR}$	$NDWI_{BR}$	CNN	RF	Predicted Border Condition
Image a	0.95	0.50	0.96	0.94	0.97	Without Clear Border
Image b	0.99	0.98	0.63	0.99	0.99	Blurred
Image c	0.94	0.95	0.85	0.95	0.97	Blurred
Image d	0.92	0.92	0.95	0.95	0.96	Blurred
Image e	0.93	0.95	0.96	0.95	0.97	Without Clear Border
Image f	0.99	0.98	0.99	0.99	0.99	Well Defined
Image g	0.99	0.91	0.99	0.99	0.99	Well Defined
Image h	0.96	0.92	0.19	0.99	0.99	Blurred
Image i	0.99	0.99	0.98	0.99	0.99	Well Defined
Image j	0.98	0.94	0.83	0.97	0.99	Without Clear Border
Image k	0.99	0.99	0.97	0.99	0.99	Blurred
Image l	0.96	0.95	0.96	0.96	0.97	Without Clear Border
All Images	0.97	0.95	0.74	0.97	0.98	-
Mean	0.97	0.92	0.85	0.97	0.98	-
σ	0.02	0.13	0.23	0.02	0.01	-

4.3.3 Binary classification of lake features

In order to produce a binary classification of lake/not-lake pixels, we apply a threshold to our probabilistic prediction. A threshold of $\rho_i \geq 0.99$ will classify all

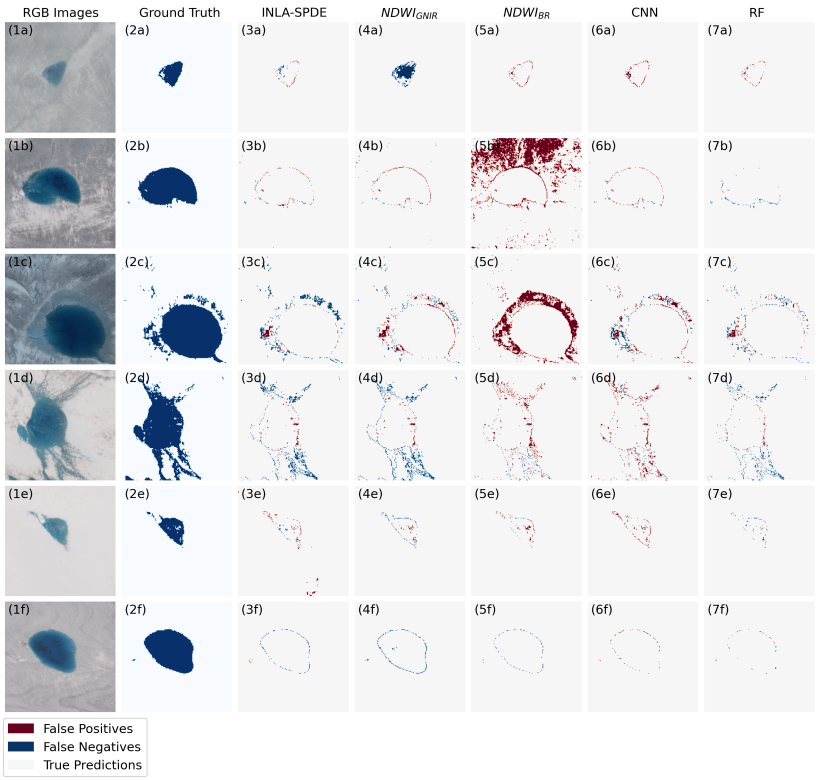


Figure 4.7: **I) Comparison between outputs of each classification algorithm and the 'true' binary labels in the first 6 test images. (1) RGB images, (2) True labels of binary water (dark blue) / non-water (pale blue), (3) Truth minus INLA-SPDE model with threshold $\rho_i \geq 0.5$, (4) Truth minus $\text{NDWI}_{\text{GNIR}}$ index, (5) Truth minus NDWI_{BR} index, (6) Truth minus CNN model, (7) Truth minus RF model. Red indicates misclassified surface water (false positives), while blue indicates misclassified non-water (false negatives), compared to the true labels.**

pixels in the image that the algorithm is 99% sure are surface water, while $\rho_i \geq 0.01$ will classify all pixels with a non-zero probability of being surface water. Here, we apply a threshold of $\rho_i \geq 0.50$ (50% probability); however, we also explore whether it is possible to further optimise the threshold (Figure 4.8). The optimal threshold in terms of the F1 score is found to be $\rho_i \geq 0.48$; however, we note that the F1 score

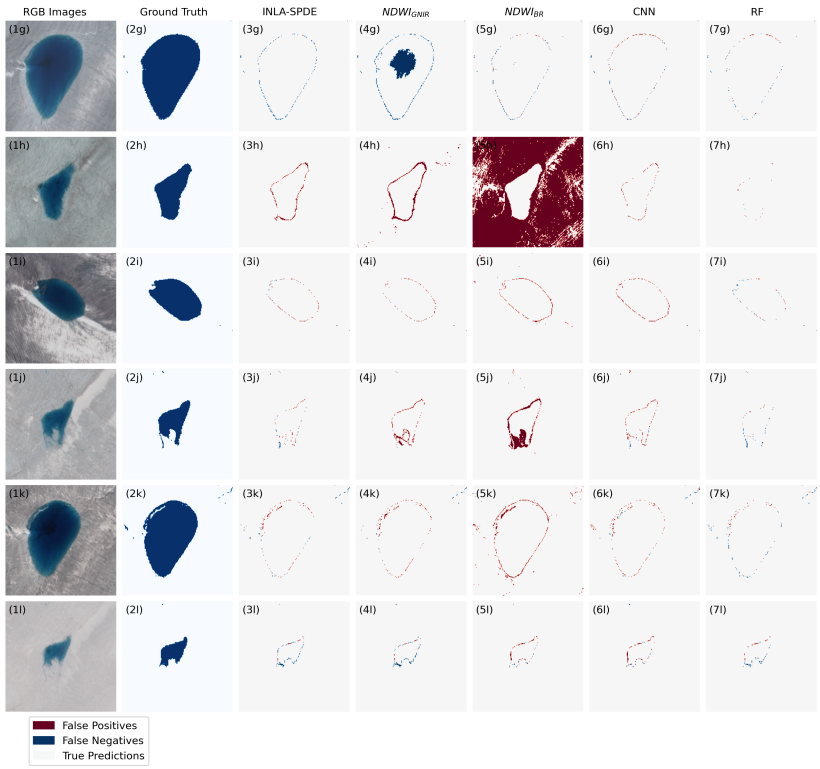


Figure 4.7: **II) Plots of the differences between outputs of each of the comparison models and the true binary labels in the remaining 6 test images. 1: RGB images; 2: True labels of binary water (dark blue) / non-water (pale blue); 3: Our INLA-SPDE model with threshold $\rho_i \geq$0.5; 4: $NDWI_{GNIR}$ index; 5 $NDWI_{BR}$ index; 6 CNN model; 7 RF model. Red indicates misclassified surface water (false positives), while blue indicates misclassified non-water (false negatives), compared to the true labels.**

associated with this threshold (0.970) is only marginally better than that associated with our threshold of 0.5 (0.969).

The optimal area ratio, i.e., the ratio of the predicted area of all images to the extent of SGL from the ground truth labels, was found to be 0.999 when the threshold was set at $\rho_i \geq$0.34. This threshold provides the best approximation for the area of

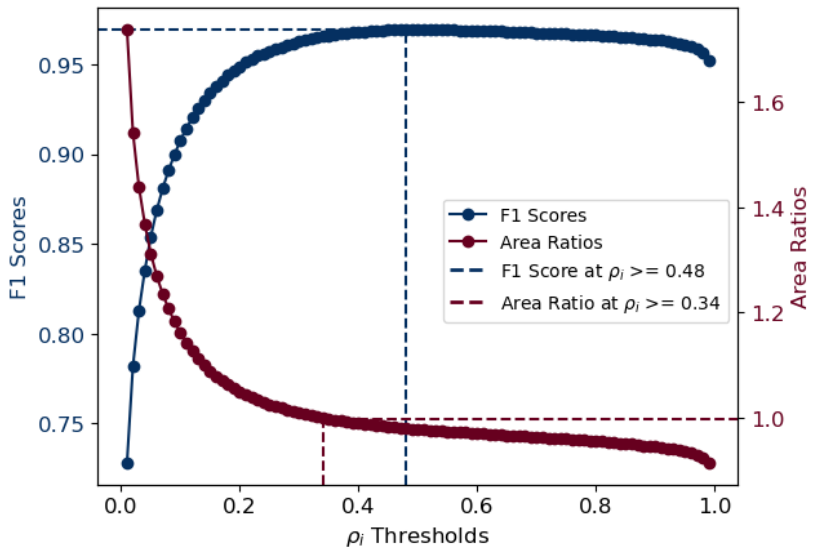


Figure 4.8: **F1 scores and ratios of the predicted area of all images to the area of SGL from the ground truth labels for varying thresholds on the probabilistic prediction (ρ_i) of supraglacial meltwater in the test images. The optimal thresholds according to the maximum F1 score ($\rho_i \geq$0.48) and the area ratio closest to 1 ($\rho_i \geq$0.34) are identified on each curve.**

surface meltwater, but may lead to false positives and false negatives because the area is not classified in the right places. This is reflected in the slightly reduced F1 score (0.966). Given the small number of samples examined in this study, it is necessary to evaluate the suitability of the chosen threshold by testing it on new data that accurately reflect the study area or on a larger group of lakes to determine whether the threshold is valid.

The difference between the area ratios for the other thresholds is small (0.980 for $\rho_i \geq$0.48 and 0.978 for $\rho_i \geq$0.50). Therefore, the choice of threshold for $\rho_i \geq$0.48 and $\rho_i \geq$0.50 has only a small effect on the result of the final binary classification, in fact, any threshold in the range 0.34$\geq \rho_i \geq$0.85 results in an F1 score that rounds to 0.97 at two decimal places. To demonstrate binary classification, we apply thresholds at

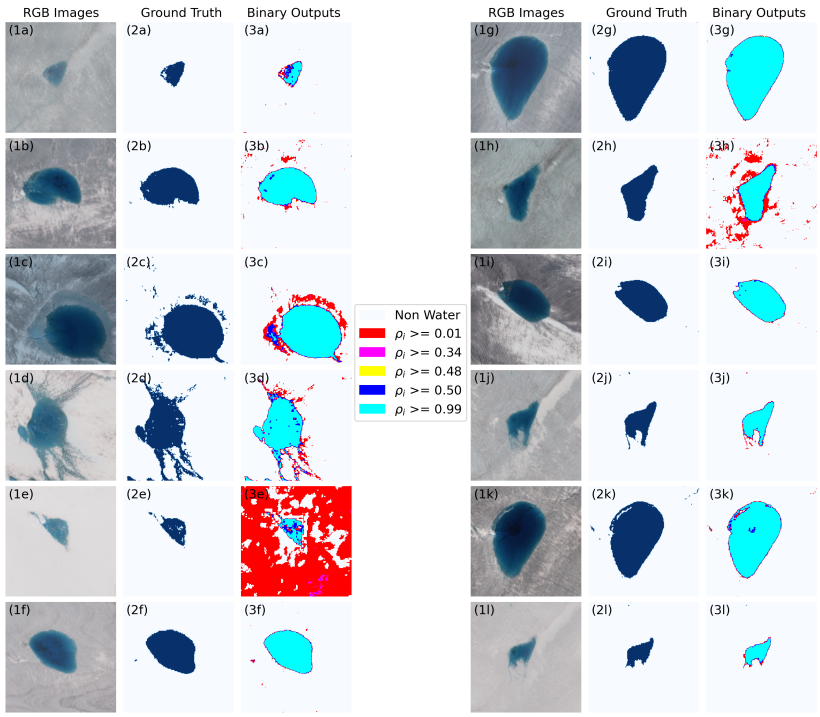


Figure 4.9: **Plots of the binary classification of our INLA-SPDE model with different thresholds overlaid, where pixels are coloured turquoise for $\rho_i \geq$0.99, the underlying pixels are coloured for all thresholds. 1: RGB images; 2: True labels of binary water (dark blue) / non-water (pale blue); 3: INLA-SPDE with thresholds $\rho_i \geq$0.01 (red); $\rho_i \geq$0.34 (pink); $\rho_i \geq$0.48 (yellow); $\rho_i \geq$0.50 (blue); and $\rho_i \geq$0.99 (cyan)**

$\rho_i \geq 0.01$ (all non-zero probabilities), $\rho_i \geq 0.34$ (optimal area ratio threshold), $\rho_i \geq$0.48 (optimal F1 score threshold), $\rho_i \geq 0.5$ (50% probability threshold), and $\rho_i \geq 0.99$ (the certain water threshold). Compared to true labels, the classification of $\rho_i \geq$0.01 is much too tolerant, with obvious false positives present in (Figure 4.9 (3b, c, e, h)). Meanwhile, the threshold for surface water classification is too strict for $\rho_i \geq$0.99, with many more false negative classifications present in most test images (Figure 4.9). The thresholds at $\rho_i \geq$0.48 and $\rho_i \geq$0.50 reveal little difference,

confirming the hypothesis that the selection of threshold for $\rho_i \geq 0.50$ is reliable.

4.4 Conclusion

This study used a Bayesian statistical technique, INLA-SPDE, to develop probabilistic classifications of water features on the Greenland Ice Sheet surface in Sentinel-2 optical satellite imagery and the associated pixel-level uncertainty, for the first time. The model demonstrated robust performance by generating probabilistic predictions regarding the likelihood that a given pixel represents surface meltwater. The predictions depend on the input satellite data and one of three predicted border conditions (well defined, blurred, and without a clear border). The result can be utilised in a variety of ways, such as establishing a limit to generate a map of features that are highly probable to be surface water or one that displays all potential surface water with any non-zero likelihood. The output characterises the uncertainty in the prediction, with the mean of the posterior distribution representing the expected value of the response variable given the model and the observed data, ranging from 0 (certain non-water) to 1 (certain water). When comparing these predictions with ground truth data, a binary cross-entropy score of 0.06 was calculated. To derive a binary classification for comparison with established models in the field, we applied a threshold ($\rho_i \geq 0.5$) at the point where the probability that the predicted pixel is more likely to be surface water than non-water is greater. In our analysis, we scrutinise the threshold selection process and find that a range of thresholds, ranging from 0.34 to 0.85 for ρ_i, yields a minimal change in the overall F1 score.

Our model outperforms the thresholding techniques of $\text{NDWI}_{\text{BlueRed}}$ and $\text{NDWI}_{\text{GreenNIR}}$, and its performance is comparable to that of the CNN and RF algorithms. However, there are areas where further improvements could be made.

The multinomial logistic regression algorithm, which constitutes the initial phase of the INLA-SPDE approach, demonstrates reasonable performance and plays a pivotal role in our methodology. However, it stands out as a potential avenue for improvement. The binary cross-entropy scores allude to room for refinement within the multinomial logistic regression algorithm. While the INLA-SPDE approach, employing a threshold of $\rho_i \geq 0.50$, performs well, it is marginally outperformed by the RF algorithm. Further analysis revealed that the optimal threshold is $\rho_i \geq 0.48$, and although the performance at $\rho_i \geq 0.50$ is comparable, there is potential for optimisation by fine-tuning the probabilistic prediction threshold.

Despite the absence of a native *predict* function, INLA is equipped with the tools to make predictions on new data using posterior distributions of model parameters. These posterior distributions encapsulate the estimated coefficients associated with the predictor variables, intercepts, and random spatial effects derived from the fitted model. Therefore, it becomes feasible to estimate the new linear predictor for any new data points. The estimated coefficients and intercepts can be obtained from the parameter *fixed effects* in the INLA model, while spatial random effects are included in the parameter *random effects*. Through this process, we can estimate predictive distributions and generate probabilistic predictions for the new data. Although this approach involves a higher level of complexity compared to the native *predict* functions available to CNN and RF algorithms, it remains a viable strategy for the scalable deployment of these models, even on an ice sheet scale. It is important to note that the training and testing of our model were confined to lakes in southwest Greenland. For future research, we recommend expanding the training data set to encompass a more representative selection of lakes on the Greenland ice sheet. This broader data collection would enhance the model's applicability to larger geographical scales. Careful consideration must be given to the computational requirements of applying the method to large amounts of

satellite data and applying the models to images larger than 192×192 pixels. To handle this Sentinel-2, tiles can be broken down into regions or patches of size 192×192 pixels. Batch processing, memory optimisation, and parallelisation can facilitate concurrent processing, address memory limitations, and improve computational efficiency. A significant advantage of the model lies in its ability to provide probabilistic predictions alongside binary classification, a feature not typically found in conventional CNN and RF models. Probabilistic predictions provide a measure of the uncertainty associated with each prediction, offering a more versatile and informative model than binary classification alone. Our approach establishes a proof of concept for employing Bayesian statistics to quantify supraglacial lakes on the Greenland Ice Sheet. This approach has the potential to provide probabilistic assessments of lake areas at the ice sheet scale. In prospective applications on the ice sheet scale, this approach can offer probabilistic assessments of lake areas. These assessments are of paramount importance, as they provide robust uncertainty estimates regarding the extent of meltwater present on the ice sheet. Such insights are critical for understanding the implications of trends in supraglacial hydrology and how they respond to the influence of a warming climate.

4.5 Data & code availability

S2 imagery are freely available at `scihub.copernicus.eu`. A script to download the data was extracted from GitHub (Hagolle, Olivier, 2015), however, with changes in the data structure in the repository, the script may no longer be effective. Alternatively, images can be downloaded from Google Cloud Storage using Python scripting (Nunes, Vasco, 2016).

The code used to produce the supraglacial hydrology features is written in R and

can be accessed on Zenodo and GitHub (D. Corr, 2023b).

Chapter 5

Synthesis

This thesis was created to address the knowledge gaps that were preventing better understanding of the distribution and temporal evolution of supraglacial hydrology on both ice sheets. To do this, I achieved the four primary objectives outlined in the thesis.(1) To quantify the presence and extent of SGH water bodies across the entire West Antarctic Ice Sheet and Antarctic Peninsula. (2) To evaluate the effectiveness of machine learning to create large SGHF datasets. (3) To identify the distribution and changes of SGHF across the entire Greenland Ice Sheet in the last decade. (4) To provide a more accurate estimate of the uncertainty in the classification of supraglacial lakes at the pixel level.

At the beginning, I provided an overview of the two major ice sheets, Antarctica and Greenland, and the hydrological system that connects their supra, en, and subglacial environments. I then examined the importance and known distributions of supraglacial hydrology and described the methods used to measure supraglacial features from remote sensors (Chapter 1). Subsequently, I presented the approach I took to develop a dual NDWI algorithm and use a threshold to distinguish supraglacial hydrology in previously unstudied areas of the West Antarctic Ice Sheet and the Antarctic Peninsula during the 2017 melting season. This produced a

high-fidelity data set that can be used to train and validate machine learning models, as well as to supplement a snapshot of supraglacial hydrology on the East Antarctic Ice Sheet (Stokes et al., 2019) during the same period (Chapter 2). Next, I explained the process in which I trained a random forest algorithm to differentiate between surface meltwater and used a radiative transfer model to determine the depth of the water and, as a result, the volume across the entire Greenland ice sheet at a monthly rate, from 2014 to 2022. This revealed seasonal and interannual trends in surface hydrology on the ice sheet scale (Chapter 3). Finally, I introduced a Bayesian inference model to examine the spatial connection between water pixels for supraglacial lakes with three different boundary conditions in western Greenland. This model provides a probabilistic evaluation of supraglacial hydrology on the Greenland Ice Sheet, quantifying the uncertainty in the prediction of supraglacial hydrology pixel by pixel, which has never been done before (Chapter 4). In this synthesis, I provide a summary of the main results of each chapter and how the thesis has achieved its goals and objectives. I then synthesise the results from all the chapters, before discussing the key limitations and explaining the decisions made during the thesis. Finally, I suggest potential areas for future research that have been identified from key findings and limitations.

5.1 Summary of principal findings

In this section, I provide a summary of the techniques used in this thesis and discuss the main discoveries and conclusions of each chapter, as well as how they contribute to our understanding of the hydrological system of the Greenland Ice Sheet.

5.1.1 An updated inventory of supraglacial lakes and channels on the West Antarctic Ice Sheet

The accurate measurement of supraglacial hydrology is of great scientific value in understanding the dynamics of the Antarctic ice sheet mass balance and its effect on global sea-level changes. The existence of meltwater on the ice surface has implications for the stability of ice shelves and the movement of grounded ice, particularly through processes such as hydrofracturing and the transportation of meltwater to the bedrock.

The availability of increased satellite data enabled the development of continent-wide inventories for the West Antarctic Ice Sheet and Antarctic Peninsula. This chapter outlines supraglacial lakes and streams in West Antarctica using a semi-automated Dual-NDWI approach applied to a large data set of over 2000 images obtained from the Sentinel-2 and Landsat-8 satellites during January 2017. A novel k-means clustering technique is used to categorise water features into lakes and channels, allowing for a more comprehensive understanding of their hydrological connections. The accuracy of the approach is evaluated by comparing it with a manually delineated reference dataset from three Antarctic test sites.

This study has identified 10,478 supraglacial features, including 10,223 lakes and 255 channels, covering an area of 119.4 km². Of these features, 27.3% are located on grounded ice, 54.9% are located on floating ice shelves, and 17.8% exist on the grounding line. This research adds to the more than 65,000 lakes (more than 1,300 km2) that were observed around the peak of the melt season in January 2017 on the East Antarctic Ice Sheet (EAIS) (Stokes et al., 2019), forming the first continent-wide assessment to help quantify the mass balance of Antarctica and its effect on global sea level rise. This inventory provides a baseline for future investigations and a reference point for monitoring Antarctica's surface hydrology in a changing climate,

as well as expanding our understanding of supraglacial hydrology in previously unexplored regions, such as the Getz Ice Shelf margin. Furthermore, the data set generated by this work is a valuable resource for Earth system research, allowing for the calibration, validation, and development of advanced approaches for real-time monitoring of lake dynamics in response to climate change. Due to the size of continental data sets, sophisticated and automated techniques, such as machine learning algorithms, are essential. This research offers a high-precision data set to train and validate these scientific endeavours.

5.1.2 Supraglacial hydrology coverage has increased on the Greenland ice sheet over the last decade

This chapter presents a comprehensive analysis of supraglacial meltwater on the Greenland Ice Sheet (GrIS) over the past decade. A random forest algorithm is adapted and validated to accurately identify meltwater from a data set of 144,000 Sentinel-2 and Landsat-8 images taken between May 1st and September 30th from 2013 to 2022. The approach used has a high spatial (10-30 m) and temporal (monthly) resolution. A radiative transfer model is then applied to estimate the surface water volume, providing monthly snapshots of meltwater distribution and volume.

Analysis of the GrIS has revealed a consistent seasonal pattern in the behaviour of supraglacial hydrology. Generally, meltwater is seen in May, reaching its peak in July and August. After that, the surface water usually refreezes or drains, mainly in September. Over the past decade, there has been a marked increase in both the extent and volume of SGHF. Although there is still considerable interannual variability, this trend implies a connection between rising temperatures, attributed to climate change, and the increased presence of supraglacial hydrology in the GrIS.

However, more research is needed over a longer period of time to better understand the long-term trends of surface hydrology. Additionally, the study found a considerable increase in total meltwater across various drainage basins on the ice sheet, particularly in the north, east, and south regions, indicating the effect of climate change on the increasing prevalence of supraglacial hydrology in these regions.

5.1.3 A new method for probabilistic prediction of supraglacial lakes on the southwest Greenland Ice Sheet

This chapter has presented a Bayesian statistical framework, the Integrated Nested Laplace Approximation with Stochastic Partial Differential Equation (INLA-SPDE), to predict the probability of the presence of supraglacial water at the pixel level within satellite imagery. Bayesian inference, as a probabilistic framework, inherently takes into account uncertainty by incorporating prior beliefs, modelling data variability with likelihood distributions, and estimating posterior distributions that represent updated beliefs about parameters while still preserving the inherent uncertainty. Multinomial logistic regression was used to categorise each pixel into one of three potential lake border conditions: well defined, blurred, or lacking a clear border. To account for the spatial relationships between pixels, a Matérn covariance function was implemented on an SPDE mesh. The model was trained to identify the complex relationships between various factors, including random spatial effects, predictor variables (Sentinel-2 indices and reflectance bands), and the corresponding water/non-water labels. The INLA-SPDE model provides a probability distribution over the likelihood that a pixel is water, instead of a binary classification (water or not water). The prediction in this method is represented by the mean (or expected value) of the probability distribution, while the uncertainty

is represented by the standard deviation. Probabilistic predictions have several advantages over binary classifications. For example, they can provide information on the likelihood of rare events and allow for more nuanced approximations of the hydrological system.

In addition, they provide the flexibility to set thresholds according to the level of confidence we wish to have in a particular binary classification. I investigated the effect of different thresholds on the probabilistic prediction to create a binary water/non-water classification. In my analysis, I found that a threshold of $\rho_i \geq 0.50$, i.e. the point at which it is more likely that a pixel is water than non-water, performs almost as well as the optimal threshold ($\rho_i \geq 0.48$). Classification using a threshold of $\rho_i \geq 0.50$ was tested against a manually labelled ground truth and the results were compared to two static thresholding approaches ($NDWI_{GNIR}$, $NDWI_{BR}$) and RF and CNN algorithms. The INLA-SPDE approach was found to be superior to static thresholding approaches and was comparable to RF and CNN. Although the sample size was small and the approach was only validated in a small region of southwest Greenland, it is encouraging to know that the INLA-SPDE approach has potential. Furthermore, it is comparable to the RF algorithm which was used in the assessment of Greenland SGHF in Chapter 3. This chapter highlights the effectiveness of the Bayesian INLA-SPDE framework, which provides a promising way to accurately predict and classify SGLs within optical satellite imagery. Despite some limitations of the model architecture, such as the lack of a native *predict* function, the process of estimating predictions by retrieving estimated coefficients, intercepts, and spatial random effects is explained.

5.2 Synthesis of principal findings

In this section, I synthesise the principal findings discussed in this thesis and explain how this research has improved our knowledge of the distribution and temporal evolution of supraglacial hydrology on the Greenland and Antarctic ice sheets, developing advanced methods that others can use to improve the mapping of supraglacial hydrology in key glaciological regions.

Through these studies, I have confirmed the widespread presence of supraglacial hydrology on the Antarctic Ice Sheet, including identifying meltwater on the Getz Ice Shelf for the first time, and that the extent of supraglacial hydrology has increased on the GrIS through the last decade, particularly in the north, east, and south. Analysis of the GrIS has improved our understanding of seasonal and interannual trends in surface meltwater more than ever before. I have confirmed the seasonal trends of supraglacial hydrology on the GrIS, with meltwater becoming apparent in May, reaching a peak in July/August and beginning to decrease throughout September in each year. I have determined that there has been a positive trend in the extent and volume of SGHF over the past decade, suggesting that temperature increases linked to climate change are leading to greater supraglacial hydrology on the Greenland Ice Sheet. Quantification of uncertainty is difficult for conventional classification algorithms. Therefore, I developed a model to provide probabilistic predictions of surface meltwater likelihood and associated pixel-level uncertainty, in addition to binary classifications. Probabilistic predictions provide valuable information on prediction uncertainties that can improve decision-making processes beyond binary classifications, which is more informative than conventional CNN and RF models.

I have developed and validated three novel models: dual-NDWI, RF, and Integrated Nested Laplace Approximation with Stochastic Partial Differential

Equation (INLA-SPDE). Data sets have been provided that can be used by the community to train, validate, and evaluate more advanced models. For the first time, a model has been provided that quantifies uncertainty at the pixel level in the identification of supraglacial hydrology on ice sheets. Through the development and application of these three models, our understanding of the distribution and temporal evolution of supraglacial hydrology on the ice sheet scale has been improved.

5.3 Major limitations

This thesis examines a number of important knowledge gaps; however, the models and methods used have certain restrictions. It is essential to recognise and address these limitations in order to accurately interpret the results and implications of this study. This will also provide justification for the decisions made during the construction of the models and writing this thesis.

5.3.1 Data availability

The use of optical satellite imagery for the classification of supraglacial hydrology has several inherent restrictions related to the availability of data, the common issue of cloud cover, and the influence of the solar elevation angle on the reflection of incoming solar radiation. The primary constraint on the availability of optical images is due to the revisit times of each sensor. Sentinel-2, consisting of twin satellites (Sentinel-2A and Sentinel-2B), operates on a short 5-day revisit cycle. Landsat-8 adheres to a less frequent 16-day revisit cycle. The shorter revisit time of Sentinel-2 results in improved temporal resolution, which affords the opportunity to capture the dynamic processes of supraglacial hydrology in a reasonable time

frame. However, large amounts of meltwater can be transported across or through the ice sheet over days or even hours through channels or moulins, respectively. Therefore, even a 5-day repeat cycle may not capture all variability in the hydrological life cycle.

A further constraint is the dependence on clear-sky observations. Cloud cover obstructs the line of sight between the satellite sensor and the surface of the ice sheet, making affected regions inaccessible for analysis and further reducing data availability. Furthermore, the unpredictability and variability of cloud cover pose significant challenges to the analysis of supraglacial hydrology by optical satellite images of the Greenland and Antarctic ice sheets during the summer melt seasons (Halberstadt et al., 2020; Lutz, Bahrami, and Braun, 2023). The dynamic nature of cloud formation and dissipation means that the extent and distribution of cloud cover can change rapidly, even within the course of a single day. This introduces temporal inconsistencies in optical data acquisition, making it difficult to obtain cloud-free imagery for specific regions of interest. In addition, when clouds are present, they scatter and absorb sunlight, altering the spectral characteristics of the incoming radiation. This can result in variations in the recorded reflectance values, which may not be solely indicative of the surface properties. Similarly, reflectance values are distorted by solar elevation angles lower than 20°(Halberstadt et al., 2020). In such images, the surface meltwater is too spectrally similar to its surroundings, and algorithms struggle to classify supraglacial lakes (M. Moussavi et al., 2020). The tilt of the Earth's axis and its orbit around the sun cause the solar elevation angles to vary significantly over Greenland and Antarctica during the melt season. This leads to data that is not usable at the start and end of the melt season. Data gaps resulting from deficient repeat cycles, cloud cover, and solar elevation angles hamper the generation of continuous time-series data sets, which are crucial for monitoring surface meltwater over extended periods. Consequently,

the pre-processing and interpretation of the collected optical satellite data is a challenge.

5.3.2 Misclassifications and model performance

The systematic validation conducted in this thesis confirms the robust performance of the developed models. However, extending these models to untested contexts, such as the East Antarctic Ice Sheet or differing years and regions, requires careful consideration. Calibration and validation with new data remain pivotal to ensure results' reliability. The models rely on spectral differences between water and terrestrial features, which pose challenges in cases of shallow or mixed water-ice compositions, leading to false classifications. The spectral similarities between surface meltwater and features such as rock, cloud, shadow, crevassing, slush, and blue ice are an additional challenge. Furthermore, atmospheric conditions, the angle of sunlight, and the surface properties of the ice sheet can all affect the spectral signature of the ice, potentially leading to misclassifications. Machine learning models require a substantial amount of labelled training data, which can be complex and time-consuming to obtain for precise ground truth data on supraglacial hydrology. Furthermore, the selection of relevant features and their interactions can influence the model's performance, and the optimal feature set may vary depending on the specific ice sheet and environmental conditions. The RF in particular assumes that the input variables are independent, which may not be true in practise, potentially affecting the classification results. To overcome the absence of a dedicated *predict* function in the INLA package, an alternative approach must be used to generate model predictions. INLA has the tools to make predictions on new data using posterior distributions of the model parameters. These distributions contain the estimated coefficients, intercepts, and random

spatial effects of the fitted model. This allows us to estimate the linear predictor for any new data points and generate probabilistic predictions. Although this approach is more complex than the *predict* functions available to CNN and RF algorithms, it is still a viable strategy to deploy these models on an ice sheet scale. Although these classifiers excel at identifying the presence of water, they lack the capability to provide depth or volume information, limiting a comprehensive understanding of ice sheet hydrology dynamics.

5.3.3 Computational requirements

The roll-out of methods on continental scales requires considerable computational resources. Remote sensing data processing and analysis require extensive computational operations, especially when dealing with large remote sensing and environmental data sets or executing complex algorithms. These computations require considerable computational resources, such as high-performance computing clusters or cloud-based platforms, which may not be accessible to all researchers or organisations. In Chapter 3, I analysed more than 144,000 satellite images (250 TB equivalent), necessitating efficient data storage and retrieval solutions, although the iterative processes I developed removed the need to download all images before application. As higher resolution products with improved temporal sampling become available, computational demands will increase and the need for efficient processing, storage, and prediction of data will become essential. In conclusion, the computational needs in remote sensing applications are considerable and can be limiting factors for researchers and organisations, particularly those with limited access to high-performance computing resources and expertise. To address these limitations, more efficient algorithms should be explored, which use distributed computing environments and

explore cloud-based solutions to democratise access to computational resources.

5.4 Suggestions for future work

This section examines how the techniques presented in this thesis can be advanced and used to generate new knowledge and insights in the future. I look at how the methods discussed in Chapters 3 and 4 can be extended to new areas such as the Antarctic Ice Sheet, high mountain glaciers, and sea ice, which would give us a more comprehensive view of surface meltwater and a more precise estimate of the uncertainty in our classification. Additionally, I consider applying models to future data acquisitions, older satellite missions, and non-optical satellite imagery. This would provide us with a better understanding of the temporal evolution of satellite imagery, with repeat monitoring in the years to come, reaching back to the 1970s, and providing estimates over winter. Furthermore, a machine learning approach to mapping supraglacial lake depth features combining ICESat-2 data and optical imagery is proposed. Lastly, I investigate the potential for applying the algorithms I have developed during my Ph.D. to applications outside the supraglacial hydrology field.

5.4.1 Expansion of methods to wider cryosphere

The use of RF and INLA-SPDE models is advantageous for determining the extent of surface meltwater in the cryosphere. These models take advantage of the power of machine learning and Bayesian inference to address the challenges of monitoring surface meltwater using remote sensing in dynamic environments. RF models are suitable for frequent and repeated monitoring of optical satellite imagery due to their capacity to process complex, high-dimensional data and non-linear

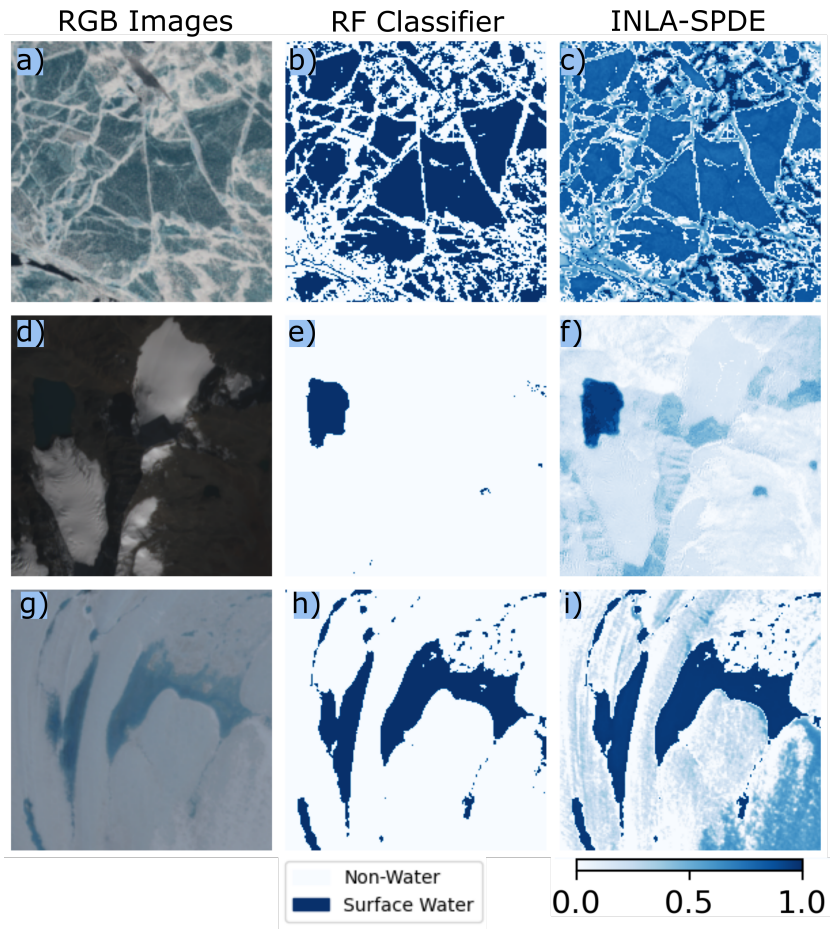


Figure 5.1: **Expansion of RF (b, e, h) and INLA-SPDE (c, f, i) models to encompass a broader range of cryospheric regions. The figure displays RGB images, binary RF classifications, and probabilistic predictions using INLA-SPDE based on Sentinel-2 imagery. The areas depicted include surface water on Arctic sea ice (a, b, c), Tibetan high mountain glaciers (d, e, f), and the Amery Ice Shelf on the EAIS (g, h, i). The Sentinel-2 tile IDs used for these regions are T09XWH_20170703T215051 for sea ice, T46RET_20221023T042831 for high mountain glaciers, and T41DPA_20220127T034619 for the Amery Ice Shelf.**

associations. Training RF models on optical satellite imagery is possible to accurately predict the extent of surface meltwater in Antarctica (M. Dirscherl et al.,

2020), Greenland (Hu et al., 2022), high mountain glaciers (C. S. R. Smith, 2022; Wendleder et al., 2021), and sea ice (Han et al., 2016; Miao et al., 2015). However, these studies represent smaller-scale studies than the continental mapping of surface meltwater I report in Chapter 3, and widespread decadal-scale monitoring of surface water in Antarctica, valley glaciers, and sea ice has yet to be undertaken. Bayesian frameworks, which estimate the extent of surface meltwater while explicitly accounting for spatial dependencies and uncertainties, have not been applied as widely. These models are particularly useful for producing probabilistic maps of the extent of surface meltwater with quantified uncertainty; however, I was unable to identify studies using Bayesian inference to predict surface meltwater in optical satellite imagery. To evaluate the performance of the models developed in Chapters 3 and 4, example images of melt ponds on Arctic sea ice, a glacial lake on a high mountain glacier in the Boshula Mountains, Tibet, and supraglacial lakes on the Amery Ice Shelf, EAIS (Figure 5.1) were used. Although the performance of the models has not been quantified due to the lack of validation data, a visual and qualitative assessment suggests that they perform satisfactorily. To ensure scientific rigour, I suggest training and validating the models on a more up-to-date data set that is more representative of the area of interest.

A major obstacle in attempting to use these techniques on a large scale in Antarctica is computational requirements. To evaluate the decadal scale of SGH on Greenland in Chapter 3, 250 TB of optical satellite data were analysed. Since the Antarctic ice sheet covers an area more than 8 times larger than Greenland, the computational needs for an assessment of this magnitude will be much greater. Detecting and classifying supraglacial lakes on high mountain glaciers and sea ice presents additional difficulties for those on ice sheets. The smaller size of lakes compared to those on ice sheets, cloud cover, shadows cast from mountains, varying levels of turbidity of glacial lakes, seasonal snows, and frozen glacial lake

surfaces add to the complexity of recognising glacial lakes on high mountain glaciers (Wangchuk and Bolch, 2020). The complexity and variability of the formation and evolution of melt ponds and the limited availability of observational data make the identification of melt ponds on sea ice a difficult challenge (Niehaus et al., 2023). With additional training and a comprehensive validation of the techniques, the approaches discussed in this thesis could be extended to other parts of the cryosphere beyond those on which I have used them.

5.4.2 Extension of methods to new data, older missions, & non-optical sensors

The RF algorithm developed in Chapter 3 is highly efficient, scalable, reliable, and accurate. To generate the results in this chapter, I created efficient data processing pipelines to download the necessary data from cloud storage repositories, apply the algorithms, and save the output to either local or cloud storage. With minor modifications, this pipeline could be used to generate results on a regular basis from newly released optical satellite data. The use of operational frameworks on optical images is widespread (e.g. Bazzi et al., 2021; Souza et al., 2022; Sudmanns et al., 2020). However, accuracy estimates can differ according to seasonal and interannual trends (Bazzi et al., 2021). Furthermore, the quality and availability of the data is not consistent in space and time due to factors such as cloudiness, location on Earth, and the acquisition plan (Sudmanns et al., 2020). The deployment of remote sensing algorithms with continuous application to newly acquired optical satellite imagery presents a variety of additional challenges. Adaptive models must take into account changing environmental conditions, allowing adjustments based on historical data for sustained accuracy. To ensure data consistency throughout melt seasons and years, automated quality assurance protocols must be

incorporated into the operational framework to identify and address sensor anomalies or seasonal discrepancies. Validation against ground-truth data should be conducted regularly to maintain the algorithm's precision and reliability throughout its operational lifespan. The algorithm must be able to scale, allowing it to efficiently handle increasing data loads without significantly increasing computational demands. Maintenance, including bug fixes, updates for new satellite missions, and efficiency improvements, is essential for long-term operational viability. Finally, cost management strategies must be implemented to ensure operational sustainability, taking into account data acquisition, storage, processing, and maintenance costs. A system to classify supraglacial hydrology on new optical satellite data could enable researchers to quickly detect and respond to changes in SGHF, help evaluate flood risks, and improve our knowledge of the cryosphere and its reaction to environmental changes. However, resolving this particular problem will not provide any insight into the long-term trends of SGHF. Expanding our models to include older optical sensors could.

Optical satellite imagery has been around since the 1970s, beginning with the Landsat-1 mission in 1972 and the AVHRR in 1970 (Table 1.2). This imagery has been used to track a 47-year history of SGHF on George VI Ice Shelf, Antarctic Peninsula (Figure 5.2, Barnes et al., 2021). However, data quality and spatial resolution can be a challenge when dealing with older sensors, as the lower resolution may limit the effectiveness of modern algorithms, and modern sensors have different multispectral bands. To create a consistent and homogeneous dataset, radiometric and geometric corrections, calibration, and cross-calibration must be performed to ensure the compatibility of the data acquired by various sensors (Kaita et al., 2022; Sousa and Small, 2017). Extending methods to previous missions and providing an approach to apply the models to new data would provide a better understanding of the SGH temporal evolution. The simplest

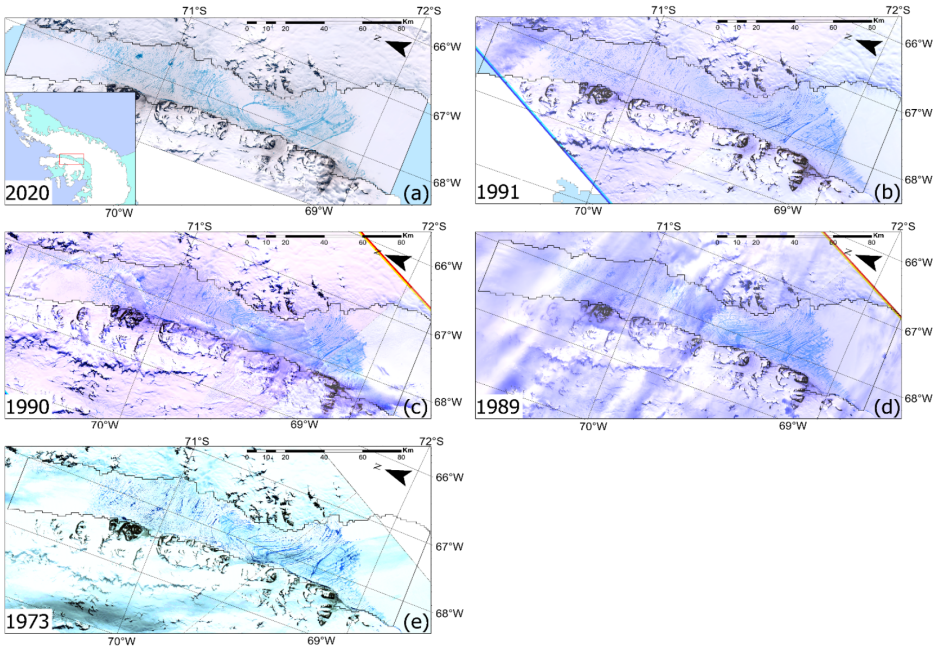


Figure 5.2: **Optical satellite imagery showing supraglacial lake coverage on George VI Ice Shelf in years when there is extensive lake coverage (Barnes et al., 2021). (a) Sentinel-2 image acquired on 19 January 2020. The inset shows the location of the study area on the AP. Landsat-4,5 images acquired on (b) 4 February 1991, (c) 15 January 1990, (d) 28 January 1989, and (e) Landsat-1 image acquired on 9 January 1973.**

approach to applying my models to these older data sets is to train once again on data representative of the area of interest. The most significant obstacle to advancing the use of older Landsat sensors, which offer nearly consistent spatial and temporal resolutions (see Table 1.2), is the creation or procurement of ground truth data for the training and validation of machine learning techniques. By manually enhancing existing data sets (Barnes et al., 2021), it is possible to use them for this purpose. If these methods were extended to the start of the optical remote sensing era, they would provide a much-needed perspective on the effects of

human-induced climate change on the SGHF of the ice sheets.

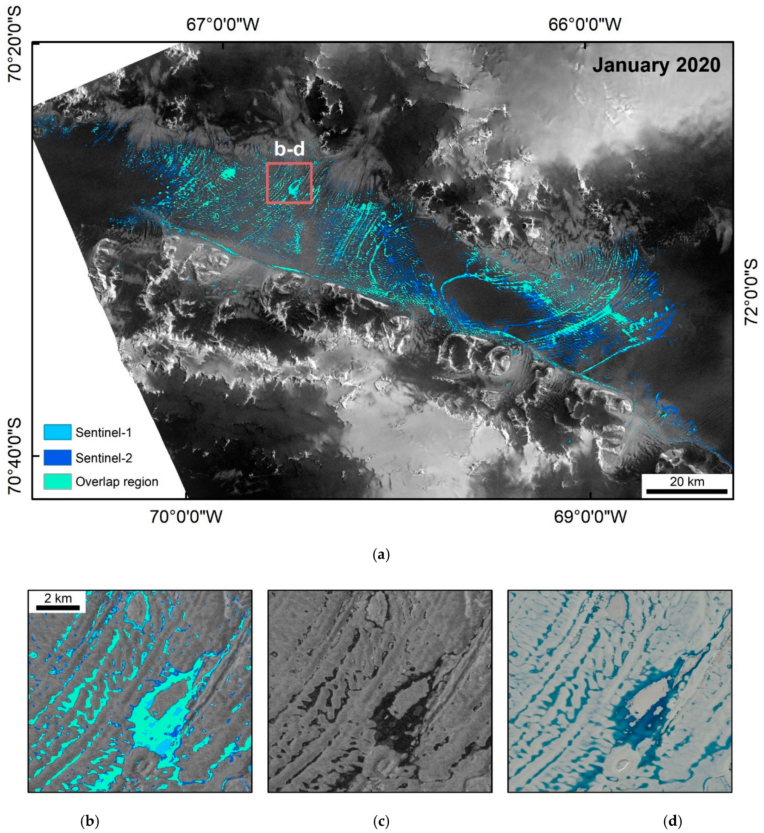


Figure 5.3: **Supraglacial meltwater identified by a dual (Sentinel-1 and Sentinel-2) sensor approach over George VI Ice Shelf, Antarctic Peninsula during January 2020. The light blue colour represents regions where only Sentinel-1 lake classifications are present, the dark blue colour illustrates regions with only Sentinel-2 lake classifications, and the green colour denotes overlapping lake mappings. The SAR image (a–c) is a Sentinel-1 monthly minimum backscattering product derived from all Sentinel-1 acquisitions of the George VI Ice Shelf in January 2020. (d) Displays A Sentinel-2 RGB image of the same area during the same month (d).**

It is essential to understand the storage and drainage of supraglacial lakes

during winter to understand the temporal dynamics of SGHF on the GrIS and AIS (Dunmire et al., 2020; Kaita et al., 2022; Koenig et al., 2015; Lampkin et al., 2020; Maier et al., 2023). Synthetic aperture radar (SAR) can be used to estimate the size of buried lakes during winter or when optical imagery is blocked by clouds, which would help us better observe short-term lake changes and measure the amount of surface water held in the ice sheet system. The radiation emitted by radar devices is not affected by clouds, darkness, or the angle of the sun. RF and INLA-SPDE algorithms have been used to calculate the extents of surface meltwater in SAR imagery (Herbert, Camps, and Vall-Llossera, 2021; Schröder et al., 2020; Wangchuk and Bolch, 2020), however, SAR imagery cannot be used to determine lake depths. Dual sensor approaches enable for a more comprehensive and detailed mapping of supraglacial lake coverage (e.g., Figure 5.3, M. Dirscherl et al., 2021). Future research could apply the models developed here for dual- or multi-sensor to better understand the temporal dynamics of SGHF on both ice sheets through all seasons and enhance our quantification of uncertainty at the pixel level. Radar devices have their own problems when used to sense supraglacial hydrology; for example, they do not detect melt over blue ice and confuse surface melt with near-surface melt (Husman et al., 2023), generally offer coarser spatial resolution and cannot retrieve depth or volume measurements.

5.4.3 Machine learning approach to estimate depth

To gain a more comprehensive understanding of the supraglacial hydrology of the GrIS, it is necessary to improve the depth estimates of supraglacial lakes (L. Melling et al., 2023). Surface reflectance and depth (Equation 1.3) used to calculate depth estimates from optical images reaches a plateau at depths between 1 and 3 m (L. Melling et al., 2023), thus necessitating a different method to quantify the depths

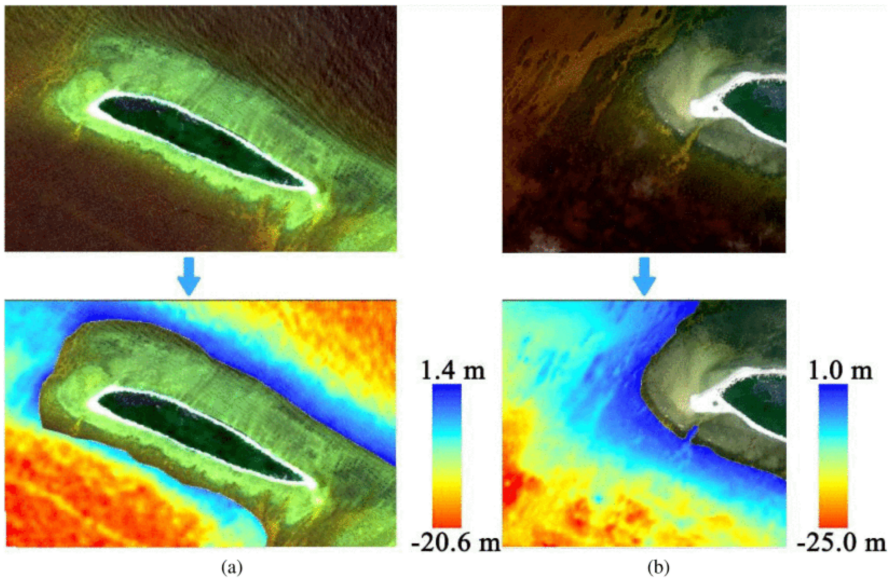


Figure 5.4: **Estimates for water depth retrieval from Resources Satellite Three (a) and Worldview-2 (b) using a CNN over marine shallow water around North Island, Xisha Islands in the South China Sea (Ai et al., 2020).**

and associated uncertainties of supraglacial features. ArcticDEM, DEM-differencing, and ICESat-2 provide more precise estimates (Das et al., 2008; Datta et al., 2019; Fair et al., 2020; Lampkin and VanderBerg, 2011; Porter et al., 2018), however, they are limited in terms of spatial and temporal coverage compared to optical imagery. A machine learning approach that uses accurate estimates (e.g., (Datta et al., 2019; Fair et al., 2020)) as a baseline dataset for training, validation, and testing could provide improved estimates of supraglacial lake depth, and therefore volume. This approach, for example, could use ICESat-2 tracks as a depth label (ground truth), and reflectance values from optical images or calculated band indices to form a relationship between the values, exploiting the relationship between surface reflectance and depth similarly to surface reflectance and depth. CNNs have been used to retrieve the depth of the water in shallow marine water from optical images in a similar approach (Figure 5.4, Ai et al., 2020).

5.4.4 Application to other fields

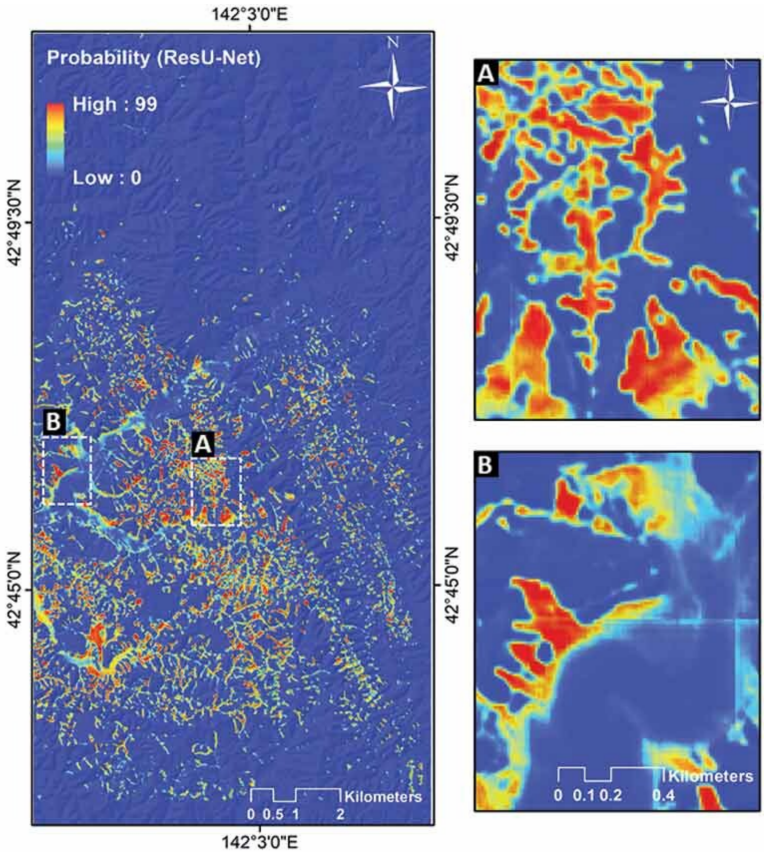


Figure 5.5: **Landslide probability prediction map obtained using a ResU-Net (a type of CNN) on multi-temporal Sentinel-2 images. The pixel values closer to 1 represent a higher probability of the landslide class (Ghorbanzadeh, Gholamnia, and Ghamisi, 2022).**

The use of optical satellite sensors in remote sensing has been demonstrated to improve our understanding of supraglacial hydrology on ice sheets. However, the methods developed and adapted can be applied to a variety of other fields, such as urban planning, disaster management, and agricultural management. For example, NDWI approaches can be used to detect floods and manage reservoirs

(Al Balasmeh, Karmaker, and Babbar, 2020; Condeça, Nascimento, and Barreiras, 2022; Liuzzo et al., 2020). Random forest algorithms are also widely used for change in land use and cover, forest tracking, crop monitoring, and mangrove analysis (Amini et al., 2022; Orynbaikyzy, Gessner, and Conrad, 2022; Sharifi, Felegari, and Tariq, 2022; Svoboda et al., 2022; R. Zhang et al., 2022). Bayesian inference has been applied to the mapping of soil matter, the classification of tree species, and the atmospheric correction (Axelsson et al., 2021; C. Yang et al., 2023; Yin, Lewis, and Gómez-Dans, 2022). Future work could apply these models to predict the probability of avalanches or landslides, which has been achieved using deep learning (Figure 5.5, Ghorbanzadeh, Gholamnia, and Ghamisi, 2022).

5.5 Concluding remarks

This thesis has significantly advanced our understanding of the distribution and temporal evolution of supraglacial hydrology on ice sheets. For the first time, it has revealed the presence of widespread supraglacial hydrology on the West Antarctic Ice Sheet and Antarctic Peninsula, providing a benchmark data set for the future development and testing of machine learning algorithms (Chapter 2). Furthermore, it has revealed increasing trends in supraglacial hydrological activity on the Greenland ice sheet since 2014 (Chapter 3). Finally, it has provided an uncertainty quantification at the pixel level, something that had not been done before (Chapter 4). To achieve these feats, I have developed three methods to classify supraglacial hydrology, dual NDWI, random forest, and Bayesian inference through INLA-SPDE. This thesis has made a natural progression in method development, from manually intensive NDWI to adaptive random forest to uncertainty quantification provided by Bayesian inference, to improve our knowledge of supraglacial meltwater systems. This knowledge is essential for understanding the dynamics of ice sheets and glaciers and their response to a changing climate. Supraglacial hydrology is a key indicator of the effects of climate change on glaciers and ice sheets. The behaviour of supraglacial meltwater, such as the formation of lakes and drainage systems on ice surfaces, is closely related to the dynamics of ice sheets and the rise of sea level. By understanding and monitoring supraglacial hydrology, we can gain insight into the vulnerability of ice sheets and their potential contribution to global sea-level changes. Additionally, this research is essential to manage and mitigate the far-reaching environmental consequences of glacial melt, from the influence of freshwater ecosystems and sediment transport to addressing the risks of glacial outburst floods that can affect downstream communities and infrastructure. In conclusion, quantifying supraglacial hydrology

is critical not only for understanding the cryosphere's reaction to climate change, but also for predicting and mitigating its broader environmental and societal impacts.

Appendix A

An inventory of supraglacial lakes & channels across the West Antarctic Ice Sheet

A.1 Summary of candidate NDWI thresholding methods

Table A.1: **Sensitivity (Sen.), Specificity (Spec.) and Accuracy (Acc.) values averaged (Mean) across three test sites: Amery, George VI and Bach Ice Shelves for S2 and L8 Sensors.**

Method	Mean S2 sen.	Mean L8 sen.	Mean S2 spec.	Mean L8 spec.	Mean S2 acc.	Mean L8 acc.
1: NDWIGNIR>0.300	0.793	0.675	0.991	0.996	0.98	0.984
2: NDWIGNIR>0.175	0.869	0.752	0.987	0.998	0.981	0.985
3: Dual-NDWI	0.853	0.776	0.991	0.997	0.983	0.987

NDWI thresholding methods (Equations 1.1 and 1.2) have been implemented using Sentinel-2 and Landsat-8 satellite imagery. Here, we summarise three

Table A.2: **Standard deviation (SD) of the values for sensitivity (sen.), specificity (spec.) and accuracy (acc.) across three test sites: Amery, George VI and Bach ice shelves for S2 and L8 Sensors.**

Method	SD S2 sen.	SD L8 sen.	SD S2 spec.	SD L8 spec.	SD S2 acc.	SD L8 acc.
1: NDWIGNIR>0.300	0.142	0.257	0.013	0.006	0.012	0.005
2: NDWIGNIR>0.175	0.07	0.222	0.014	0.003	0.013	0.009
3: Dual-NDWI	0.069	0.176	0.01	0.004	0.011	0.005

candidate thresholding approaches that were assessed during methodological assessment. To determine the thresholds, we compared the output from two separate thresholds on the $NDWI_{GNIR}$, >0.300 (Method 1) (Stokes et al., 2019) and a lower threshold of >0.175 (Method 2) to maximise the delineated lake area, with the Dual-NDWI thresholding approach (Method 3) presented in Chapter 2.2 Data and methods. In addition to the threshold on $NDWI_{GNIR}$, we explored the use of band filters (M. Moussavi et al., 2020) to remove false positives from rock, cloud and other problem pixels as discussed in Section 2.2.1.2 Cloud, rock masking and elimination of slush, blue-ice and shaded pixels.

Method 1 comprises a simple thresholding of $NDWI_{GNIR}$ classification, by excluding any pixels with an $NDWI_{GNIR}$ value less than or equal to 0.300. This method was used to map lakes in January 2017 in East Antarctica 1, and is the basis for three methods considered here. Method 2 utilised a lower threshold of 0.175 on $NDWI_{GNIR}$, for more complete delineation of supraglacial hydrology. However, due to the lower threshold, more non-lake pixels were misclassified as lake pixels. To reduce such misclassifications, band filters were introduced to remove some cloud, rock, slush and shaded areas (M. Moussavi et al., 2020). Method 3 combines the $NDWI_{GNIR}$ (lower threshold of 0.16) with a second, $NDWI_{BR}$ classification was introduced. This $NDWI_{BR}$ classifier, was given a threshold of 0.18 (M. Moussavi

et al., 2020). Additional band filters were applied in Method 3 as in Method 2.

As in the analysis in Section 2.2 Accuracy assessment, we compared the output of the three methods to the manually delineated lakes and channels. We computed the mean values for sensitivity, specificity and accuracy (Equations 2.3, 2.4 and 2.5) for each test site (Amery, George VI and Bach) across both sensors (S2 and L8) (Table A.1). Based upon this assessment, the two best-performing methods are Methods 2 and 3. However, we selected the Dual-NDWI method because it performs well not only in terms of the average values, but also in terms of the standard deviation for both sensors (Table A.2). This indicates greater stability between sites, which is important when applying the method across other sites during the ice sheet wide roll out.

Appendix B

Supraglacial hydrology coverage has increased on the Greenland ice sheet over the last decade

B.1 Random forest

Random forest (RF) is a supervised learning algorithm that requires labelled training data (Breiman, 2001). It is made up of a group (forest) of unrelated decision trees. Decision trees are used to plot some statistical probability analysis made up of three types of nodes, connected by branches. The first node (the 'root') is the part from which all other nodes eventually branch, that is, it is the first question asked. The end of the decision path, where there are no more divisions or branches, is named the 'leaf' node. There exists any number of internal nodes between the root and leaf nodes. Each node represents a decision or a test on an attribute, while the branches, which connect them, denote the outcomes. Individual decision trees are built on a randomly sampled subset of training data. Each tree, which is grown to

some maximum extent, provides a classification decision or vote. RF contains many decision trees in the form of an ensemble classifier. After a given number of trees are generated, the most popular classification is determined. RF resists overfitting, which can be a major sources of error in machine learning processes (Breiman, 2001). For large data sets, such as the S2 and L8 archives used in this study, the major memory requirement is the storage of the data itself. The error rate in RF depends on two main factors: the correlation between any two trees and the strength of each tree in the forest. Increasing the correlation increases the forest error rate, while increasing the strength of an individual tree decreases the forest error rate (Belgiu and Drăguţ, 2016; Breiman, 2001; Kamiński, Jakubczyk, and Szufel, 2018; Pal, 2005).

B.2 Training data

Here we display the Sentinel-2 and Landsat-8 tiles used in the training, validation, and testing of the random forest algorithms.

Table B.1: **All tiles used in the training of the L8 and S2 RF algorithms. S2 tiles ending with '*' are also used in the transferability tests.**

L8 Tile Name	S2 Tile Name
LC08_L1TP_016002_20130428_20180301_01_T1	S2A_MSIL1C_20170530T152911_N0205_R111_T22WEB_20170530T152909
LC08_L1TP_005015_20130501_20170504_01_T1	S2A_MSIL1C_20170531T145921_N0205_R125_T22WEA_20170531T150133
LC08_L1TP_010003_20130504_20170504_01_T2	S2A_MSIL1C_20170531T145921_N0205_R125_T22WEV_20170531T150133
LC08_L1TP_010010_20130504_20170504_01_T1	S2A_MSIL1C_20170531T145921_N0205_R125_T22WFA_20170531T150133
LC08_L1TP_007013_20130819_20170502_01_T1	S2A_MSIL1C_20170531T145921_N0205_R125_T22WFV_20170531T150133
LC08_L1GT_019004_20130823_20170502_01_T2	S2A_MSIL1C_20170612T153911_N0205_R011_T26XNN_20170612T153906
LC08_L1TP_010003_20130824_20170502_01_T1	S2A_MSIL1C_20170615T141011_N0205_R053_T25WDS_20170615T141200

LC08_L1TP_007013_20140416_20170423_01_T1	S2A_MSIL1C_20170619T234731_N0205_R116_T21XWB_20170619T234734
LC08_L1TP_007013_20140502_20170423_01_T1	S2A_MSIL1C_20170710T145911_N0205_R125_T22WEA_20170710T150120
LC08_L1TP_016002_20140517_20170422_01_T1	S2A_MSIL1C_20170710T145911_N0205_R125_T22WEV_20170710T150120
LC08_L1TP_010003_20140523_20170422_01_T1	S2A_MSIL1C_20170710T145911_N0205_R125_T22WFA_20170710T150120
LC08_L1TP_010010_20140523_20170422_01_T1	S2A_MSIL1C_20170710T145911_N0205_R125_T22WFV_20170710T150120
LC08_L1TP_016002_20140602_20170422_01_T1	S2A_MSIL1C_20170722T171901_N0205_R012_T20XNK_20170722T171857
LC08_L1GT_019004_20140607_20170422_01_T2	S2A_MSIL1C_20170723T150911_N0205_R025_T22WEB_20170723T151115
LC08_L1TP_010003_20140608_20170422_01_T1	S2A_MSIL1C_20170724T143921_N0205_R039_T27XVE_20170724T143924
LC08_L1TP_005015_20140621_20170421_01_T1	S2A_MSIL1C_20170724T175911_N0205_R041_T20XMN_20170724T175913
LC08_L1TP_010003_20140624_20170421_01_T1	S2A_MSIL1C_20170730T145921_N0205_R125_T22WFB_20170730T150103
LC08_L1TP_016002_20140704_20170421_01_T1	S2A_MSIL1C_20170730T145921_N0205_R125_T26XNN_20170730T145915
LC08_L1TP_010010_20140710_20170421_01_T1	S2A_MSIL1C_20170731T160901_N0205_R140_T21XWB_20170731T160903
LC08_L1TP_016002_20140720_20170421_01_T1	S2A_MSIL1C_20170809T145921_N0205_R125_T22WEV_20170809T150205
LC08_L1TP_005015_20140723_20170421_01_T1	S2A_MSIL1C_20170812T150911_N0205_R025_T22WEA_20170812T150912
LC08_L1GT_019004_20140725_20170420_01_T2	S2A_MSIL1C_20170812T150911_N0205_R025_T22WEV_20170812T150912
LC08_L1TP_007013_20140806_20170420_01_T1	S2A_MSIL1C_20170812T150911_N0205_R025_T22WFA_20170812T150912
LC08_L1TP_005015_20140808_20180129_01_T1	S2A_MSIL1C_20170812T150911_N0205_R025_T22WFV_20170812T150912
LC08_L1TP_016002_20140821_20170420_01_T1	S2A_MSIL1C_20170825T165851_N0205_R069_T20XNK_20170825T170139
LC08_L1GT_019004_20140826_20170420_01_T2	S2A_MSIL1C_20170830T142931_N0205_R139_T27XVE_20170830T142933
LC08_L1TP_010010_20140827_20170420_01_T1	S2A_MSIL1C_20170908T145911_N0205_R125_T22WEA_20170908T150039
LC08_L1TP_005015_20150421_20170409_01_T1	S2A_MSIL1C_20170908T145911_N0205_R125_T22WEV_20170908T150039
LC08_L1GT_019004_20150423_20170409_01_T2	S2A_MSIL1C_20170908T145911_N0205_R125_T22WFA_20170908T150039
LC08_L1TP_016002_20150504_20170409_01_T1	S2A_MSIL1C_20170908T145911_N0205_R125_T22WFV_20170908T150039
LC08_L1TP_005015_20150507_20170411_01_T1	S2A_MSIL1C_20180526T145921_N0206_R125_T22WEA_20180526T201815*
LC08_L1TP_010003_20150510_20170409_01_T1	S2A_MSIL1C_20180526T145921_N0206_R125_T22WEV_20180526T201815*
LC08_L1TP_010010_20150510_20170409_01_T1	S2A_MSIL1C_20180526T145921_N0206_R125_T22WFA_20180526T201815*

LC08_L1TP_007013_20150521_20170408_01_T1	S2A_MSIL1C_20180526T145921_N0206_R125_T22WFV_20180526T201815*
LC08_L1TP_010003_20150611_20170408_01_T1	S2B_MSIL1C_20180622T185919_N0206_R013_T20XNQ_20180706T183955
LC08_L1TP_016002_20150621_20170407_01_T1	S2B_MSIL1C_20180710T150009_N0206_R125_T22WEA_20180710T202212*
LC08_L1TP_005015_20150624_20170407_01_T1	S2B_MSIL1C_20180710T150009_N0206_R125_T22WEV_20180710T202212*
LC08_L1GT_019004_20150626_20170407_01_T2	S2B_MSIL1C_20180710T150009_N0206_R125_T22WFA_20180710T202212*
LC08_L1TP_010010_20150627_20170407_01_T1	S2B_MSIL1C_20180710T150009_N0206_R125_T22WFV_20180710T202212*
LC08_L1TP_007013_20150708_20170407_01_T1	S2B_MSIL1C_20180715T140739_N0206_R053_T24WWU_20180715T175138
LC08_L1GT_019004_20150712_20170407_01_T2	S2B_MSIL1C_20180819T145959_N0206_R125_T22WEA_20180819T200111*
LC08_L1TP_010003_20150713_20170407_01_T1	S2B_MSIL1C_20180819T145959_N0206_R125_T22WEV_20180819T200111*
LC08_L1TP_010010_20150713_20170407_01_T1	S2B_MSIL1C_20180819T145959_N0206_R125_T22WFA_20180819T200111*
LC08_L1TP_016002_20150723_20170406_01_T1	S2B_MSIL1C_20180819T145959_N0206_R125_T22WFV_20180819T200111*
LC08_L1TP_010003_20150814_20170406_01_T1	S2B_MSIL1C_20180822T150749_N0206_R025_T22WEB_20180822T202152
LC08_L1TP_010010_20150814_20170406_01_T1	S2A_MSIL1C_20180913T145911_N0206_R125_T22WEA_20180913T171024*
LC08_L1TP_007013_20150825_20170405_01_T1	S2A_MSIL1C_20180913T145911_N0206_R125_T22WEV_20180913T171024*
LC08_L1GT_019004_20150829_20170405_01_T2	S2A_MSIL1C_20180913T145911_N0206_R125_T22WFA_20180913T171024*
LC08_L1TP_005015_20150912_20170404_01_T1	S2A_MSIL1C_20180913T145911_N0206_R125_T22WFV_20180913T171024*
LC08_L1TP_005015_20150928_20170403_01_T1	S2B_MSIL1C_20180921T151119_N0206_R025_T22WEV_20180921T201342
LC08_L1TP_016002_20160506_20170326_01_T1	S2B_MSIL1C_20190529T150809_N0207_R025_T22WEA_20190529T201431*
LC08_L1TP_010003_20160512_20180130_01_T1	S2B_MSIL1C_20190529T150809_N0207_R025_T22WEV_20190529T201431*
LC08_L1TP_007013_20160523_20180130_01_T1	S2B_MSIL1C_20190529T150809_N0207_R025_T22WFA_20190529T201431*
LC08_L1GT_019004_20160527_20170324_01_T2	S2B_MSIL1C_20190529T150809_N0207_R025_T22WFV_20190529T201431*
LC08_L1TP_010010_20160528_20170324_01_T1	S2B_MSIL1C_20190601T151809_N0207_R068_T27XVJ_20190601T171240*
LC08_L1TP_007013_20160608_20170324_01_T1	S2A_MSIL1C_20190601T160901_N0207_R140_T27XVH_20190601T180236*
LC08_L1TP_016002_20160623_20170323_01_T1	S2B_MSIL1C_20190613T141749_N0207_R096_T24WXV_20190613T152413
LC08_L1GT_019004_20160628_20170323_01_T2	S2B_MSIL1C_20190708T150809_N0208_R025_T22WEA_20190708T185223*
LC08_L1TP_010010_20160629_20170323_01_T1	S2B_MSIL1C_20190708T150809_N0208_R025_T22WEV_20190708T185223*

LC08_L1TP_007013_20160710_20170323_01_T1	S2B_MSIL1C_20190708T150809_N0208_R025_T22WFA_20190708T185223*
LC08_L1TP_005015_20160712_20180129_01_T1	S2B_MSIL1C_20190708T150809_N0208_R025_T22WFV_20190708T185223*
LC08_L1TP_010010_20160715_20180130_01_T1	S2A_MSIL1C_20190712T153911_N0208_R011_T21XXA_20190712T185647
LC08_L1TP_016002_20160725_20170322_01_T1	S2B_MSIL1C_20190713T155829_N0208_R097_T27XVH_20190713T193729*
LC08_L1TP_007013_20160726_20180130_01_T1	S2A_MSIL1C_20190713T164901_N0208_R026_T27XVJ_20190713T201443*
LC08_L1TP_005015_20160728_20180129_01_T1	S2B_MSIL1C_20190804T150019_N0208_R125_T22WEA_20190804T165643*
LC08_L1TP_010003_20160731_20170322_01_T1	S2B_MSIL1C_20190804T150019_N0208_R125_T22WEV_20190804T165643*
LC08_L1TP_010010_20160731_20170322_01_T1	S2B_MSIL1C_20190804T150019_N0208_R125_T22WFA_20190804T165643*
LC08_L1GT_019004_20160815_20170322_01_T2	S2B_MSIL1C_20190804T150019_N0208_R125_T22WFV_20190804T165643*
LC08_L1TP_010003_20160816_20170322_01_T1	S2B_MSIL1C_20190813T152819_N0208_R111_T27XVH_20190813T172427*
LC08_L1TP_005015_20160829_20170321_01_T1	S2B_MSIL1C_20190813T152819_N0208_R111_T27XVJ_20190813T172427*
LC08_L1TP_007013_20160912_20170321_01_T1	S2A_MSIL1C_20190901T150911_N0208_R025_T22WEA_20190901T170819*
	S2A_MSIL1C_20190901T150911_N0208_R025_T22WEV_20190901T170819*
	S2A_MSIL1C_20190901T150911_N0208_R025_T22WFA_20190901T170819*
	S2A_MSIL1C_20190901T150911_N0208_R025_T22WFV_20190901T170819*
	S2B_MSIL1C_20190903T150009_N0208_R125_T27XVH_20190903T183932*
	S2A_MSIL1C_20190903T154911_N0208_R054_T27XVJ_20190903T193001*
	S2A_MSIL1C_20200518T150921_N0209_R025_T22WEB_20200518T170904
	S2A_MSIL1C_20200528T150921_N0209_R025_T22WEV_20200528T170824
	S2B_MSIL1C_20200614T140739_N0209_R053_T26WME_20200614T160505
	S2B_MSIL1C_20200620T142749_N0209_R139_T27XVE_20200620T144408
	S2A_MSIL1C_20200623T170901_N0209_R112_T20XNK_20200623T204029
	S2A_MSIL1C_20200702T142001_N0209_R096_T27XVE_20200702T143521
	S2A_MSIL1C_20200717T150921_N0209_R025_T22WEB_20200717T170914
	S2A_MSIL1C_20200717T150921_N0209_R025_T22WEV_20200717T170914
	S2A_MSIL1C_20200717T150921_N0209_R025_T26XNN_20200717T170914

	S2A_MSIL1C_20200718T161911_N0209_R040_T21XWB_20200718T181502
	S2B_MSIL1C_20200718T170849_N0209_R112_T20XNK_20200718T190620
	S2A_MSIL1C_20200803T145921_N0209_R125_T27XVH_20200803T170904
	S2B_MSIL1C_20200905T155829_N0209_R097_T21XWB_20200905T193742
	S2B_MSIL1C_20200909T153819_N0209_R011_T22WED_20200909T193104
	S2B_MSIL1C_20200909T153819_N0209_R011_T26XNN_20200909T193104

B.3 Additional Validation

B.3.1 Validation of the algorithms and transferability testing

Here we show the confusion matrices for each of the validation tests and for each of the transferability tests discussed in Chapter 3.

Table B.2: **Example of a confusion matrix showing the labelled training data and predicted outcomes for the SGH and non-SGH classes.**

Labelled Data	Model Prediction: Non-SGH	Model Prediction: SGH
Non-SGH	True Negative	False Positive
SGH	False Negative	True Positive

Table B.3: **Confusion matrix for the validation of the algorithm trained and tested on the three distinct subsets of the data from the Landsat-8 training data set: a) Test 1; b) Test 2; c) Test 3.**

a)

Labelled Data	Model Prediction: Non-SGH	Model Prediction: SGH
Non-SGH	7539	78
SGH	8	717

b)

Labelled Data	Model Prediction: Non-SGH	Model Prediction: SGH
Non-SGH	7545	72
SGH	13	712

c)

Labelled Data	Model Prediction: Non-SGH	Model Prediction: SGH
Non-SGH	7525	92
SGH	23	702

Table B.4: **Confusion matrix for the validation of the algorithm trained and tested on the three distinct subsets of the data from the Sentinel-2 training data set: a) Test 1; b) Test 2; c) Test 3.**

a)

Labelled Data \ Model Prediction	Non-SGH	SGH
Non-SGH	67214	2370
SGH	249	16357

b)

Labelled Data \ Model Prediction	Non-SGH	SGH
Non-SGH	67171	2413
SGH	269	16337

c)

Labelled Data \ Model Prediction	Non-SGH	SGH
Non-SGH	67184	2400
SGH	311	16295

Table B.5: **Confusion matrix for the transferability testing of the algorithm a) trained on data from 2019 and tested on unseen data from 2018; b) trained on data from 2018 and tested on unseen data from 2019.**

a)

Labelled Data \ Model Prediction	Non-SGH	SGH
Non-SGH	49635	2824
SGH	305	24103

b)

Labelled Data \ Model Prediction	Non-SGH	SGH
Non-SGH	50774	970
SGH	2522	35771

Table B.6: **Confusion matrix for the transferability testing of the algorithm a) trained on data from July, August, and September and tested on unseen data from May; b) trained on data from May, August, and September and tested on unseen data from July; c) trained on data from May, July, and September and tested on unseen data from August; d) trained on data from May, July, and August and tested on unseen data from September e) trained on data from May, July, and August 2018, July, August, and September 2019 and tested on unseen data from September 2018 and May 2019.**

a) Model Prediction

Labelled Data	Non-SGH	SGH
Non-SGH	25743	1261
SGH	388	6722

b) Model Prediction

Labelled Data	Non-SGH	SGH
Non-SGH	24714	976
SGH	265	25137

c) Model Prediction

Labelled Data	Non-SGH	SGH
Non-SGH	25021	386
SGH	2215	23008

d) Model Prediction

Labelled Data	Non-SGH	SGH
Non-SGH	25487	615
SGH	631	4335

e) Model Prediction

Labelled Data	Non-SGH	SGH
Non-SGH	25061	256
SGH	541	7144

Table B.7: **Confusion matrix for the transferability testing of the algorithm trained on all data from the Watson River region and tested on unseen data from the NEGIS region.**

Model Prediction

Labelled Data	Non-SGH	SGH
Non-SGH	20565	383
SGH	807	12841

B.3.2 Comparing the sensors

Table B.8: **Comparison of the classification and volume estimates for Landsat-8 and Sentinel-2 sensors over various regions on the Greenland Ice Sheet.**

Date	Region	L8 Tile	L8 Area [km^2]	L8 Volume [km^3]	S2 Tile	S2 Area [km^2]	S2 Volume [km^3]
20170806	EG	232011	32.94	0.0207	T25WDS	78.24	0.0363
20170831	EG	232011	9.86	0.0059	T25WDS	52.01	0.0254
20180622	EG	232011	14.59	0.0071	T25WDS	26.04	0.0140
20200706	EG	232011	14.56	0.0173	T25WDS	30.11	0.0255
20200830	EG	232011	7.50	0.0056	T25WDS	28.18	0.0166
20170531	NWG	042001	4.55	0.0020	T21XVK	16.26	0.0071
20190521	NWG	042001	7.25	0.0019	T21XVK	13.04	0.0062
20200624	NWG	042001	30.44	0.0314	T21XVK	45.61	0.0426
20170512	NWG	037002	1.25	0.0002	T21XVK	2.95	0.0021
20180616	NWG	037002	7.67	0.0028	T21XVK	19.29	0.0087
20210710	NWG	037002	31.43	0.0323	T21XVK	55.16	0.0462
20190502	NEGIS	021001	3.08	0.0005	T25XEK	1.49	0.0006
20200715	NEGIS	013248	87.31	0.0191	T25XEK	38.09	0.0153
20170805	Watson	008013	20.45	0.0249	T22WEV	8.31	0.0068
20190608	Watson	008013	81.87	0.0905	T22WEV	91.35	0.0935
20190710	Watson	008013	23.49	0.0221	T22WEV	31.25	0.0258
20190928	Watson	008013	1.56	0.0006	T22WEV	25.41	0.0145
20200712	Watson	008013	64.31	0.1144	T22WEV	78.80	0.1197
20200813	Watson	008013	3.82	0.0050	T22WEV	78.80	0.0042
20210528	Watson	008013	10.68	0.0026	T22WEV	29.94	0.0097
20210917	Watson	008013	12.83	0.0138	T22WEV	8.18	0.0048
20170814	Watson	007013	40.28	0.0558	T22WEV	67.44	0.0645
20180918	Watson	007013	2.45	0.0007	T22WEV	5.74	0.0034
20190804	Watson	007013	13.96	0.0115	T22WEV	46.98	0.0238
20200619	Watson	007013	51.98	0.0513	T22WEV	55.15	0.0459
20210505	Watson	007013	9.74	0.0022	T22WEV	52.22	0.0149
20210622	Watson	007013	44.48	0.0503	T22WEV	56.03	0.0543
20170908	Watson	006013	4.08	0.0044	T22WEV	8.85	0.0052
20190610	Watson	006013	91.65	0.1149	T22WEV	130.17	0.1307
20190829	Watson	006013	4.75	0.0039	T22WEV	11.83	0.0074
20210530	Watson	006013	15.38	0.0054	T22WEV	27.01	0.0110
Cumulative			**750.20**	**0.7210**		**1219.92**	**0.8867**

B.4 Results

Table B.9: **The Pearson correlation coefficient (r) and corresponding P-Values for monthly trends in SGHF extents on GrIS drainage basins (Zwally et al., 2012) from 2014 through 2022.**

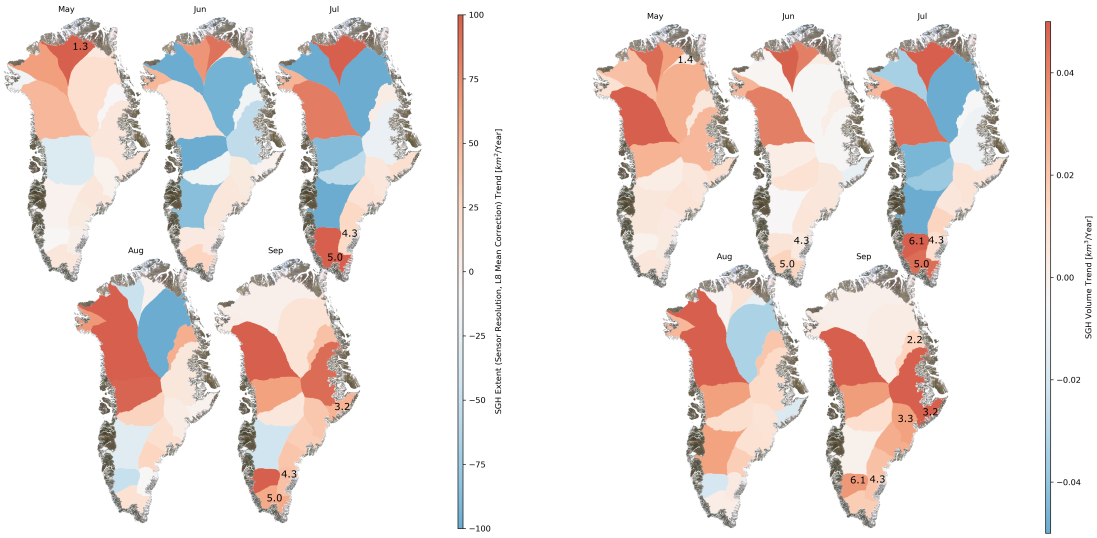
Basin	May		June		July		August		September		All Months	
	Pearson's r	P-Value	Pearson's r	P-Value	Pearson's r	P-Value	Pearson's r	P-Value	Pearson's r	P-Value	Pearson's r	P-Value
1.1	0.774	0.014	0.026	0.948	-0.576	0.104	0.486	0.185	NoData	NoData	0.013	0.934
1.2	0.885	0.002	0.596	0.090	-0.161	0.679	0.097	0.804	NoData	NoData	0.127	0.405
1.3	0.912	0.001	0.660	0.053	-0.061	0.875	0.111	0.776	NoData	NoData	0.159	0.296
1.4	0.732	0.025	0.349	0.358	-0.343	0.366	-0.150	0.699	NoData	NoData	-0.013	0.934
2.1	0.871	0.002	-0.174	0.654	-0.566	0.112	-0.354	0.350	0.295	0.441	-0.092	0.547
2.2	0.675	0.046	0.044	0.910	-0.596	0.090	0.468	0.204	0.600	0.087	0.135	0.376
3.1	0.724	0.028	0.296	0.439	0.185	0.633	0.391	0.298	0.782	0.013	0.383	0.009
3.2	0.634	0.066	0.653	0.056	0.750	0.020	0.499	0.172	0.875	0.002	0.601	0.000
3.3	0.618	0.076	0.691	0.039	0.766	0.016	0.604	0.085	0.903	0.001	0.549	0.000
4.1	0.763	0.017	0.635	0.066	0.761	0.017	0.587	0.097	0.827	0.006	0.642	0.000
4.2	0.839	0.005	0.699	0.036	0.622	0.074	0.431	0.246	0.886	0.001	0.494	0.001
4.3	0.708	0.033	0.689	0.040	0.933	0.000	0.631	0.068	0.912	0.001	0.649	0.000
5.0	0.783	0.013	0.750	0.020	0.884	0.002	0.631	0.068	0.813	0.008	0.629	0.000
6.1	0.226	0.559	0.193	0.618	0.679	0.044	0.149	0.702	0.772	0.015	0.285	0.058
6.2	0.316	0.408	-0.019	0.962	-0.464	0.208	0.538	0.135	0.349	0.357	0.052	0.734
7.1	0.366	0.333	0.086	0.826	-0.437	0.239	0.524	0.148	0.515	0.156	0.104	0.495
7.2	0.599	0.088	-0.102	0.794	-0.564	0.114	0.498	0.173	0.679	0.044	0.122	0.426
8.1	0.698	0.036	0.128	0.744	-0.241	0.532	0.768	0.016	0.548	0.127	0.191	0.210
8.2	0.814	0.008	0.754	0.019	0.086	0.826	0.679	0.044	0.227	0.556	0.301	0.045

Table B.10: **The Pearson correlation coefficient (r) and corresponding P-Values for seasonal and yearly trends in SGHF volume on GrIS drainage basins (Zwally et al., 2012) from 2014 through 2022.**

Basin	May		June		July		August		September		All Months	
	Pearson's r	P-Value	Pearson's r	P-Value	Pearson's r	P-Value	Pearson's r	P-Value	Pearson's r	P-Value	Pearson's r	P-Value
1.1	0.774	0.014	0.026	0.948	-0.576	0.104	0.486	0.185	NoData	NoData	0.013	0.934
1.2	0.885	0.002	0.596	0.090	-0.161	0.679	0.097	0.804	NoData	NoData	0.127	0.405
1.3	0.912	0.001	0.660	0.053	-0.061	0.875	0.111	0.776	NoData	NoData	0.159	0.296
1.4	0.732	0.025	0.349	0.358	-0.343	0.366	-0.150	0.699	NoData	NoData	-0.013	0.934
2.1	0.871	0.002	-0.174	0.654	-0.566	0.112	-0.354	0.350	0.295	0.441	-0.092	0.547
2.2	0.675	0.046	0.044	0.910	-0.596	0.090	0.468	0.204	0.600	0.087	0.135	0.376
3.1	0.724	0.028	0.296	0.439	0.185	0.633	0.391	0.298	0.782	0.013	0.383	0.009
3.2	0.634	0.066	0.653	0.056	0.750	0.020	0.499	0.172	0.875	0.002	0.601	0.000
3.3	0.618	0.076	0.691	0.039	0.766	0.016	0.604	0.085	0.903	0.001	0.549	0.000
4.1	0.763	0.017	0.635	0.066	0.761	0.017	0.587	0.097	0.827	0.006	0.642	0.000
4.2	0.839	0.005	0.699	0.036	0.622	0.074	0.431	0.246	0.886	0.001	0.494	0.001
4.3	0.708	0.033	0.689	0.040	0.933	0.000	0.631	0.068	0.912	0.001	0.649	0.000
5.0	0.783	0.013	0.750	0.020	0.884	0.002	0.631	0.068	0.813	0.008	0.629	0.000
6.1	0.226	0.559	0.193	0.618	0.679	0.044	0.149	0.702	0.772	0.015	0.285	0.058
6.2	0.316	0.408	-0.019	0.962	-0.464	0.208	0.538	0.135	0.349	0.357	0.052	0.734
7.1	0.366	0.333	0.086	0.826	-0.437	0.239	0.524	0.148	0.515	0.156	0.104	0.495
7.2	0.599	0.088	-0.102	0.794	-0.564	0.114	0.498	0.173	0.679	0.044	0.122	0.426
8.1	0.698	0.036	0.128	0.744	-0.241	0.532	0.768	0.016	0.548	0.127	0.191	0.210
8.2	0.814	0.008	0.754	0.019	0.086	0.826	0.679	0.044	0.227	0.556	0.301	0.045

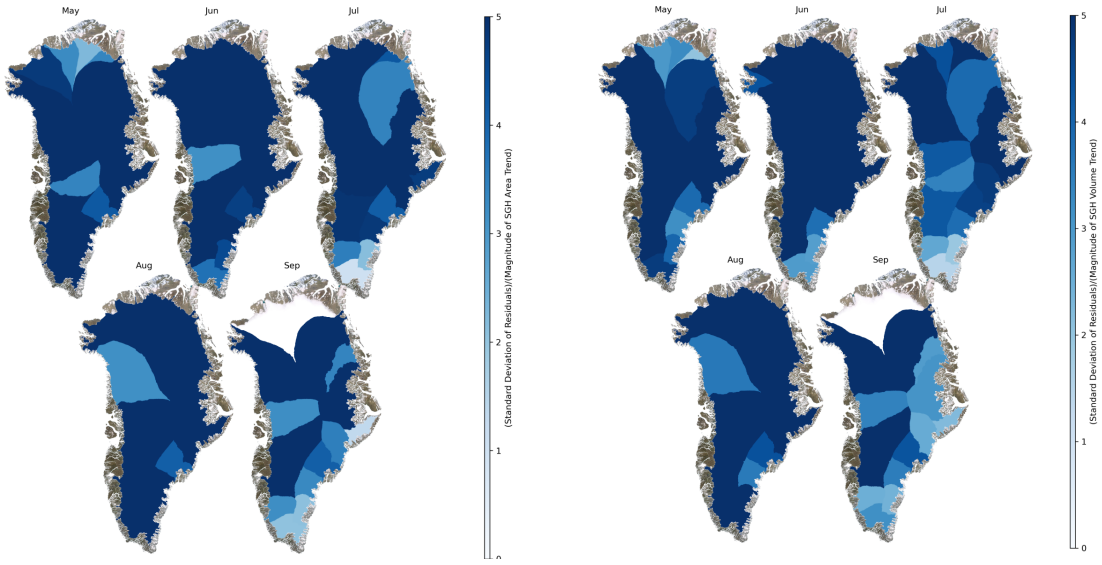
Table B.11: **The Pearson correlation coefficient (r) and corresponding P-Values for seasonal and yearly trends in SGHF occurrence on GrIS drainage basins (Zwally et al., 2012) from 2014 through 2022.**

Basin	May		June		July		August		September		All Months	
	Pearson's r	P-Value	Pearson's r	P-Value	Pearson's r	P-Value	Pearson's r	P-Value	Pearson's r	P-Value	Pearson's r	P-Value
1.1	0.774	0.014	0.026	0.948	-0.576	0.104	0.486	0.185	NoData	NoData	0.013	0.934
1.2	0.885	0.002	0.596	0.090	-0.161	0.679	0.097	0.804	NoData	NoData	0.127	0.405
1.3	0.912	0.001	0.660	0.053	-0.061	0.875	0.111	0.776	NoData	NoData	0.159	0.296
1.4	0.732	0.025	0.349	0.358	-0.343	0.366	-0.150	0.699	NoData	NoData	-0.013	0.934
2.1	0.871	0.002	-0.174	0.654	-0.566	0.112	-0.354	0.350	0.295	0.441	-0.092	0.547
2.2	0.675	0.046	0.044	0.910	-0.596	0.090	0.468	0.204	0.600	0.087	0.135	0.376
3.1	0.724	0.028	0.296	0.439	0.185	0.633	0.391	0.298	0.782	0.013	0.383	0.009
3.2	0.634	0.066	0.653	0.056	0.750	0.020	0.499	0.172	0.875	0.002	0.601	0.000
3.3	0.618	0.076	0.691	0.039	0.766	0.016	0.604	0.085	0.903	0.001	0.549	0.000
4.1	0.763	0.017	0.635	0.066	0.761	0.017	0.587	0.097	0.827	0.006	0.642	0.000
4.2	0.839	0.005	0.699	0.036	0.622	0.074	0.431	0.246	0.886	0.001	0.494	0.001
4.3	0.708	0.033	0.689	0.040	0.933	0.000	0.631	0.068	0.912	0.001	0.649	0.000
5.0	0.783	0.013	0.750	0.020	0.884	0.002	0.631	0.068	0.813	0.008	0.629	0.000
6.1	0.226	0.559	0.193	0.618	0.679	0.044	0.149	0.702	0.772	0.015	0.285	0.058
6.2	0.316	0.408	-0.019	0.962	-0.464	0.208	0.538	0.135	0.349	0.357	0.052	0.734
7.1	0.366	0.333	0.086	0.826	-0.437	0.239	0.524	0.148	0.515	0.156	0.104	0.495
7.2	0.599	0.088	-0.102	0.794	-0.564	0.114	0.498	0.173	0.679	0.044	0.122	0.426
8.1	0.698	0.036	0.128	0.744	-0.241	0.532	0.768	0.016	0.548	0.127	0.191	0.210
8.2	0.814	0.008	0.754	0.019	0.086	0.826	0.679	0.044	0.227	0.556	0.301	0.045



(a) **Trend in the SGHF extent.** *(b)* **Trend in the SGHF volume.**

Figure B.1: **Trends for SGHF extent (a) and volume (b) in the Zwally storage basins over the decade from 2014-2022, calculated for the products available at the source resolution of each sensor and with the sensor correction applied to Landsat-8 era data (2014-2016). Labelled basins are those where we find statistically significant trends in the extent and volume of SGHF.**



(a) **Ratio for SGHF extent.** *(b)* **Ratio for SGHF volume.**

Figure B.2: **The ratios of the standard deviation of the residuals of the trend to the magnitude of the trend for the SGHF extent (a) and volume (b) in the Zwally storage basins. In most cases, where the basins (Figures B.1a, B.1b) show no significant trend, this ratio is many times that of the basins with a significant trend.**

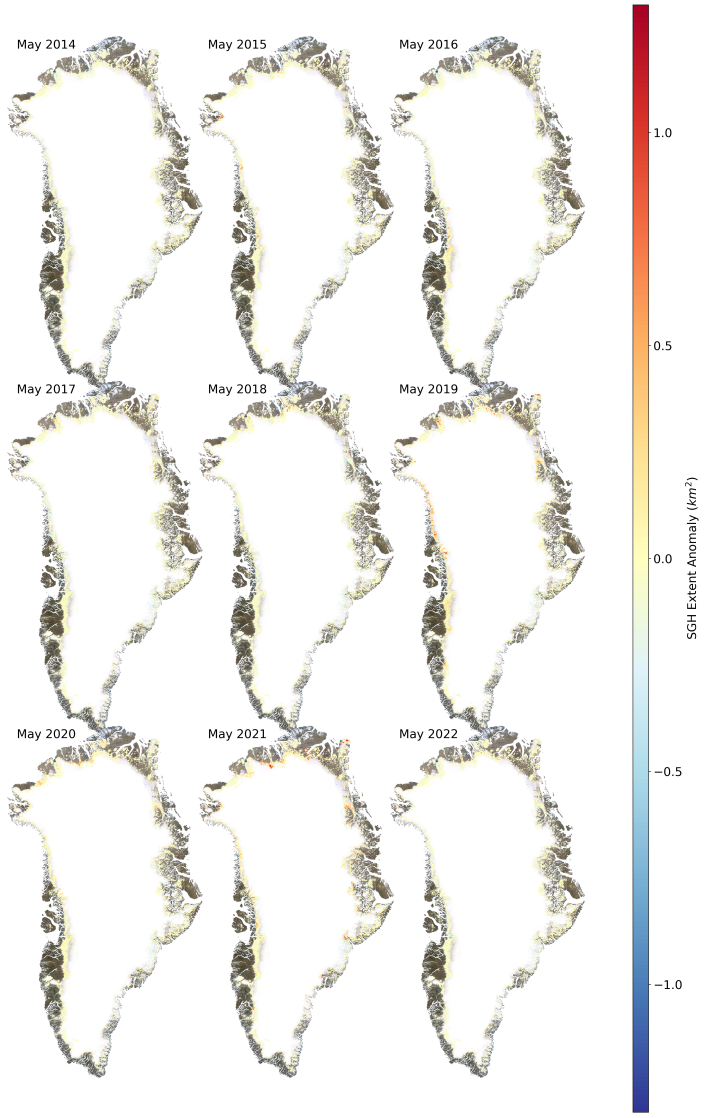


Figure B.3: **(a) The supraglacial hydrology extent anomaly for May compared to the mean extent in 1200 m grid cells on the Greenland Ice Sheet. The anomaly was calculated, taking the difference of the mean cell value of the 9 years in each cell from the extent value for each year.**

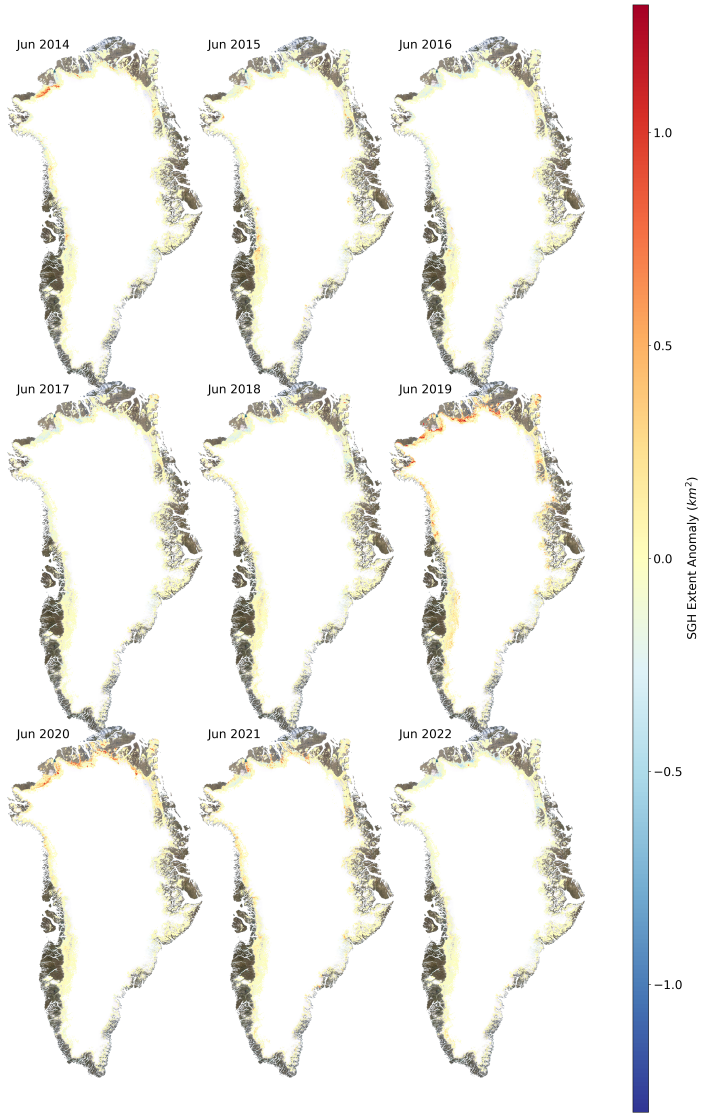


Figure B.3: **(b) The supraglacial hydrology extent anomaly for June compared to the mean extent in 1200 m grid cells on the Greenland Ice Sheet. The anomaly was calculated, taking the difference of the mean cell value of the 9 years in each cell from the extent value for each year.**

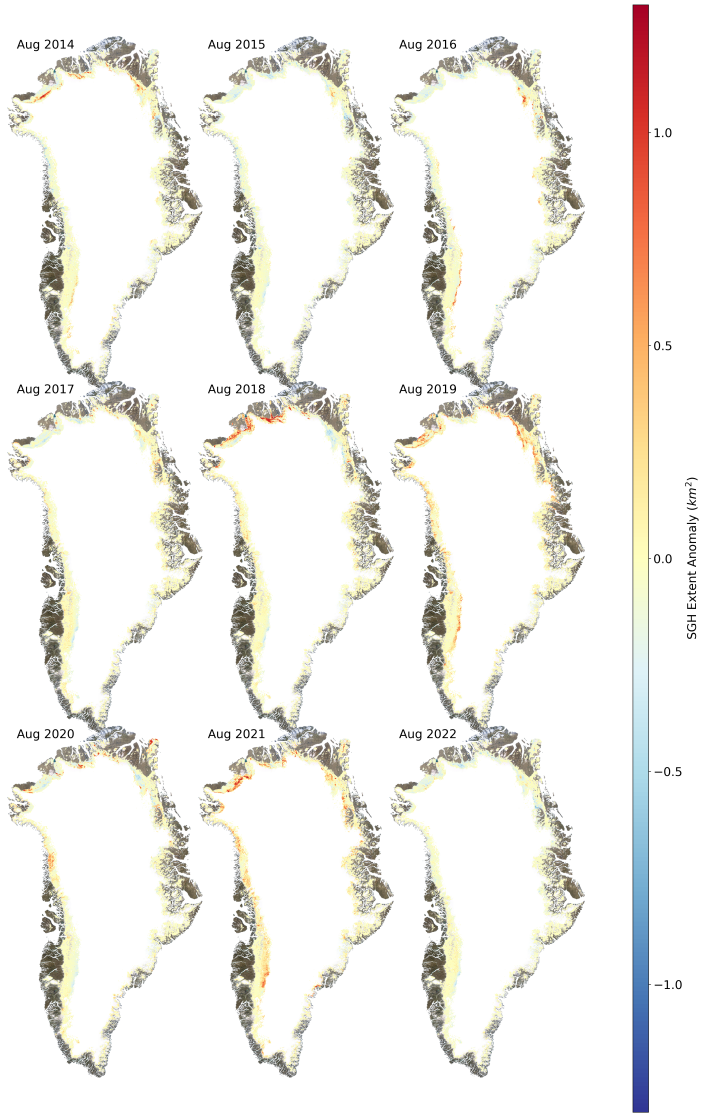


Figure B.3: **(c) The supraglacial hydrology extent anomaly for August compared to the mean extent in 1200 m grid cells on the Greenland Ice Sheet. The anomaly was calculated, taking the difference of the mean cell value of the 9 years in each cell from the extent value for each year.**

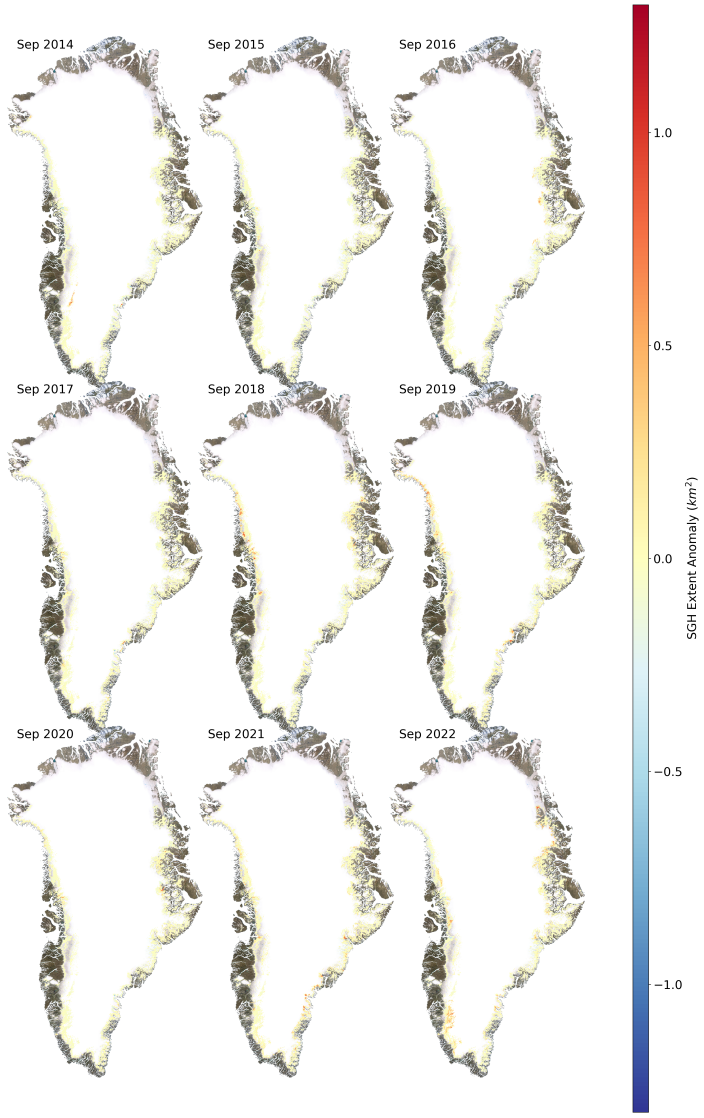
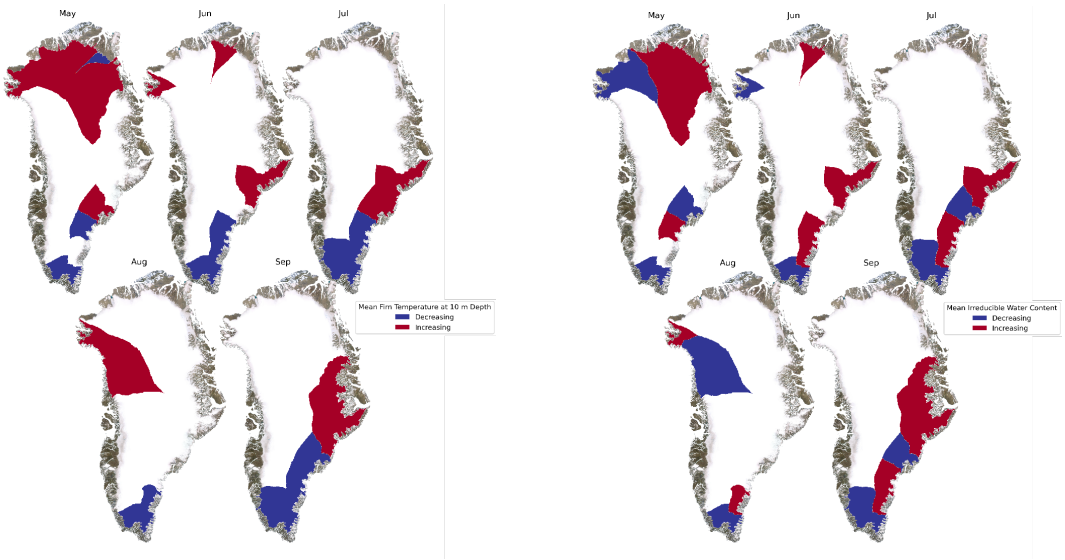
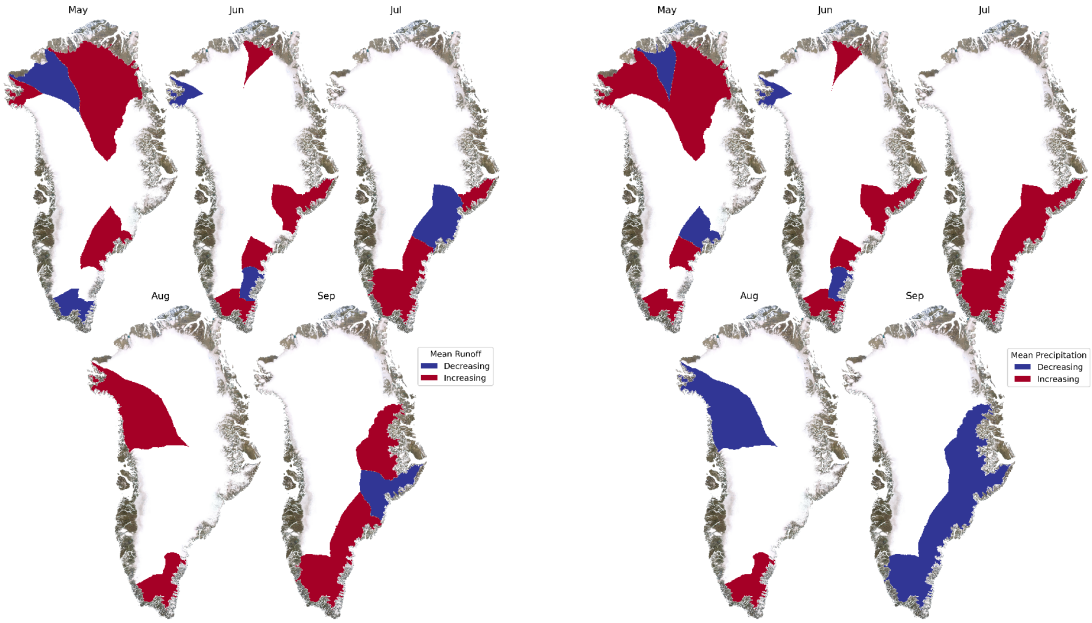


Figure B.3: **(d)The supraglacial hydrology extent anomaly for September compared to the mean extent in 1200 m grid cells on the Greenland Ice Sheet. The anomaly was calculated, taking the difference of the mean cell value of the 9 years in each cell from the extent value for each year.**



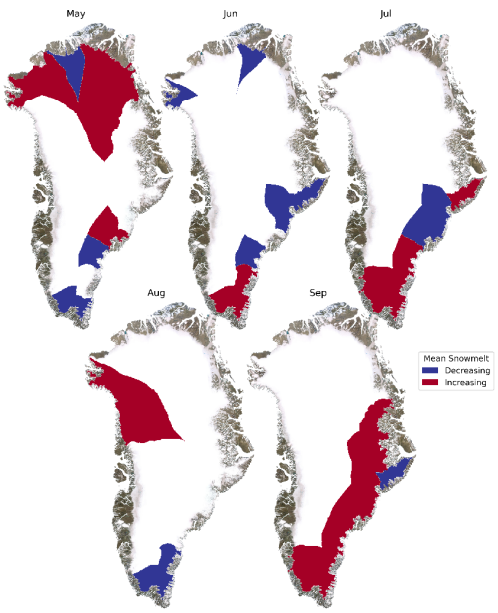
(a) **Trends in the mean firn temperature.** *(b)* **Trends in the irreducible water content.**

Figure B.3: **Trends for the mean firn temperature at 10 m depth and mean irreducible water content for the years 2014-2020 calculated from IMAU-FDM (Firn Densification Model) v1.2G (Brils et al., 2022), where we find a significant trend for the occurrence of SGH (Figure 3.8a) across the years 2014-2022 for each Zwally basin.**

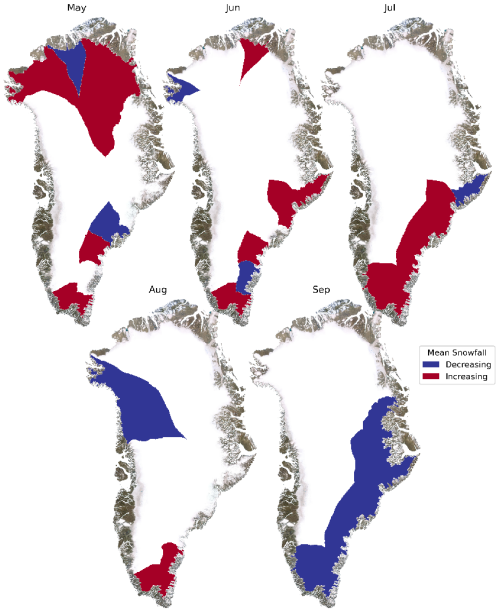


(a) **Trends in mean runoff.** *(b)* **Trends in mean precipitation.**

Figure B.4: **Trends for the mean runoff and mean precipitation for the years 2014-2022 calculated from regional climate model (RACMO2, B. Noël et al., 2018), where we find a significant trend for the occurrence of SGHF (Figure 3.8a) across the years 2014-2022 for each Zwally basin.**

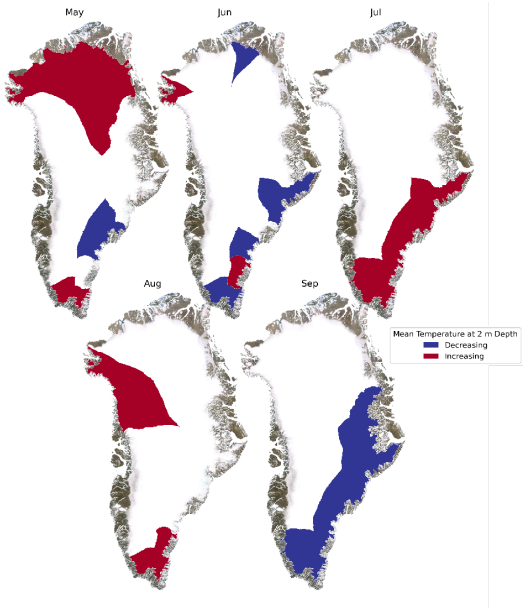


(c) **Trends in the mean snowmelt.**



(d) **Trends in the mean snowfall.**

Figure B.4: **Trends for the mean snowmelt and mean snowfall for the years 2014-2022 calculated from regional climate model (RACMO2, B. Noël et al., 2018), where we find a significant trend for the occurrence of SGHF (Figure 3.8a) across the years 2014-2022 for each Zwally basin.**



(e) **Trends in the mean 2 m temperature.**

Figure B.4: **Trends for the mean temperature of air at 2m above the ice surface for the years 2014-2022 calculated from regional climate model (RACMO2, B. Noël et al., 2018), where we find a significant trend for the occurrence of SGHF (Figure 3.8a) across the years 2014-2022 for each Zwally basin.**

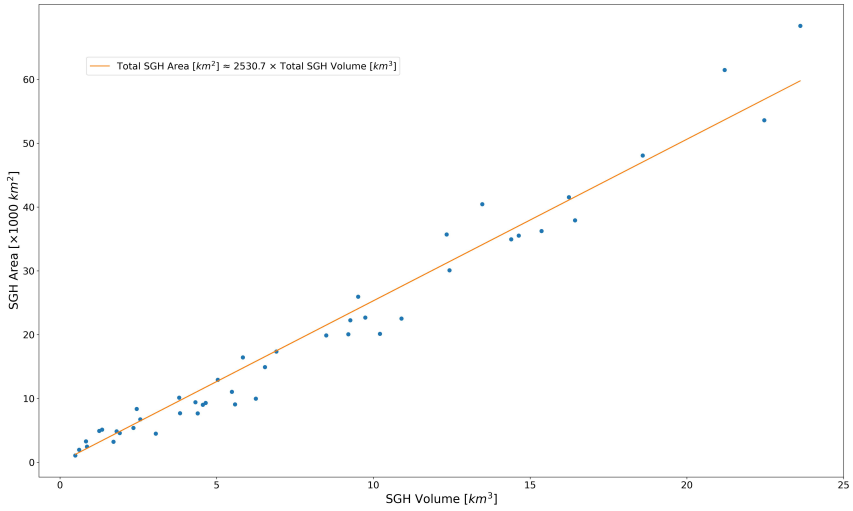


Figure B.5: **Relationship between total monthly SGHF area (in km^2) and volume (in km^3) on the GrIS is found to be linear with the equation for the line of best fit: SGHF Area (km^2) = 2530.7 $\times$ SGHF Volume (km^3).**

Appendix C

A new method for probabilistic prediction of supraglacial lakes on the southwest Greenland Ice Sheet

C.1 Training & testing data sets

Table C.1: **The geographic coordinates, dates, data sets, and border types for the Seninel-2 training data set.**

Latitude	Longitude	Date	Border Type	Data Set
67.546 N	49.533 W	10 06 2019	Defined Border (1a)	Training
66.811 N	48.539 W	03 07 2019	Defined Border (1b)	Training
67.529 N	49.323 W	29 07 2021	Defined Border (1c)	Training
67.546 N	49.533 W	30 06 2022	Defined Border (1d)	Training
67.413 N	49.204 W	24 07 2022	Defined Border (1e)	Training
66.772 N	48.683 W	28 07 2017	Defined Border (1f)	Training
67.413 N	49.204 W	10 07 2018	Defined Border (1g)	Training
67.546 N	49.533 W	05 06 2019	Defined Border (1h)	Training
67.413 N	49.204 W	18 06 2019	Defined Border (1i)	Training
67.546 N	49.533 W	13 06 2017	Blurred Border (2a)	Training
66.83 N	48.776 W	29 07 2021	Blurred Border (2b)	Training
66.811 N	48.539 W	24 07 2022	Blurred Border (2c)	Training
66.772 N	48.683 W	01 08 2017	Blurred Border (2d)	Training
67.529 N	49.323 W	11 06 2019	Blurred Border (2e)	Training
67.546 N	49.533 W	12 06 2020	Blurred Border (2f)	Training
67.413 N	49.204 W	10 07 2021	Blurred Border (2g)	Training
67.529 N	49.323 W	10 07 2021	Blurred Border (2h)	Training
66.83 N	48.776 W	19 07 2022	Blurred Border (2i)	Training
67.546 N	49.533 W	26 06 2017	NoClear Border (3a)	Training
66.83 N	48.776 W	12 06 2018	NoClear Border (3b)	Training
66.772 N	48.683 W	20 06 2019	NoClear Border (3c)	Training
67.546 N	49.533 W	20 06 2019	NoClear Border (3d)	Training
66.83 N	48.776 W	12 07 2021	NoClear Border (3e)	Training
66.83 N	48.776 W	01 08 2021	NoClear Border (3f)	Training
66.83 N	48.776 W	24 07 2022	NoClear Border (3g)	Training
67.236 N	48.737 W	14 08 2017	NoClear Border (3h)	Training
67.413 N	49.204 W	30 07 2021	NoClear Border (3i)	Training

Table C.2: **The geographic coordinates, dates, and data sets for the Seninel-2 testing data set.**

Latitude	Longitude	Date	Data Set
67.546 N	49.533 W	03 06 2017	Testing (1a)
67.529 N	49.323 W	26 06 2017	Testing (1b)
67.236 N	48.737 W	12 08 2017	Testing (1c)
66.83 N	48.776 W	02 07 2021	Testing (1d)
66.811 N	48.539 W	02 07 2021	Testing (1e)
67.529 N	49.323 W	14 07 2022	Testing (1f)
67.413 N	49.204 W	17 07 2022	Testing (1g)
67.546 N	49.533 W	18 06 2019	Testing (1h)
66.772 N	48.683 W	30 07 2017	Testing (1i)
67.546 N	49.533 W	02 07 2021	Testing (1j)
67.413 N	49.204 W	02 07 2020	Testing (1k)
67.546 N	49.533 W	10 07 2021	Testing (1l)

References

Ahmadzadeh, A. et al. (2017). "Improving the functionality of tamura directionality on solar images". In: *2017 IEEE International Conference on Big Data (Big Data)*, pp. 2518–2526.

Ai, B. et al. (2020). "Convolutional Neural Network to Retrieve Water Depth in Marine Shallow Water Area From Remote Sensing Images". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13, pp. 2888–2898. ISSN: 2151-1535.

Al Balasmeh, O. I., T. Karmaker, and R. Babbar (2020). "Estimation of Reservoir Capacity and Sediment Deposition Using Remote Sensing Data". EN. In: pp. 154–161.

Alley, K. E. et al. (2018). "Quantifying vulnerability of Antarctic ice shelves to hydrofracture using microwave scattering properties". en. In: *Remote Sensing of Environment* 210, pp. 297–306. ISSN: 0034-4257.

Alley, R. B. et al. (2005). "Access of surface meltwater to beds of sub-freezing glaciers: preliminary insights". en. In: *Annals of Glaciology* 40, pp. 8–14. ISSN: 0260-3055, 1727-5644.

Amini, S. et al. (2022). "Urban Land Use and Land Cover Change Analysis Using Random Forest Classification of Landsat Time Series". en. In: *Remote Sensing* 14.11, p. 2654. ISSN: 2072-4292.

Antarctic weather (2019). en.

Arthur, D. and S. Vassilvitskii (2007). "k-means++: The Advantages of Careful Seeding". en. In: p. 9.

Arthur, J. F. et al. (2020a). "Recent understanding of Antarctic supraglacial lakes using satellite remote sensing". en. In: *Progress in Physical Geography: Earth and Environment* 44.6, pp. 837–869. ISSN: 0309-1333.

Arthur, J. F. et al. (2020b). "Distribution and seasonal evolution of supraglacial lakes on Shackleton Ice Shelf, East Antarctica". English. In: *The Cryosphere* 14.11, pp. 4103–4120. ISSN: 1994-0416.

Axelsson, A. et al. (2021). "Tree species classification using Sentinel-2 imagery and Bayesian inference". In: *International Journal of Applied Earth Observation and Geoinformation* 100, p. 102318. ISSN: 1569-8432.

Bamber, J. L. et al. (2018). "The land ice contribution to sea level during the satellite era". en. In: *Environmental Research Letters* 13.6, p. 063008. ISSN: 1748-9326.

Banda, J. M. and R. A. Angryk (2010). "Selection of Image Parameters as the First Step towards Creating a CBIR System for the Solar Dynamics Observatory". In: *2010 International Conference on Digital Image Computing: Techniques and Applications*, pp. 528–534.

Banwell, A. F., D. R. MacAyeal, and O. V. Sergienko (2013). "Breakup of the Larsen B Ice Shelf triggered by chain reaction drainage of supraglacial lakes". en. In: *Geophysical Research Letters* 40.22, pp. 5872–5876. ISSN: 1944-8007.

Banwell, A. F. et al. (2014). "Supraglacial lakes on the Larsen B ice shelf, Antarctica, and at Paakitsoq, West Greenland: a comparative study". en. In: *Annals of Glaciology* 55.66, pp. 1–8. ISSN: 0260-3055, 1727-5644.

Banwell, A. F. et al. (2019). "Direct measurements of ice-shelf flexure caused by surface meltwater ponding and drainage". en. In: *Nature Communications* 10.1, p. 730. ISSN: 2041-1723.

Barnes, T. J. et al. (2021). “Changes in Supraglacial Lakes on George VI Ice Shelf, Antarctic Peninsula: 1973–2020”. English. In: *The Cryosphere Discussions*, pp. 1–19.

Barral, H. et al. (2014). “Blowing snow in coastal Adélie Land, Antarctica: three atmospheric-moisture issues”. English. In: *The Cryosphere* 8.5, pp. 1905–1919. ISSN: 1994-0416.

Bartholomew, I. et al. (2010). “Seasonal evolution of subglacial drainage and acceleration in a Greenland outlet glacier”. In: *Nature Geoscience* 3.6, pp. 408–411. ISSN: 1752-0894.

Bartholomew, I. et al. (2011). “Supraglacial forcing of subglacial drainage in the ablation zone of the Greenland ice sheet”. In: *Geophysical Research Letters* 38. ISSN: 0094-8276.

Bartholomew, I. et al. (2012). “Short-term variability in Greenland Ice Sheet motion forced by time-varying meltwater drainage: Implications for the relationship between subglacial drainage system behavior and ice velocity”. In: *Journal of Geophysical Research-Earth Surface* 117. ISSN: 0148-0227.

Bazzi, H. et al. (2021). “An Operational Framework for Mapping Irrigated Areas at Plot Scale Using Sentinel-1 and Sentinel-2 Data”. en. In: *Remote Sensing* 13.13, p. 2584. ISSN: 2072-4292.

Belgiu, M. and L. Drăguţ (2016). “Random forest in remote sensing: A review of applications and future directions”. en. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 114, pp. 24–31. ISSN: 0924-2716.

Bell, R. E. et al. (2017). “Antarctic ice shelf potentially stabilized by export of meltwater in surface river”. en. In: *Nature* 544.7650, pp. 344–348. ISSN: 1476-4687.

Bell, R. E. et al. (2018). “Antarctic surface hydrology and impacts on ice-sheet mass balance”. en. In: *Nature Climate Change* 8.12, pp. 1044–1052. ISSN: 1758-6798.

Benedek, C. L. and I. C. Willis (2021). "Winter drainage of surface lakes on the Greenland Ice Sheet from Sentinel-1 SAR imagery". English. In: *The Cryosphere* 15.3, pp. 1587–1606. ISSN: 1994-0416.

Bi, H., Y. Liang, and X. Chen (2023). "Distinct Role of a Spring Atmospheric Circulation Mode in the Arctic Sea Ice Decline in Summer". en. In: *Journal of Geophysical Research: Atmospheres* 128.6, e2022JD037477. ISSN: 2169-8996.

Bintanja, R. et al. (2013). "Important role for ocean warming and increased ice-shelf melt in Antarctic sea-ice expansion". en. In: *Nature Geoscience* 6.5, pp. 376–379. ISSN: 1752-0908.

Boening, C. et al. (2012). "Snowfall-driven mass change on the East Antarctic ice sheet". en. In: *Geophysical Research Letters* 39.21. ISSN: 1944-8007.

Böhning, D. (1992). "Multinomial logistic regression algorithm". en. In: *Annals of the Institute of Statistical Mathematics* 44.1, pp. 197–200. ISSN: 1572-9052.

Bolch, T. et al. (2008). "Identification of glacier motion and potentially dangerous glacial lakes in the Mt. Everest region/Nepal using spaceborne imagery". English. In: *Natural Hazards and Earth System Sciences* 8.6, pp. 1329–1340. ISSN: 1561-8633.

Booth, A. D. et al. (2012). "Thin-layer effects in glaciological seismic amplitude-versus-angle (AVA) analysis: implications for characterising a subglacial till unit, Russell Glacier, West Greenland". English. In: *The Cryosphere* 6.4, pp. 909–922. ISSN: 1994-0416.

Boulton, G. S. and A. S. Jones (1979). "Stability of Temperate Ice Caps and Ice Sheets Resting on Beds of Deformable Sediment". en. In: *Journal of Glaciology* 24.90, pp. 29–43. ISSN: 0022-1430, 1727-5652.

Box, J. E. et al. (2012). "Greenland ice sheet albedo feedback: thermodynamics and atmospheric drivers". English. In: *The Cryosphere* 6.4, pp. 821–839. ISSN: 1994-0416.

Box, J. E. and K. Ski (2007). "Remote sounding of Greenland supraglacial melt lakes: implications for subglacial hydraulics". en. In: *Journal of Glaciology* 53.181, pp. 257–265. ISSN: 0022-1430, 1727-5652.

Box, J. E. and K. Steffen (2001). "Sublimation on the Greenland Ice Sheet from automated weather station observations". en. In: *Journal of Geophysical Research: Atmospheres* 106.D24, pp. 33965–33981. ISSN: 2156-2202.

Box, J. E. et al. (2022). "Greenland ice sheet climate disequilibrium and committed sea-level rise". en. In: *Nature Climate Change* 12.9, pp. 808–813. ISSN: 1758-6798.

Braithwaite, R. J. (1995). "Positive Degree-Day Factors for Ablation on the Greenland Ice-Sheet Studied by Energy-Balance Modeling". English. In: *Journal of Glaciology* 41.137, pp. 153–160. ISSN: 0022-1430.

Braithwaite, R. J. and S. C. B. Raper (2009). "Estimating equilibrium-line altitude (ELA) from glacier inventory data". en. In: *Annals of Glaciology* 50.53, pp. 127–132. ISSN: 0260-3055, 1727-5644.

Breiman, L. (2001). "Random Forests". en. In: *Machine Learning* 45.1, pp. 5–32. ISSN: 1573-0565.

Brils, M. et al. (2022). "Improved representation of the contemporary Greenland ice sheet firn layer by IMAU-FDM v1.2G". English. In: *Geoscientific Model Development* 15.18, pp. 7121–7138. ISSN: 1991-959X.

Brownlee, J. (2019). *A Gentle Introduction to Cross-Entropy for Machine Learning*. en-US.

Buch, E. (2002). "Present oceanographic conditions in Greenland Waters". en. In.

Chander, G., B. L. Markham, and D. L. Helder (2009). "Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors". en. In: *Remote Sensing of Environment* 113.5, pp. 893–903. ISSN: 00344257.

Chandler, D. M. and A. Hubbard (2023). "Widespread partial-depth hydrofractures in ice sheets driven by supraglacial streams". en. In: *Nature Geoscience* 16.7, pp. 605–611. ISSN: 1752-0908.

Charlebois, D., D. Goodenough, and S. Matwin (1993). "Machine learning from remote sensing analysis". In: *Proceedings of IGARSS '93 - IEEE International Geoscience and Remote Sensing Symposium*, 165–172 vol.1.

Chen, Y. (2020). "Two Sets of Simple Formulae to Estimating Fractal Dimension of Irregular Boundaries". In: *Mathematical Problems in Engineering* 2020, pp. 1–15. ISSN: 1024-123X, 1563-5147.

Choi, Y. et al. (2021). "Ice dynamics will remain a primary driver of Greenland ice sheet mass loss over the next century". en. In: *Communications Earth & Environment* 2.1, pp. 1–9. ISSN: 2662-4435.

Christoffersen, P. et al. (2018). "Cascading lake drainage on the Greenland Ice Sheet triggered by tensile shock and fracture". In: *Nature Communications* 9. ISSN: 2041-1723.

Chu, W. et al. (2016). "Extensive winter subglacial water storage beneath the Greenland Ice Sheet". en. In: *Geophysical Research Letters* 43.24, pp. 12, 484–12, 492. ISSN: 1944-8007.

Chudley, T. R. et al. (2019). "Supraglacial lake drainage at a fast-flowing Greenlandic outlet glacier". In: *Proceedings of the National Academy of Sciences* 116.51, pp. 25468–25477.

Chylek, P., J. E. Box, and G. Lesins (2004). "Global Warming and the Greenland Ice Sheet". en. In: *Climatic Change* 63.1, pp. 201–221. ISSN: 1573-1480.

Clarkson, D., E. Eastoe, and A. Leeson (2021). *Melt probabilities and surface temperature trends on the Greenland ice sheet using a Gaussian mixture model*. en. preprint. Ice sheets/Greenland.

Colgan, W. et al. (2011). "An increase in crevasse extent, West Greenland: Hydrologic implications". en. In: *Geophysical Research Letters* 38.18. ISSN: 1944-8007.

Colliander, A. et al. (2023). "Spatial and temporal differences in surface and subsurface meltwater distribution over Greenland ice sheet using multi-frequency passive microwave observations". In: *Remote Sensing of Environment* 295, p. 113705. ISSN: 0034-4257.

Comiso, J. C. and F. Nishio (2008). "Trends in the sea ice cover using enhanced and compatible AMSR-E, SSM/I, and SMMR data". en. In: *Journal of Geophysical Research: Oceans* 113.C2. ISSN: 2156-2202.

Comiso, J. C. et al. (2008). "Accelerated decline in the Arctic sea ice cover". en. In: *Geophysical Research Letters* 35.1. ISSN: 1944-8007.

Condeça, J., J. Nascimento, and N. Barreiras (2022). "Monitoring the Storage Volume of Water Reservoirs Using Google Earth Engine". en. In: *Water Resources Research* 58.3, e2021WR030026. ISSN: 1944-7973.

Cook, A. J. and D. G. Vaughan (2010). "Overview of areal changes of the ice shelves on the Antarctic Peninsula over the past 50 years". English. In: *The Cryosphere* 4.1, pp. 77–98. ISSN: 1994-0416.

Cooper, A. P. R. (1997). "Historical observations of Prince Gustav Ice Shelf". en. In: *Polar Record* 33.187, pp. 285–294. ISSN: 1475-3057, 0032-2474.

Cornish, S. B. et al. (2022). "Rise and fall of sea ice production in the Arctic Ocean's ice factories". en. In: *Nature Communications* 13.1, p. 7800. ISSN: 2041-1723.

Corr, D. (2021). *diarmuidcorr/Lake-Channel-Identifier: S2 and L8 SGL and Channel Classifier.*

— (2023a). "Classifying_SGH_from_Optical_Imagery". In.

— (2023b). "Probabilistic_Prediction_SGL". In.

— (2023c). *Supraglacial hydrology extent and depth estimates for Greenland ice sheet from 2013 through 2022.*

Corr, D. et al. (2021). *Supraglacial lakes and channels in West Antarctica and Antarctic Peninsula during January 2017*. eng.

Corr, D. et al. (2022). "An inventory of supraglacial lakes and channels across the West Antarctic Ice Sheet". en. In: *Earth System Science Data* 14.1, pp. 209–228. ISSN: 1866-3516.

Corr, D. e. a. (n.d.). "Supraglacial hydrology coverage has increased on the Greenland ice sheet over the last decade". In: *Nature Climate Change (Intended)* ().

Covi, F., R. Hock, and C. H. Reijmer (2023). "Challenges in modeling the energy balance and melt in the percolation zone of the Greenland ice sheet". en. In: *Journal of Glaciology* 69.273, pp. 164–178. ISSN: 0022-1430, 1727-5652.

Culberg, R., D. M. Schroeder, and W. Chu (2021). "Extreme melt season ice layers reduce firn permeability across Greenland". en. In: *Nature Communications* 12.1, p. 2336. ISSN: 2041-1723.

Cullen, N. J. et al. (2014). "Assessing the role of sublimation in the dry snow zone of the Greenland ice sheet in a warming world". en. In: *Journal of Geophysical Research: Atmospheres* 119.11, pp. 6563–6577. ISSN: 2169-8996.

Curry, J. A., J. L. Schramm, and E. E. Ebert (1995). "Sea Ice-Albedo Climate Feedback Mechanism". EN. In: *Journal of Climate* 8.2, pp. 240–247. ISSN: 0894-8755, 1520-0442.

Das, S. B. et al. (2008). "Fracture Propagation to the Base of the Greenland Ice Sheet During Supraglacial Lake Drainage". In: *Science* 320.5877, pp. 778–781.

Datta, R. T. and B. Wouters (2021). "Supraglacial lake bathymetry automatically derived from ICESat-2 constraining lake depth estimates from multi-source satellite imagery". English. In: *The Cryosphere* 15.11, pp. 5115–5132. ISSN: 1994-0416.

Datta, R. T. et al. (2019). "The Effect of Foehn-Induced Surface Melt on Firn Evolution Over the Northeast Antarctic Peninsula". en. In: *Geophysical Research Letters* 46.7, pp. 3822–3831. ISSN: 1944-8007.

David Knighton, A. (1972). "Meandering Habit of Supraglacial Streams". In: *GSA Bulletin* 83.1, pp. 201–204. ISSN: 0016-7606.

Davison, B. J. et al. (2019). "The Influence of Hydrology on the Dynamics of Land-Terminating Sectors of the Greenland Ice Sheet". In: *Frontiers in Earth Science* 7. ISSN: 2296-6463.

Dawson, E. J. et al. (2022). "Ice mass loss sensitivity to the Antarctic ice sheet basal thermal state". en. In: *Nature Communications* 13.1, p. 4957. ISSN: 2041-1723.

De Angelis, H. and P. Skvarca (2003). "Glacier Surge After Ice Shelf Collapse". In: *Science* 299.5612, pp. 1560–1562.

De la Peña, S. et al. (2015). "Changes in the firn structure of the western Greenland Ice Sheet caused by recent warming". English. In: *The Cryosphere* 9.3, pp. 1203–1211. ISSN: 1994-0416.

Dell, R. L. et al. (2022). "Supervised classification of slush and ponded water on Antarctic ice shelves using Landsat 8 imagery". en. In: *Journal of Glaciology* 68.268, pp. 401–414. ISSN: 0022-1430, 1727-5652.

Di, J. et al. (2021). "Automatic Extraction of Supraglacial Lake using SAR Imagery and Deep Learning". In: *2021 CIE International Conference on Radar (Radar)*, pp. 2496–2499.

Dirscherl, M. et al. (2020). "Automated Mapping of Antarctic Supraglacial Lakes Using a Machine Learning Approach". en. In: *Remote Sensing* 12.7, p. 1203. ISSN: 2072-4292.

— (2021). "A Novel Method for Automated Supraglacial Lake Mapping in Antarctica Using Sentinel-1 SAR Imagery and Deep Learning". en. In: *Remote Sensing* 13.2, p. 197. ISSN: 2072-4292.

Dirscherl, M. C., A. J. Dietz, and C. Kuenzer (2021). "Seasonal evolution of Antarctic supraglacial lakes in 2015–2021 and links to environmental controls". English. In: *The Cryosphere* 15.11, pp. 5205–5226. ISSN: 1994-0416.

Doyle, S. H. et al. (2013). "Ice tectonic deformation during the rapid in situ drainage of a supraglacial lake on the Greenland Ice Sheet". In: *Cryosphere* 7.1, pp. 129–140. ISSN: 1994-0416.

Dunmire, D. et al. (2020). "Observations of Buried Lake Drainage on the Antarctic Ice Sheet". en. In: *Geophysical Research Letters* 47.15, e2020GL087970. ISSN: 1944-8007.

Dunmire, D. et al. (2021). "Contrasting regional variability of buried meltwater extent over 2 years across the Greenland Ice Sheet". English. In: *The Cryosphere* 15.6, pp. 2983–3005. ISSN: 1994-0416.

Echelmeyer, K., T. S. Clarke, and W. D. Harrison (1991). "Surficial glaciology of Jakobshavns Isbræ, West Greenland: Part I. Surface morphology". en. In: *Journal of Glaciology* 37.127, pp. 368–382. ISSN: 0022-1430, 1727-5652.

Eisen, O. et al. (2020). "Basal roughness of the East Antarctic Ice Sheet in relation to flow speed and basal thermal state". en. In: *Annals of Glaciology* 61.81, pp. 162–175. ISSN: 0260-3055, 1727-5644.

Ely, J. C. et al. (2017). "Insights on the formation of longitudinal surface structures on ice sheets from analysis of their spacing, spatial distribution, and relationship to ice thickness and flow". In: *Journal of Geophysical Research-Earth Surface* 122.4, pp. 961–972. ISSN: 2169-9003.

Enderlin, E. M. et al. (2014). "An improved mass budget for the Greenland ice sheet". In: *Geophysical Research Letters* 41.3, pp. 866–872. ISSN: 0094-8276.

Engelhardt, H. (2004). "Thermal regime and dynamics of the West Antarctic ice sheet". en. In: *Annals of Glaciology* 39, pp. 85–92. ISSN: 0260-3055, 1727-5644.

Ettema, J. et al. (2009). "Higher surface mass balance of the Greenland ice sheet revealed by high-resolution climate modeling". en. In: *Geophysical Research Letters* 36.12. ISSN: 1944-8007.

Fahnestock, M. et al. (2001). "High Geothermal Heat Flow, Basal Melt, and the Origin of Rapid Ice Flow in Central Greenland". In: *Science* 294.5550, pp. 2338–2342.

Fair, Z. et al. (2020). "Using ICESat-2 and Operation IceBridge altimetry for supraglacial lake depth retrievals". English. In: *The Cryosphere* 14.11, pp. 4253–4263. ISSN: 1994-0416.

Feng, D. (2009). "Study on information extraction of water body with a new water index(NWI)". en. In: *undefined*.

Feyisa, G. L. et al. (2014). "Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery". In: *Remote Sensing of Environment* 140, pp. 23–35. ISSN: 0034-4257.

Fitzpatrick, A. A. W. et al. (2014). "A decade (2002-2012) of supraglacial lake volume estimates across Russell Glacier, West Greenland". In: *Cryosphere* 8.1, pp. 107–121. ISSN: 1994-0416.

Flament, T., E. Berthier, and F. Rémy (2014). "Cascading water underneath Wilkes Land, East Antarctic ice sheet, observed using altimetry and digital elevation models". English. In: *The Cryosphere* 8.2, pp. 673–687. ISSN: 1994-0416.

Foley, K. M. et al. (2013). *Coastal-change and glaciological map of the Amery Ice Shelf area, Antarctica: 1961–2004*. USGS Numbered Series 2600-Q. Reston, VA: U.S. Geological Survey.

Forster, R. R. et al. (2014). "Extensive liquid meltwater storage in firn within the Greenland ice sheet". In: *Nature Geoscience* 7.2, pp. 95–98. ISSN: 1752-0894.

Fowler, A. C. (1987). "Sliding with Cavity Formation". en. In: *Journal of Glaciology* 33.115, pp. 255–267. ISSN: 0022-1430, 1727-5652.

Fox-Kemper, B. et al. (2021). *2021: Ocean, Cryosphere and Sea Level Change. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]* Tech. rep. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1211–1362.

Fretwell, P. et al. (2013). "Bedmap2: improved ice bed, surface and thickness datasets for Antarctica". English. In: *The Cryosphere* 7.1, pp. 375–393. ISSN: 1994-0416.

Fricker, H. A. and T. Scambos (2009). "Connected subglacial lake activity on lower Mercer and Whillans Ice Streams, West Antarctica, 2003–2008". en. In: *Journal of Glaciology* 55.190, pp. 303–315. ISSN: 0022-1430, 1727-5652.

Fricker, H. A. et al. (2002). "Redefinition of the Amery Ice Shelf, East Antarctica, grounding zone". en. In: *Journal of Geophysical Research: Solid Earth* 107.B5, ECV 1–1–ECV 1–9. ISSN: 2156-2202.

Fuerst, J. J. et al. (2016). "The safety band of Antarctic ice shelves". In: *Nature Climate Change* 6.5, pp. 479–482. ISSN: 1758-678X.

Fyke, J. et al. (2018). "An Overview of Interactions and Feedbacks Between Ice Sheets and the Earth System". en. In: *Reviews of Geophysics* 56.2, pp. 361–408. ISSN: 1944-9208.

Gantayat, P. et al. (2023). "A new model for supraglacial hydrology evolution and drainage for the Greenland ice sheet (SHED v1.0)". English. In: *Geoscientific Model Development Discussions*, pp. 1–30.

Georgiou, S. et al. (2009). "Seasonal evolution of supraglacial lake volume from ASTER imagery". en. In: *Annals of Glaciology* 50.52, pp. 95–100. ISSN: 0260-3055, 1727-5644.

Ghorbanzadeh, O., K. Gholamnia, and P. Ghamisi (2022). "The application of ResU-net and OBIA for landslide detection from multi-temporal Sentinel-2 images". In: *Big Earth Data* 0.0, pp. 1–25. ISSN: 2096-4471.

Gilbert, A. et al. (2012). "The influence of snow cover thickness on the thermal regime of Tête Rousse Glacier (Mont Blanc range, 3200 m a.s.l.): Consequences for outburst flood hazards and glacier response to climate change". en. In: *Journal of Geophysical Research: Earth Surface* 117.F4. ISSN: 2156-2202.

Glasser, N. F. and G. H. Gudmundsson (2012). "Longitudinal surface structures (flowstripes) on Antarctic glaciers". In: *Cryosphere* 6.2, pp. 383–391. ISSN: 1994-0416.

Glasser, N. F. and T. A. Scambos (2008). "A structural glaciological analysis of the 2002 Larsen B ice-shelf collapse". en. In: *Journal of Glaciology* 54.184, pp. 3–16. ISSN: 0022-1430, 1727-5652.

Gleason, C. J. et al. (2016). "Characterizing supraglacial meltwater channel hydraulics on the Greenland Ice Sheet from in situ observations". In: *Earth Surface Processes and Landforms* 41.14, pp. 2111–2122. ISSN: 0197-9337.

Gleason, C. J. et al. (2020). *Hourly surface meltwater routing for a Greenlandic supraglacialcatchment across hillslopes and through a dense topological channelnetwork*. en. preprint. Ice sheets/Glacier Hydrology.

Gledhill, L. A. and A. G. Williamson (2018). "Inland advance of supraglacial lakes in north-west Greenland under recent climatic warming". en. In: *Annals of Glaciology* 59.76pt1, pp. 66–82. ISSN: 0260-3055, 1727-5644.

Gloersen, P. (1992). *Arctic and Antarctic sea ice, 1978-1987: Satellite passive-microwave observations and analysis*. 511. Scientific, Technical Information Program, National Aeronautics, and Space ...

Goelzer, H. et al. (2020). "The future sea-level contribution of the Greenland ice sheet: a multi-model ensemble study of ISMIP6". English. In: *The Cryosphere* 14.9, pp. 3071–3096. ISSN: 1994-0416.

Gómez-Rubio, V. (2020). *Bayesian inference with INLA*. Chapman & Hall/CRC Press. Boca Raton, FL.

Gooch, B. T., D. A. Young, and D. D. Blankenship (2016). "Potential groundwater and heterogeneous heat source contributions to ice sheet dynamics in critical submarine basins of East Antarctica". en. In: *Geochemistry, Geophysics, Geosystems* 17.2, pp. 395–409. ISSN: 1525-2027.

Good, I. J. (1952). "Rational Decisions". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 14.1, pp. 107–114. ISSN: 0035-9246.

Gou, R., C. Pennelly, and P. G. Myers (2022). "The Changing Behavior of the West Greenland Current System in a Very High-Resolution Model". en. In: *Journal of Geophysical Research: Oceans* 127.8, e2022JC018404. ISSN: 2169-9291.

Gray, L. et al. (2016). *Improved processing and calibration of the interferometric mode ofthe CryoSat radar altimeter allows height measurements ofsupraglacial lakes in west Greenland*. en. preprint. Remote Sensing.

Grazioli, J. et al. (2017). "Katabatic winds diminish precipitation contribution to the Antarctic ice mass balance". In: *Proceedings of the National Academy of Sciences* 114.41, pp. 10858–10863.

Greuell, W., C. H. Reijmer, and J. Oerlemans (2002). "Narrowband-to-broadband albedo conversion for glacier ice and snow based on aircraft and near-surface measurements". en. In: *Remote Sensing of Environment* 82.1, pp. 48–63. ISSN: 0034-4257.

Gudmundsson, G. H. (2003). "Transmission of basal variability to a glacier surface". In: *Journal of Geophysical Research-Solid Earth* 108.B5. ISSN: 2169-9313.

Gudmundsson, G. H. (2013). "Ice-shelf buttressing and the stability of marine ice sheets". English. In: *The Cryosphere* 7.2, pp. 647–655. ISSN: 1994-0416.

Hagen, J., O. Korsen, and G. Vatne (1991). "Drainage pattern in a subpolar glacier: Brøggerbreen, Svalbard". In: *Arctic Hydrology, Present and Future Tasks*, pp. 121–131.

Hagolle, Olivier (2014). *LANDSAT-Download: Python package for download of Landsat satellite imagery.*

— (2015). *Sentinel-Download: Python package for download of Sentinel-2 satellite imagery.*

Halberstadt, A. R. W. et al. (2020). "Antarctic Supraglacial Lake Identification Using Landsat-8 Image Classification". en. In: *Remote Sensing* 12.8, p. 1327.

Hall, D. K., G. A. Riggs, and V. V. Salomonson (1995). "Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data". en. In: *Remote Sensing of Environment* 54.2, pp. 127–140. ISSN: 0034-4257.

Hambrey, M. J. (1977). "Supraglacial drainage and its relationship to structure, with particular reference to Charles Rabots Bre, Okstindan, Norway". In: *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography* 31.2, pp. 69–77. ISSN: 0029-1951.

Han, H. et al. (2016). "Retrieval of Melt Ponds on Arctic Multiyear Sea Ice in Summer from TerraSAR-X Dual-Polarization Data Using Machine Learning Approaches: A Case Study in the Chukchi Sea with Mid-Incidence Angle Data". en. In: *Remote Sensing* 8.1, p. 57. ISSN: 2072-4292.

Hanna, E. et al. (2013). "Ice-sheet mass balance and climate change". en. In: *Nature* 498.7452, pp. 51–59. ISSN: 1476-4687.

Hanna, E. et al. (2021). "Greenland surface air temperature changes from 1981 to 2019 and implications for ice-sheet melt and mass-balance change". en. In: *International Journal of Climatology* 41.S1, E1336–E1352. ISSN: 1097-0088.

Hao, G. et al. (2021). "Rapid decrease in Antarctic sea ice in recent years". en. In: *Acta Oceanologica Sinica* 40.7, pp. 119–128. ISSN: 1869-1099.

Harper, J. et al. (2012). "Greenland ice-sheet contribution to sea-level rise buffered by meltwater storage in firn". en. In: *Nature* 491.7423, pp. 240–243. ISSN: 1476-4687.

Harper, J. T. et al. (2017). "Borehole measurements indicate hard bed conditions, Kangerlussuaq sector, western Greenland Ice Sheet". en. In: *Journal of Geophysical Research: Earth Surface* 122.9, pp. 1605–1618. ISSN: 2169-9011.

Håvik, L. et al. (2017). "Evolution of the East Greenland Current from Fram Strait to Denmark Strait: Synoptic measurements from summer 2012". en. In: *Journal of Geophysical Research: Oceans* 122.3, pp. 1974–1994. ISSN: 2169-9291.

Hellmer, H. H. (2004). "Impact of Antarctic ice shelf basal melting on sea ice and deep ocean properties". en. In: *Geophysical Research Letters* 31.10. ISSN: 1944-8007.

Herbert, C., A. Camps, and M. Vall-Llossera (2021). "Probabilistic Inference Method to Discriminate Closed Water from Sea Ice Using SENTINEL-1 Sar Signatures". In: *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pp. 5546–5549.

Hochreuther, P. et al. (2021). "Fully Automated Detection of Supraglacial Lake Area for Northeast Greenland Using Sentinel-2 Time-Series". en. In: *Remote Sensing* 13.2, p. 205. ISSN: 2072-4292.

Hoffman, M. J. et al. (2011). "Links between acceleration, melting, and supraglacial lake drainage of the western Greenland Ice Sheet". In: *Journal of Geophysical Research-Earth Surface* 116. ISSN: 2169-9003.

Hoffman, M. J. et al. (2018). "Widespread Moulin Formation During Supraglacial Lake Drainages in Greenland". en. In: *Geophysical Research Letters* 45.2, pp. 778–788. ISSN: 1944-8007.

Holdsworth, G. (1969). "Primary Transverse Crevasses". en. In: *Journal of Glaciology* 8.52, pp. 107–129. ISSN: 0022-1430, 1727-5652.

Holland, D. M. et al. (2008). "Acceleration of Jakobshavn Isbræ triggered by warm subsurface ocean waters". en. In: *Nature Geoscience* 1.10, pp. 659–664. ISSN: 1752-0908.

How, P. et al. (2021). "Greenland-wide inventory of ice marginal lakes using a multi-method approach". en. In: *Scientific Reports* 11.1, p. 4481. ISSN: 2045-2322.

How cold is the Antarctic? (2017). en.

Howat, I. M. et al. (2013). ""Expansion of meltwater lakes on the Greenland Ice Sheet"". In: *Cryosphere* 7.1, pp. 201–204. ISSN: 1994-0416.

Hu, J. et al. (2022). "Distribution and Evolution of Supraglacial Lakes in Greenland during the 2016–2018 Melt Seasons". en. In: *Remote Sensing* 14.1, p. 55. ISSN: 2072-4292.

Huang, X., J. R. Jensen, and J. Mackey (1995). *A machine learning approach to automated construction of knowledge bases for expert systems for remote sensing image analysis with GIS data*. English. Tech. rep. WSRC-MS-95-0433; CONF-9604133-1. Savannah River Site (SRS), Aiken, SC (United States).

Hubbard, B. et al. (2016). "Massive subsurface ice formed by refreezing of ice-shelf melt ponds". In: *Nature Communications* 7. ISSN: 2041-1723.

Huete, A. R. (1988). "A soil-adjusted vegetation index (SAVI)". en. In: *Remote Sensing of Environment* 25.3, pp. 295–309. ISSN: 0034-4257.

Humbert, A. and M. Braun (2008). "The Wilkins Ice Shelf, Antarctica: break-up along failure zones". en. In: *Journal of Glaciology* 54.188, pp. 943–944. ISSN: 0022-1430, 1727-5652.

Humphrey, N. F., J. T. Harper, and W. T. Pfeffer (2012). "Thermal tracking of meltwater retention in Greenland's accumulation area". en. In: *Journal of Geophysical Research: Earth Surface* 117.F1. ISSN: 2156-2202.

Husman, S. d. R. et al. (2023). "Remote Sensing of Surface Melt on Antarctica: Opportunities and Challenges". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 16, pp. 2462–2480. ISSN: 2151-1535.

Igneczi, A. et al. (2016). "Northeast sector of the Greenland Ice Sheet to undergo the greatest inland expansion of supraglacial lakes during the 21st century". en. In: *Geophysical Research Letters* 43.18, pp. 9729–9738. ISSN: 1944-8007.

Iken, A. and R. A. Bindschadler (1986). "Combined measurements of Subglacial Water Pressure and Surface Velocity of Findelengletscher, Switzerland: Conclusions about Drainage System and Sliding Mechanism". en. In: *Journal of Glaciology* 32.110, pp. 101–119. ISSN: 0022-1430, 1727-5652.

Inc, A. (2010). *Redrawing the Map on Redistricting: A National Study*.

Irvine-Fynn, T. D. L. et al. (2011). "Polythermal Glacier Hydrology: A Review". en. In: *Reviews of Geophysics* 49.4. ISSN: 1944-9208.

Jawak, S. and A. Luis (2014). "A Semiautomatic Extraction of Antarctic Lake Features Using Worldview-2 Imagery". In: *Photogrammetric Engineering and Remote Sensing* 80, pp. 939–952.

Jenkins, A., K. W. Nicholls, and H. F. J. Corr (2010). "Observation and Parameterization of Ablation at the Base of Ronne Ice Shelf, Antarctica". EN. In: *Journal of Physical Oceanography* 40.10, pp. 2298–2312. ISSN: 0022-3670, 1520-0485.

Jenkins, A. et al. (2010). "Observations beneath Pine Island Glacier in West Antarctica and implications for its retreat". en. In: *Nature Geoscience* 3.7, pp. 468–472. ISSN: 1752-0908.

Jiang, D. et al. (2022). "Automatic Supraglacial Lake Extraction in Greenland Using Sentinel-1 SAR Images and Attention-Based U-Net". en. In: *Remote Sensing* 14.19, p. 4998. ISSN: 2072-4292.

Johannessen, O. M. et al. (2005). "Recent Ice-Sheet Growth in the Interior of Greenland". In: *Science* 310.5750, pp. 1013–1016.

Johansson, A. M., P. Jansson, and I. A. Brown (2013). "Spatial and temporal variations in lakes on the Greenland Ice Sheet". In: *Journal of Hydrology* 476, pp. 314–320. ISSN: 0022-1694.

Johansson, A. M. and I. A. Brown (2013). "Adaptive Classification of Supra-Glacial Lakes on the West Greenland Ice Sheet". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6.4, pp. 1998–2007. ISSN: 2151-1535.

Johansson, A. M. and I. A. Brown (2012). "Observations of supra-glacial lakes in west Greenland using winter wide swath Synthetic Aperture Radar". In: *Remote Sensing Letters* 3.6, pp. 531–539. ISSN: 2150-704X.

Joughin, I. et al. (2009). "Basal conditions for Pine Island and Thwaites Glaciers, West Antarctica, determined using satellite and airborne data". en. In: *Journal of Glaciology* 55.190, pp. 245–257. ISSN: 0022-1430, 1727-5652.

Juen, M. et al. (2014). "Impact of varying debris cover thickness on ablation: a case study for Koxkar Glacier in the Tien Shan". English. In: *The Cryosphere* 8.2, pp. 377–386. ISSN: 1994-0416.

Kaita, E. et al. (2022). "Landsat 9 Cross Calibration Under-Fly of Landsat 8: Planning, and Execution". en. In: *Remote Sensing* 14.21, p. 5414. ISSN: 2072-4292.

Kamb, B. (1987). "Glacier surge mechanism based on linked cavity configuration of the basal water conduit system". en. In: *Journal of Geophysical Research: Solid Earth* 92.B9, pp. 9083–9100. ISSN: 2156-2202.

Kamiński, B., M. Jakubczyk, and P. Szufel (2018). "A framework for sensitivity analysis of decision trees". In: *Central European Journal of Operations Research* 26.1, pp. 135–159. ISSN: 1435-246X.

Karlsson, N. B. et al. (2021). "A first constraint on basal melt-water production of the Greenland ice sheet". en. In: *Nature Communications* 12.1, p. 3461. ISSN: 2041-1723.

Karlstrom, L. and K. Yang (2016). "Fluvial supraglacial landscape evolution on the Greenland Ice Sheet". In: *Geophysical Research Letters* 43.6, pp. 2683–2692. ISSN: 0094-8276.

Kashiwase, H. et al. (2017). "Evidence for ice-ocean albedo feedback in the Arctic Ocean shifting to a seasonal ice zone". en. In: *Scientific Reports* 7.1, p. 8170. ISSN: 2045-2322.

Kauth, R. and G. Thomas (1976). "The tasselled cap - A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat". In: *Environmental Science*.

Keshri, A. K., A. Shukla, and R. P. Gupta (2009). "ASTER ratio indices for supraglacial terrain mapping". In: *International Journal of Remote Sensing* 30.2, pp. 519–524. ISSN: 0143-1161.

Khan, S. A. et al. (2014). "Sustained mass loss of the northeast Greenland ice sheet triggered by regional warming". en. In: *Nature Climate Change* 4.4, pp. 292–299. ISSN: 1758-6798.

Khan, S. A. et al. (2022). "Extensive inland thinning and speed-up of Northeast Greenland Ice Stream". en. In: *Nature* 611.7937, pp. 727–732. ISSN: 1476-4687.

Kibria, H. B. and A. Matin (2022). "The severity prediction of the binary and multi-class cardiovascular disease - A machine learning-based fusion approach". en. In: *Computational Biology and Chemistry* 98, p. 107672. ISSN: 1476-9271.

Kingslake, J. et al. (2017). "Widespread movement of meltwater onto and across Antarctic ice shelves". en. In: *Nature* 544.7650, pp. 349–352. ISSN: 1476-4687.

Kittel, C. et al. (2021). "Diverging future surface mass balance between the Antarctic ice shelves and grounded ice sheet". English. In: *The Cryosphere* 15.3, pp. 1215–1236. ISSN: 1994-0416.

Koenig, L. S. et al. (2015). "Wintertime storage of water in buried supraglacial lakes across the Greenland Ice Sheet". English. In: *The Cryosphere* 9.4, pp. 1333–1342. ISSN: 1994-0416.

Koenig, L. S. et al. (2016). "Annual Greenland accumulation rates (2009 – 2012) from airborne snow radar". English. In: *The Cryosphere* 10.4, pp. 1739–1752. ISSN: 1994-0416.

Kostrzewski, A. and Z. Zwolinski (1995). "Hydraulic geometry of a supraglacial stream, Ragnarbreen, Spitsbergen". EN. In: *Quaestiones Geographicae. Zeszyt Specjalny* 04. ISSN: 0137-477X.

Kottek, M. et al. (2006). "World Map of the Köppen-Geiger climate classification updated". en. In: *Meteorologische Zeitschrift* 15.3, pp. 259–263. ISSN: 0941-2948.

Koziol, C. et al. (2017). "Quantifying supraglacial meltwater pathways in the Paakitsoq region, West Greenland". en. In: *Journal of Glaciology* 63.239, pp. 464–476. ISSN: 0022-1430, 1727-5652.

Krainski, E. T. et al. (2019). *Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA*. Chapman & Hall/CRC Press. Boca Raton, FL.

Kuipers Munneke, P. et al. (2014). "Firn air depletion as a precursor of Antarctic ice-shelf collapse". In: *Journal of Glaciology* 60.220, pp. 205–214. ISSN: 0022-1430.

Kulessa, B. et al. (2017). "Seismic evidence for complex sedimentary control of Greenland Ice Sheet flow". In: *Science Advances* 3.8, e1603071.

Kwon, Y. et al. (2020). "Uncertainty quantification using Bayesian neural networks in classification: Application to biomedical image segmentation". In: *Computational Statistics & Data Analysis* 142, p. 106816. ISSN: 0167-9473.

LaBarbera, C. H. and D. R. MacAyeal (2011). "Traveling supraglacial lakes on George VI Ice Shelf, Antarctica". In: *Geophysical Research Letters* 38. ISSN: 0094-8276.

Laffin, M. K. et al. (2021). "Climatology and Evolution of the Antarctic Peninsula Föhn Wind-Induced Melt Regime From 1979–2018". en. In: *Journal of Geophysical Research: Atmospheres* 126.4, e2020JD033682. ISSN: 2169-8996.

Laffin, M. K. et al. (2022). "The role of föhn winds in eastern Antarctic Peninsula rapid ice shelf collapse". English. In: *The Cryosphere* 16.4, pp. 1369–1381. ISSN: 1994-0416.

Laffin, M. K. (2022). "The Roles of föhn and Katabatic Winds in Ice Sheet Surface Melt". English. Ph.D. United States – California: University of California, Irvine.

Lai, C.-Y. et al. (2020). "Vulnerability of Antarctica's ice shelves to meltwater-driven fracture". en. In: *Nature* 584.7822, pp. 574–578. ISSN: 1476-4687.

Lampkin, D. J. and J. VanderBerg (2011). "A preliminary investigation of the influence of basal and surface topography on supraglacial lake distribution near Jakobshavn Isbrae, western Greenland". en. In: *Hydrological Processes* 25.21, pp. 3347–3355. ISSN: 1099-1085.

— (2014). "Supraglacial melt channel networks in the Jakobshavn Isbrae region during the 2007 melt season". In: *Hydrological Processes* 28.25, pp. 6038–6053. ISSN: 0885-6087.

Lampkin, D. J. et al. (2020). "Investigating Controls on the Formation and Distribution of Wintertime Storage of Water in Supraglacial Lakes". In: *Frontiers in Earth Science* 8. ISSN: 2296-6463.

Langley, E. S. et al. (2016). "Seasonal evolution of supraglacial lakes on an East Antarctic outlet glacier". en. In: *Geophysical Research Letters* 43.16, pp. 8563–8571. ISSN: 1944-8007.

Larsen, N. K. et al. (2018). "Instability of the Northeast Greenland Ice Stream over the last 45,000 years". en. In: *Nature Communications* 9.1, p. 1872. ISSN: 2041-1723.

Leeson, A. A. et al. (2012). "Simulating the growth of supraglacial lakes at the western margin of the Greenland ice sheet". English. In: *The Cryosphere* 6.5, pp. 1077–1086. ISSN: 1994-0416.

Leeson, A. A. et al. (2015). "Supraglacial lakes on the Greenland ice sheet advance inland under warming climate". In: *Nature Climate Change* 5.1, pp. 51–55. ISSN: 1758-678X.

Leeson, A. A. et al. (2020). "Evolution of Supraglacial Lakes on the Larsen B Ice Shelf in the Decades Before it Collapsed". en. In: *Geophysical Research Letters* 47.4, e2019GL085591. ISSN: 1944-8007.

Leeson, A. A. et al. (2013). "A comparison of supraglacial lake observations derived from MODIS imagery at the western margin of the Greenland ice sheet". In: *Journal of Glaciology* 59.218, pp. 1179–1188. ISSN: 0022-1430.

Leidman, S. Z. et al. (2021). "The Presence and Widespread Distribution of Dark Sediment in Greenland Ice Sheet Supraglacial Streams Implies Substantial Impact of Microbial Communities on Sediment Deposition and Albedo". en. In: *Geophysical Research Letters* 48.1, 2020GL088444. ISSN: 1944-8007.

Leidman, S. Z. et al. (2023). "Intra-seasonal variability in supraglacial stream sediment on the Greenland Ice Sheet". In: *Frontiers in Earth Science* 11. ISSN: 2296-6463.

Lenaerts, J. T. M. et al. (2012). "A new, high-resolution surface mass balance map of Antarctica (1979-2010) based on regional atmospheric climate modeling". In: *Geophysical Research Letters* 39. ISSN: 0094-8276.

Lenaerts, J. T. M. et al. (2017). "Meltwater produced by wind-albedo interaction stored in an East Antarctic ice shelf". In: *Nature Climate Change* 7.1, pp. 58–+. ISSN: 1758-678X.

Lenaerts, J. T. M. et al. (2015). "Representing Greenland ice sheet freshwater fluxes in climate models". en. In: *Geophysical Research Letters* 42.15, pp. 6373–6381. ISSN: 1944-8007.

Lenaerts, J. T. M. et al. (2016). "Present-day and future Antarctic ice sheet climate and surface mass balance in the Community Earth System Model". en. In: *Climate Dynamics* 47.5, pp. 1367–1381. ISSN: 1432-0894.

Lenaerts, J. T. M. et al. (2019). "Observing and Modeling Ice Sheet Surface Mass Balance". en. In: *Reviews of Geophysics* 57.2, pp. 376–420. ISSN: 1944-9208.

Li, H. et al. (2016). "A general variational framework considering cast shadows for the topographic correction of remote sensing imagery". en. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 117, pp. 161–171. ISSN: 0924-2716.

Li, W., M. F. Goodchild, and R. Church (2013). "An efficient measure of compactness for two-dimensional shapes and its application in regionalization problems". In: *International Journal of Geographical Information Science* 27.6, pp. 1227–1250. ISSN: 1365-8816.

Li, Y. et al. (2022). "Surface meltwater runoff routing through a coupled supraglacial-proglacial drainage system, Inglefield Land, northwest Greenland". en. In: *International Journal of Applied Earth Observation and Geoinformation* 106, p. 102647. ISSN: 1569-8432.

Liang, D. et al. (2021). "Analyzing Antarctic ice sheet snowmelt with dynamic Big Earth Data". In: *International Journal of Digital Earth* 14.1, pp. 88–105. ISSN: 1753-8947.

Liang, L. et al. (2023). "Delayed Antarctic melt season reduces albedo feedback". In: *National Science Review* 10.9, nwad157. ISSN: 2095-5138.

Liang, Y.-L. et al. (2012). "A decadal investigation of supraglacial lakes in West Greenland using a fully automatic detection and tracking algorithm". en. In: *Remote Sensing of Environment* 123, pp. 127–138. ISSN: 0034-4257.

Liljedahl, L. C. et al. (2021). "Rapid and sensitive response of Greenland's groundwater system to ice sheet change". en. In: *Nature Geoscience* 14.10, pp. 751–755. ISSN: 1752-0908.

Lindbäck, K. et al. (2015). "Subglacial water drainage, storage, and piracy beneath the Greenland ice sheet". en. In: *Geophysical Research Letters* 42.18, pp. 7606–7614. ISSN: 1944-8007.

Lindgren, F., H. Rue, and J. Lindström (2011). "An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach". en. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73.4, pp. 423–498. ISSN: 1467-9868.

Liu, H., L. Wang, and K. C. Jezek (2006). "Spatiotemporal variations of snowmelt in Antarctica derived from satellite scanning multichannel microwave radiometer and Special Sensor Microwave Imager data (1978–2004)". en. In: *Journal of Geophysical Research: Earth Surface* 111.F1. ISSN: 2156-2202.

Liuzzo, L. et al. (2020). "Parameterization of a Bayesian Normalized Difference Water Index for Surface Water Detection". en. In: *Geosciences* 10.7, p. 260. ISSN: 2076-3263.

Livingstone, S. J. et al. (2022). "Subglacial lakes and their changing role in a warming climate". en. In: *Nature Reviews Earth & Environment* 3.2, pp. 106–124. ISSN: 2662-138X.

Lliboutry, L. (1968). "General Theory of Subglacial Cavitation and Sliding of Temperate Glaciers". en. In: *Journal of Glaciology* 7.49, pp. 21–58. ISSN: 0022-1430, 1727-5652.

Lucchitta, B. K. and C. E. Rosanova (1998). "Retreat of northern margins of George VI and Wilkins Ice Shelves, Antarctic Peninsula". en. In: *Annals of Glaciology* 27, pp. 41–46. ISSN: 0260-3055, 1727-5644.

Luckman, A. et al. (2014). "Surface melt and ponding on Larsen C Ice Shelf and the impact of föhn winds". en. In: *Antarctic Science* 26.6, pp. 625–635. ISSN: 0954-1020, 1365-2079.

Lüthje, M. et al. (2006). "Modelling the evolution of supraglacial lakes on the West Greenland ice-sheet margin". en. In: *Journal of Glaciology* 52.179, pp. 608–618. ISSN: 0022-1430, 1727-5652.

Lutz, K., Z. Bahrami, and M. Braun (2023). "Supraglacial Lake Evolution over Northeast Greenland Using Deep Learning Methods". en. In: *Remote Sensing* 15.17, p. 4360. ISSN: 2072-4292.

Macdonald, G. J., A. F. Banwell, and D. R. MacAyeal (2018). "Seasonal evolution of supraglacial lakes on a floating ice tongue, Petermann Glacier, Greenland". en. In: *Annals of Glaciology* 59.76pt1, pp. 56–65. ISSN: 0260-3055, 1727-5644.

Macdonald, G. J. et al. (2019). "Formation of pedestalled, relict lakes on the McMurdo Ice Shelf, Antarctica". en. In: *Journal of Glaciology* 65.250, pp. 337–343. ISSN: 0022-1430, 1727-5652.

MacFerrin, M. et al. (2019). "Rapid expansion of Greenland's low- permeability ice slabs". en. In: *Nature* 573.7774, pp. 403–407. ISSN: 1476-4687.

MacGregor, J. A. et al. (2016). "A synthesis of the basal thermal state of the Greenland Ice Sheet". en. In: *Journal of Geophysical Research: Earth Surface* 121.7, pp. 1328–1350. ISSN: 2169-9011.

Machguth, H., A. J. Tedstone, and E. Mattea (2023). "Daily variations in Western Greenland slush limits, 2000–2021". en. In: *Journal of Glaciology* 69.273, pp. 191–203. ISSN: 0022-1430, 1727-5652.

Machguth, H. et al. (2016). "Greenland meltwater storage in firn limited by near-surface ice formation". en. In: *Nature Climate Change* 6.4, pp. 390–393. ISSN: 1758-6798.

Machguth, H. et al. (2018). "Editorial: Melt Water Retention Processes in Snow and Firn on Ice Sheets and Glaciers: Observations and Modeling". In: *Frontiers in Earth Science* 6. ISSN: 2296-6463.

Madsen, M. S. et al. (2022). "The role of an interactive Greenland ice sheet in the coupled climate-ice sheet model EC-Earth-PISM". en. In: *Climate Dynamics* 59.3, pp. 1189–1211. ISSN: 1432-0894.

Maier, N. et al. (2023). "Wintertime Supraglacial Lake Drainage Cascade Triggers Large-Scale Ice Flow Response in Greenland". en. In: *Geophysical Research Letters* 50.4, e2022GL102251. ISSN: 1944-8007.

Mantelli, E., C. Camporeale, and L. Ridolfi (2015). "Supraglacial channel inception: Modeling and processes". en. In: *Water Resources Research* 51.9, pp. 7044–7063. ISSN: 1944-7973.

Maritorena, S., A. Morel, and B. Gentili (1994). "Diffuse reflectance of oceanic shallow waters: Influence of water depth and bottom albedo". en. In: *Limnology and Oceanography* 39.7, pp. 1689–1703. ISSN: 1939-5590.

Marshall, J. and K. Speer (2012). "Closure of the meridional overturning circulation through Southern Ocean upwelling". en. In: *Nature Geoscience* 5.3, pp. 171–180. ISSN: 1752-0908.

Marston, R. A. (1983). "Supraglacial Stream Dynamics on the Juneau Icefield". en. In: *Annals of the Association of American Geographers* 73.4, pp. 597–608. ISSN: 1467-8306.

Martins, T. G. et al. (2013). "Bayesian computing with INLA: New features". In: *Computational Statistics & Data Analysis* 67, pp. 68–83. ISSN: 0167-9473.

Massom, R. A. and S. E. Stammerjohn (2010). "Antarctic sea ice change and variability – Physical and ecological implications". In: *Polar Science*. Antarctic Biology in the 21st Century - Advances in and beyond IPY 4.2, pp. 149–186. ISSN: 1873-9652.

Massom, R. A. et al. (2018). "Antarctic ice shelf disintegration triggered by sea ice loss and ocean swell". en. In: *Nature* 558.7710, pp. 383–389. ISSN: 1476-4687.

Mattingly, K. et al. (2021). "The contribution of föhn winds to northeast Greenland summer melt and their relationship with atmospheric rivers". In: EGU21–8002.

McFeeters, S. K. (1996). "The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features". In: *International Journal of Remote Sensing* 17.7, pp. 1425–1432. ISSN: 0143-1161.

McGrath, D. et al. (2011). "Assessing the summer water budget of a moulin basin in the Sermeq Avannarleq ablation region, Greenland ice sheet". en. In: *Journal of Glaciology* 57.205, pp. 954–964. ISSN: 0022-1430, 1727-5652.

McGrath, D. et al. (2012). "Basal crevasses on the Larsen C Ice Shelf, Antarctica: Implications for meltwater ponding and hydrofracture". en. In: *Geophysical Research Letters* 39.16. ISSN: 1944-8007.

McIlhattan, E. A. et al. (2020). "Satellite observations of snowfall regimes over the Greenland Ice Sheet". English. In: *The Cryosphere* 14.12, pp. 4379–4404. ISSN: 1994-0416.

McMillan, M. et al. (2007). "Seasonal evolution of supra-glacial lakes on the Greenland Ice Sheet". en. In: *Earth and Planetary Science Letters* 262.3, pp. 484–492. ISSN: 0012-821X.

Medina-Lopez, E. (2020). "Machine Learning and the End of Atmospheric Corrections: A Comparison between High-Resolution Sea Surface Salinity in Coastal Areas from Top and Bottom of Atmosphere Sentinel-2 Imagery". en. In: *Remote Sensing* 12.18, p. 2924. ISSN: 2072-4292.

Medley, B. and E. R. Thomas (2019). "Increased snowfall over the Antarctic Ice Sheet mitigated twentieth-century sea-level rise". en. In: *Nature Climate Change* 9.1, pp. 34–39. ISSN: 1758-6798.

Melling, L. et al. (2023). *Evaluation of satellite methods for estimating supraglacial lake depth in southwest Greenland*. en. preprint. Ice sheets/Glacier Hydrology.

Miao, X. et al. (2015). "Object-based detection of Arctic sea ice and melt ponds using high spatial resolution aerial photographs". In: *Cold Regions Science and Technology* 119, pp. 211–222. ISSN: 0165-232X.

Miège, C. et al. (2013). "Southeast Greenland high accumulation rates derived from firn cores and ground-penetrating radar". en. In: *Annals of Glaciology* 54.63, pp. 322–332. ISSN: 0260-3055, 1727-5644.

Miles, B. W. J., C. R. Stokes, and S. S. R. Jamieson (2017). "Simultaneous disintegration of outlet glaciers in Porpoise Bay (Wilkes Land), East Antarctica, driven by sea ice break-up". In: *Cryosphere* 11.1, pp. 427–442. ISSN: 1994-0416.

Miles, K. E. et al. (2017). "Toward Monitoring Surface and Subsurface Lakes on the Greenland Ice Sheet Using Sentinel-1 SAR and Landsat-8 OLI Imagery". In: *Frontiers in Earth Science* 5, p. 58. ISSN: 2296-6463.

Mioduszewski, J. R. et al. (2016). "Atmospheric drivers of Greenland surface melt revealed by self-organizing maps". en. In: *Journal of Geophysical Research: Atmospheres* 121.10, pp. 5095–5114. ISSN: 2169-8996.

Montgomery, L. et al. (2020). "Hydrologic Properties of a Highly Permeable Firn Aquifer in the Wilkins Ice Shelf, Antarctica". en. In: *Geophysical Research Letters* 47.22, e2020GL089552. ISSN: 1944-8007.

Moon, T. A. et al. (2020). "Rapid Reconfiguration of the Greenland Ice Sheet Coastal Margin". en. In: *Journal of Geophysical Research: Earth Surface* 125.11, e2020JF005585. ISSN: 2169-9011.

Morlighem, M. et al. (2017). "BedMachine v3: Complete Bed Topography and Ocean Bathymetry Mapping of Greenland From Multibeam Echo Sounding Combined With Mass Conservation". en. In: *Geophysical Research Letters* 44.21, pp. 11, 051–11, 061. ISSN: 1944-8007.

Morriss, B. F. et al. (2013). "A ten-year record of supraglacial lake evolution and rapid drainage in West Greenland using an automated processing algorithm for multispectral imagery". English. In: *The Cryosphere* 7.6, pp. 1869–1877. ISSN: 1994-0416.

Mouginot, J. and U. O. C. Irvine (2017). *MEaSURES Antarctic Boundaries for IPY 2007-2009 from Satellite Radar, Version 2*.

Mouginot, J., B. Scheuchl, and E. Rignot (2012). "Mapping of Ice Motion in Antarctica Using Synthetic-Aperture Radar Data". en. In: *Remote Sensing* 4.9, pp. 2753–2767. ISSN: 2072-4292.

Moussavi, M. et al. (2020). "Antarctic Supraglacial Lake Detection Using Landsat 8 and Sentinel-2 Imagery: Towards Continental Generation of Lake Volumes". en. In: *Remote Sensing* 12.1, p. 134.

Moussavi, M. S. et al. (2016). "Derivation and validation of supraglacial lake volumes on the Greenland Ice Sheet from high-resolution satellite imagery". en. In: *Remote Sensing of Environment* 183, pp. 294–303. ISSN: 0034-4257.

Moussavi, Mahsa (2019). *Satellite image processing - Automated detection of lakes in Antarctica: This repository contains all the codes and dataset required for processing of data collected by Landsat 8 and Sentinel-2 satellites.*

Münchow, A., K. K. Falkner, and H. Melling (2015). "Baffin Island and West Greenland Current Systems in northern Baffin Bay". In: *Progress in Oceanography*. Oceanography of the Arctic and North Atlantic Basins 132, pp. 305–317. ISSN: 0079-6611.

Niehaus, H. et al. (2023). "Sea Ice Melt Pond Fraction Derived From Sentinel-2 Data: Along the MOSAiC Drift and Arctic-Wide". en. In: *Geophysical Research Letters* 50.5, e2022GL102102. ISSN: 1944-8007.

Noël, B. et al. (2015). "Evaluation of the updated regional climate model RACMO2.3: summer snowfall impact on the Greenland Ice Sheet". English. In: *The Cryosphere* 9.5, pp. 1831–1844. ISSN: 1994-0416.

Noël, B. et al. (2018). "Modelling the climate and surface mass balance of polar ice sheets using RACMO2 – Part 1: Greenland (1958–2016)". English. In: *The Cryosphere* 12.3, pp. 811–831. ISSN: 1994-0416.

Noël, B. P. Y. (2019). *Rapid ablation zone expansion amplifies north Greenland mass loss: modelled (RACMO2) and observed (MODIS) data sets*.

Nolin, A. W. and M. C. Payne (2007). "Classification of glacier zones in western Greenland using albedo and surface roughness from the Multi-angle Imaging SpectroRadiometer (MISR)". en. In: *Remote Sensing of Environment*. Multi-angle Imaging SpectroRadiometer (MISR) Special Issue 107.1, pp. 264–275. ISSN: 0034-4257.

Notz, D. (2009). "The future of ice sheets and sea ice: Between reversible retreat and unstoppable loss". In: *Proceedings of the National Academy of Sciences* 106.49, pp. 20590–20595.

Nunes, Vasco (2016). *fetchLandsatSentinelFromGoogleCloud: Python package to find and download Landsat and Sentinel-2 data from public storage on Google Cloud.*

Onesti, L. J. and E. Hestnes (1989). "Slush-Flow Questionnaire". en. In: *Annals of Glaciology* 13, pp. 226–230. ISSN: 0260-3055, 1727-5644.

Orheim, O. and B. Lucchitta (1990). "Investigating Climate Change by Digital Analysis of Blue Ice Extent on Satellite Images of Antarctica". en. In: *Annals of Glaciology* 14, pp. 211–215. ISSN: 0260-3055, 1727-5644.

Orheim, O. and B. K. Lucchitta (1987). "Snow and Ice Studies By Thematic Mapper and Multispectral Scanner Landsat Images". en. In: *Annals of Glaciology* 9, pp. 109–118. ISSN: 0260-3055, 1727-5644.

Orynbaikyzy, A., U. Gessner, and C. Conrad (2022). "Spatial Transferability of Random Forest Models for Crop Type Classification Using Sentinel-1 and Sentinel-2". en. In: *Remote Sensing* 14.6, p. 1493. ISSN: 2072-4292.

Otosaka, I. N. et al. (2022). "Mass Balance of the Greenland and Antarctic Ice Sheets from 1992 to 2020". English. In: *Earth System Science Data Discussions*, pp. 1–33.

Otto, J., F. A. Holmes, and N. Kirchner (2022). "Supraglacial lake expansion, intensified lake drainage frequency, and first observation of coupled lake drainage, during 1985–2020 at Ryder Glacier, Northern Greenland". In: *Frontiers in Earth Science* 10. ISSN: 2296-6463.

Pal, M. (2005). "Random forest classifier for remote sensing classification". In: *International Journal of Remote Sensing* 26.1, pp. 217–222. ISSN: 0143-1161.

Parizek, B. R. and R. B. Alley (2004). "Implications of increased Greenland surface melt under global-warming scenarios: ice-sheet simulations". In: *Quaternary Science Reviews* 23.9, pp. 1013–1027. ISSN: 0277-3791.

Pattyn, F. (2010). "Antarctic subglacial conditions inferred from a hybrid ice sheet/ice stream model". en. In: *Earth and Planetary Science Letters* 295.3, pp. 451–461. ISSN: 0012-821X.

Payne, M. and A. Nolin (2008). "A MISR-based Surface Roughness Map of the Greenland Ice Sheet". In: *AGU Fall Meeting Abstracts*.

Pedregosa, F. et al. (2018). "Scikit-learn: Machine Learning in Python". In: *arXiv:1201.0490 [cs]*.

Peng, J. et al. (2021). "Machine Learning Techniques for Personalised Medicine Approaches in Immune-Mediated Chronic Inflammatory Diseases: Applications and Challenges". eng. In: *Frontiers in Pharmacology* 12, p. 720694. ISSN: 1663-9812.

Pfeffer, W. T. and N. F. Humphrey (1998). "Formation of ice layers by infiltration and refreezing of meltwater". en. In: *Annals of Glaciology* 26, pp. 83–91. ISSN: 0260-3055, 1727-5644.

Pfeffer, W. T., M. F. Meier, and T. H. Illangasekare (1991). "Retention of Greenland runoff by refreezing: Implications for projected future sea level change". en. In: *Journal of Geophysical Research: Oceans* 96.C12, pp. 22117–22124. ISSN: 2156-2202.

Phillips, H. A. (1998). "Surface meltstreams on the Amery Ice Shelf, East Antarctica". In: *Annals of Glaciology, Vol 27, 1998*. Ed. by W. F. Budd. Vol. 27. Annals of Glaciology, pp. 177–181. ISBN: 0-946417-22-9.

Phillips, T., H. Rajaram, and K. Steffen (2010). "Cryo-hydrologic warming: A potential mechanism for rapid thermal response of ice sheets". In: *Geophysical Research Letters* 37.20. ISSN: 0094-8276.

Phillips, T. et al. (2013). "Evaluation of cryo-hydrologic warming as an explanation for increased ice velocities in the wet snow zone, Sermeq Avannarleq, West Greenland". en. In: *Journal of Geophysical Research: Earth Surface* 118.3, pp. 1241–1256. ISSN: 2169-9011.

Philpot, W. D. (1989). "Bathymetric mapping with passive multispectral imagery". EN. In: *Applied Optics* 28.8, pp. 1569–1578. ISSN: 2155-3165.

Picard, G. and M. Fily (2006). "Surface melting observations in Antarctica by microwave radiometers: Correcting 26-year time series from changes in acquisition hours". In: *Remote Sensing of Environment* 104.3, pp. 325–336. ISSN: 0034-4257.

Pitcher, L. H. and L. C. Smith (2019). "Supraglacial Streams and Rivers". In: *Annual Review of Earth and Planetary Sciences* 47.1, pp. 421–452.

Polsby, D. D. and R. Popper (1991). *The Third Criterion: Compactness as a Procedural Safeguard Against Partisan Gerrymandering*. en. SSRN Scholarly Paper ID 2936284. Rochester, NY: Social Science Research Network.

Pope, A. et al. (2016). "Estimating supraglacial lake depth in West Greenland using Landsat 8 and comparison with other multispectral methods". English. In: *The Cryosphere* 10.1, pp. 15–27. ISSN: 1994-0416.

Pope, A. (2016). “Reproducibly estimating and evaluating supraglacial lake depth with Landsat 8 and other multispectral sensors”. In: *Earth and Space Science* 3.4, pp. 176–188. ISSN: 2333-5084.

Porter, C. et al. (2018). *ArcticDEM*. en.

Porter, C. et al. (2023). *ArcticDEM - Mosaics, Version 4.1*. en.

Pritchard, H. D. et al. (2012). “Antarctic ice-sheet loss driven by basal melting of ice shelves”. en. In: *Nature* 484.7395, pp. 502–505. ISSN: 1476-4687.

Purich, A. and E. W. Doddridge (2023). “Record low Antarctic sea ice coverage indicates a new sea ice state”. en. In: *Communications Earth & Environment* 4.1, pp. 1–9. ISSN: 2662-4435.

Rahmani, S. et al. (2010). “An Adaptive IHS Pan-Sharpening Method”. In: *IEEE Geoscience and Remote Sensing Letters* 7.4, pp. 746–750. ISSN: 1558-0571.

Ramadhan, W., S. Astri Novianty, and S. Casi Setianingsih (2017). “Sentiment analysis using multinomial logistic regression”. In: *2017 International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC)*, pp. 46–49.

Rennermalm, A. K. et al. (2013). “Understanding Greenland ice sheet hydrology using an integrated multi-scale approach”. en. In: *Environmental Research Letters* 8.1, p. 015017. ISSN: 1748-9326.

Reock, E. C. (1961). “A Note: Measuring Compactness as a Requirement of Legislative Apportionment”. In: *Midwest Journal of Political Science* 5.1, pp. 70–74. ISSN: 0026-3397.

Reyes, M. F., P. D’Angelo, and F. Fraundorfer (2022). “SyntCities: A Large Synthetic Remote Sensing Dataset for Disparity Estimation”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 15, pp. 10087–10098. ISSN: 1939-1404, 2151-1535.

Rignot, E., J. Mouginot, and B. Scheuchl (2011). "Ice Flow of the Antarctic Ice Sheet". In: *Science* 333.6048, pp. 1427–1430.

Rignot, E. et al. (2004). "Accelerated ice discharge from the Antarctic Peninsula following the collapse of Larsen B ice shelf". en. In: *Geophysical Research Letters* 31.18. ISSN: 1944-8007.

Rignot, E. et al. (2008a). "Mass balance of the Greenland ice sheet from 1958 to 2007". en. In: *Geophysical Research Letters* 35.20. ISSN: 1944-8007.

Rignot, E. et al. (2013). "Ice-shelf melting around Antarctica." en. In: *Science (New York, N.Y.)* 341.6143, pp. 266–270. ISSN: 0036-8075.

Rignot, E. and U. O. C. Irvine (2017). *MEaSUREs InSAR-Based Antarctica Ice Velocity Map, Version 2*.

Rignot, E. et al. (2008b). "Recent Antarctic ice mass loss from radar interferometry and regional climate modelling". en. In: *Nature Geoscience* 1.2, pp. 106–110. ISSN: 1752-0908.

Rignot, E. et al. (2019). "Four decades of Antarctic Ice Sheet mass balance from 1979–2017". In: *Proceedings of the National Academy of Sciences* 116.4, pp. 1095–1103.

Rintoul, S. R. (2018). "The global influence of localized dynamics in the Southern Ocean". en. In: *Nature* 558.7709, pp. 209–218. ISSN: 1476-4687.

Rippin, D. M., A. Pomfret, and N. King (2015). "High resolution mapping of supra-glacial drainage pathways reveals link between micro-channel drainage density, surface roughness and surface reflectance". In: *Earth Surface Processes and Landforms* 40.10, pp. 1279–1290. ISSN: 0197-9337.

Roberts, S. J. et al. (2008). "The Holocene history of George VI Ice Shelf, Antarctic Peninsula from clast-provenance analysis of epishelf lake sediments". en. In: *Palaeogeography, Palaeoclimatology, Palaeoecology*. Lake systems: sedimentary archives of climate change and tectonics 259.2, pp. 258–283. ISSN: 0031-0182.

Rott, H., P. Skvarca, and T. Nagler (1996). "Rapid Collapse of Northern Larsen Ice Shelf, Antarctica". In: *Science* 271.5250, pp. 788–792.

Rue, H., S. Martino, and N. Chopin (2009). "Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations". en. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 71.2, pp. 319–392. ISSN: 1467-9868.

Sazonau, V. (2012). *Implementation and Evaluation of a Random Forest Machine Learning Algorithm*. en.

Scambos, T. A. et al. (2004). "Glacier acceleration and thinning after ice shelf collapse in the Larsen B embayment, Antarctica". en. In: *Geophysical Research Letters* 31.18. ISSN: 1944-8007.

Scambos, T., C. Hulbe, and M. Fahnestock (2003). "Climate-Induced Ice Shelf Disintegration in the Antarctic Peninsula". en. In: *Antarctic Peninsula Climate Variability: Historical and Paleoenvironmental Perspectives*. American Geophysical Union (AGU), pp. 79–92. ISBN: 978-1-118-66845-0.

Scambos, T. et al. (2009). "Ice shelf disintegration by plate bending and hydro-fracture: Satellite observations and model results of the 2008 Wilkins ice shelf break-ups". en. In: *Earth and Planetary Science Letters* 280.1, pp. 51–60. ISSN: 0012-821X.

Scambos, T. A. et al. (2000). "The link between climate warming and break-up of ice shelves in the Antarctic Peninsula". en. In: *Journal of Glaciology* 46.154, pp. 516–530. ISSN: 0022-1430, 1727-5652.

Schoof, C. (2010). "Ice-sheet acceleration driven by melt supply variability". eng. In: *Nature* 468.7325, pp. 803–806. ISSN: 1476-4687.

Schröder, L. et al. (2020). "Perennial Supraglacial Lakes in Northeast Greenland Observed by Polarimetric SAR". en. In: *Remote Sensing* 12.17, p. 2798. ISSN: 2072-4292.

Schwatke, C., D. Scherer, and D. Dettmering (2019). "Automated Extraction of Consistent Time-Variable Water Surfaces of Lakes and Reservoirs Based on Landsat and Sentinel-2". en. In: *Remote Sensing* 11.9, p. 1010.

Seguinot, J. et al. (2020). "Englacial Warming Indicates Deep Crevassing in Bowdoin Glacier, Greenland". In: *Frontiers in Earth Science* 8. ISSN: 2296-6463.

Selmes, N., T. Murray, and T. D. James (2011). "Fast draining lakes on the Greenland Ice Sheet". In: *Geophysical Research Letters* 38.15. ISSN: 0094-8276.

Sergienko, O. V. (2013). "Normal modes of a coupled ice-shelf/sub-ice-shelf cavity system". en. In: *Journal of Glaciology* 59.213, pp. 76–80. ISSN: 0022-1430, 1727-5652.

Sergienko, O. V. and C. L. Hulbe (2011). "'Sticky spots' and subglacial lakes under ice streams of the Siple Coast, Antarctica". en. In: *Annals of Glaciology* 52.58, pp. 18–22. ISSN: 0260-3055, 1727-5644.

Sharifi, A., S. Felegari, and A. Tariq (2022). "Mangrove forests mapping using Sentinel-1 and Sentinel-2 satellite images". en. In: *Arabian Journal of Geosciences* 15.20, p. 1593. ISSN: 1866-7538.

Shean, D. E. et al. (2019). "Ice shelf basal melt rates from a high-resolution digital elevation model (DEM) record for Pine Island Glacier, Antarctica". English. In: *The Cryosphere* 13.10, pp. 2633–2656. ISSN: 1994-0416.

Shepherd, A., D. Wingham, and E. Rignot (2004). "Warm ocean is eroding West Antarctic Ice Sheet". en. In: *Geophysical Research Letters* 31.23. ISSN: 1944-8007.

Shepherd, A. et al. (2009). "Greenland ice sheet motion coupled with daily melting in late summer". en. In: *Geophysical Research Letters* 36.1. ISSN: 1944-8007.

Shepherd, A. et al. (2012). "A Reconciled Estimate of Ice-Sheet Mass Balance". In: *Science* 338.6111, pp. 1183–1189.

Shepherd, A. et al. (2018). "Mass balance of the Antarctic Ice Sheet from 1992 to 2017". en. In: *Nature* 558.7709, pp. 219–222. ISSN: 1476-4687.

Shepherd, A. et al. (2020). "Mass balance of the Greenland Ice Sheet from 1992 to 2018". en. In: *Nature* 579.7798, pp. 233–239. ISSN: 1476-4687.

Siegert, M. et al. (2019). "The Antarctic Peninsula Under a 1.5°C Global Warming Scenario". In: *Frontiers in Environmental Science* 7, p. 102. ISSN: 2296-665X.

Simpson, D., F. Lindgren, and H. Rue (2012a). "In order to make spatial statistics computationally feasible, we need to forget about the covariance function". en. In: *Environmetrics* 23.1, pp. 65–74. ISSN: 1099-095X.

— (2012b). "Think continuous: Markovian Gaussian models in spatial statistics". In: *Spatial Statistics* 1, pp. 16–29. ISSN: 2211-6753.

Slater, T. et al. (2021). "Review article: Earth's ice imbalance". English. In: *The Cryosphere* 15.1, pp. 233–246. ISSN: 1994-0416.

Smith, A. M. et al. (2018). "Evidence for the long-term sedimentary environment in an Antarctic subglacial lake". en. In: *Earth and Planetary Science Letters* 504, pp. 139–151. ISSN: 0012-821X.

Smith, B. et al. (2020). "Pervasive ice sheet mass loss reflects competing ocean and atmosphere processes". In: *Science* 368.6496, pp. 1239–1242.

Smith, C. S. R. (2022). "Observing the Seasonal Evolution of Supraglacial Ponds in High Mountain Asia: A Supervised Classification Approach". en. In.

Smith, L. C. et al. (2015). "Efficient meltwater drainage through supraglacial streams and rivers on the southwest Greenland ice sheet". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.4, pp. 1001–1006. ISSN: 0027-8424.

Smith, L. C. et al. (2021). "Supraglacial River Forcing of Subglacial Water Storage and Diurnal Ice Sheet Motion". en. In: *Geophysical Research Letters* 48.7, e2020GL091418. ISSN: 1944-8007.

Smith, L. C. et al. (2017). "Direct measurements of meltwater runoff on the Greenland ice sheet surface". In: *Proceedings of the National Academy of Sciences* 114.50, E10622–E10631.

Sneed, W. A. and G. S. Hamilton (2007). "Evolution of melt pond volume on the surface of the Greenland Ice Sheet". en. In: *Geophysical Research Letters* 34.3. ISSN: 1944-8007.

Song, B. (2015). "ROC random forest and its application". PhD Thesis. The Graduate School, Stony Brook University: Stony Brook, NY.

Sousa, D. and C. Small (2017). "Global cross-calibration of Landsat spectral mixture models". In: *Remote Sensing of Environment* 192, pp. 139–149. ISSN: 0034-4257.

Souza, W. d. O. et al. (2022). "Analysis of Environmental and Atmospheric Influences in the Use of SAR and Optical Imagery from Sentinel-1, Landsat-8, and Sentinel-2 in the Operational Monitoring of Reservoir Water Level". en. In: *Remote Sensing* 14.9, p. 2218. ISSN: 2072-4292.

Spergel, J. J. et al. (2021). "Surface meltwater drainage and ponding on Amery Ice Shelf, East Antarctica, 1973–2019". en. In: *Journal of Glaciology* 67.266, pp. 985–998. ISSN: 0022-1430, 1727-5652.

Steffen, K., W. Abdalati, and J. Stroeve (1993). "Climate sensitivity studies of the Greenland ice sheet using satellite AVHRR, SMMR, SSM/I and in situ data". en. In: *Meteorology and Atmospheric Physics* 51.3, pp. 239–258. ISSN: 1436-5065.

Stevens, L. A. et al. (2015). "Greenland supraglacial lake drainages triggered by hydrologically induced basal slip". In: *Nature* 522.7554, 73–U163. ISSN: 0028-0836.

Stevens, L. A. et al. (2022). "Tidewater-glacier response to supraglacial lake drainage". en. In: *Nature Communications* 13.1, p. 6065. ISSN: 2041-1723.

Stewart, C. L. et al. (2019). "Basal melting of Ross Ice Shelf from solar heat absorption in an ice-front polynya". en. In: *Nature Geoscience* 12.6, pp. 435–440. ISSN: 1752-0908.

Stokes, C. R. et al. (2019). "Widespread distribution of supraglacial lakes around the margin of the East Antarctic Ice Sheet". en. In: *Scientific Reports* 9.1, p. 13823. ISSN: 2045-2322.

Stroeve, J. et al. (2007). "Arctic sea ice decline: Faster than forecast". en. In: *Geophysical Research Letters* 34.9. ISSN: 1944-8007.

Studinger, M. et al. (2022). "High-resolution imaging of supraglacial hydrological features on the Greenland Ice Sheet with NASA's Airborne Topographic Mapper (ATM) instrument suite". English. In: *The Cryosphere* 16.9, pp. 3649–3668. ISSN: 1994-0416.

Sudmanns, M. et al. (2020). "Assessing global Sentinel-2 coverage dynamics and data availability for operational Earth observation (EO) applications using the EO-Compass". In: *International Journal of Digital Earth* 13.7, pp. 768–784. ISSN: 1753-8947.

Sun, Y., B. Riel, and B. Minchew (2023). "Disintegration and Buttressing Effect of the Landfast Sea Ice in the Larsen B Embayment, Antarctic Peninsula". en. In: *Geophysical Research Letters* 50.16, e2023GL104066. ISSN: 1944-8007.

Sundal, A. V. et al. (2009). "Evolution of supra-glacial lakes across the Greenland Ice Sheet". In: *Remote Sensing of Environment* 113.10, pp. 2164–2171. ISSN: 0034-4257.

Sundal, A. V. et al. (2011). "Melt-induced speed-up of Greenland ice sheet offset by efficient subglacial drainage". en. In: *Nature* 469.7331, pp. 521–524. ISSN: 1476-4687.

Surawy-Stepney, T. et al. (2023). "Mapping Antarctic Crevasses and their Evolution with Deep Learning Applied to Satellite Radar Imagery". English. In: *The Cryosphere Discussions*, pp. 1–32.

Svoboda, J. et al. (2022). "Random Forest Classification of Land Use, Land-Use Change and Forestry (LULUCF) Using Sentinel-2 Data—A Case Study of Czechia". en. In: *Remote Sensing* 14.5, p. 1189. ISSN: 2072-4292.

Tedesco, M. and N. Steiner (2011). "In-situ multispectral and bathymetric measurements over a supraglacial lake in western Greenland using a remotely controlled watercraft". English. In: *The Cryosphere* 5.2, pp. 445–452. ISSN: 1994-0416.

Tedesco, M. et al. (2012). "Measurement and modeling of ablation of the bottom of supraglacial lakes in western Greenland". en. In: *Geophysical Research Letters* 39.2. ISSN: 1944-8007.

Tedesco, M. (2015). *Remote Sensing of the Cryosphere*. en. John Wiley & Sons. ISBN: 978-1-118-36885-5.

Tedesco, M. and X. Fettweis (2020). "Unprecedented atmospheric conditions (1948–2019) drive the 2019 exceptional melting season over the Greenland ice sheet". English. In: *The Cryosphere* 14.4, pp. 1209–1223. ISSN: 1994-0416.

Tedesco, M. et al. (2013). "Ice dynamic response to two modes of surface lake drainage on the Greenland ice sheet". en. In: *Environmental Research Letters* 8.3, p. 034007. ISSN: 1748-9326.

Tedstone, A. J. and H. Machguth (2022). "Increasing surface runoff from Greenland's firn areas". en. In: *Nature Climate Change* 12.7, pp. 672–676. ISSN: 1758-6798.

Trusel, L. D., K. E. Frey, and S. B. Das (2012). "Antarctic surface melting dynamics: Enhanced perspectives from radar scatterometer data". In: *Journal of Geophysical Research-Earth Surface* 117. ISSN: 2169-9003.

Trusel, L. D., Z. Pan, and M. Moussavi (2022). "Repeated Tidally Induced Hydrofracture of a Supraglacial Lake at the Amery Ice Shelf Grounding Zone". en. In: *Geophysical Research Letters* 49.7, e2021GL095661. ISSN: 1944-8007.

Trusel, L. D. et al. (2015). "Divergent trajectories of Antarctic surface melt under two twenty-first-century climate scenarios". en. In: *Nature Geoscience* 8.12, pp. 927–932. ISSN: 1752-0908.

Trusel, L. D. et al. (2018). "Nonlinear rise in Greenland runoff in response to post-industrial Arctic warming". en. In: *Nature* 564.7734, pp. 104–108. ISSN: 1476-4687.

Tuckett, P. A. et al. (2019). "Rapid accelerations of Antarctic Peninsula outlet glaciers driven by surface melt". en. In: *Nature Communications* 10.1, p. 4311. ISSN: 2041-1723.

Turner, J. et al. (2020). "Antarctic temperature variability and change from station data". en. In: *International Journal of Climatology* 40.6, pp. 2986–3007. ISSN: 1097-0088.

Turton, J. V. et al. (2020). "The influence of föhn winds on annual and seasonal surface melt on the Larsen C Ice Shelf, Antarctica". English. In: *The Cryosphere* 14.11, pp. 4165–4180. ISSN: 1994-0416.

Turton, J. V. et al. (2021). "The distribution and evolution of supraglacial lakes on 79°N Glacier (north-eastern Greenland) and interannual climatic controls". English. In: *The Cryosphere* 15.8, pp. 3877–3896. ISSN: 1994-0416.

Van de Berg, W. J. et al. (2006). "Reassessment of the Antarctic surface mass balance using calibrated output of a regional atmospheric climate model". en. In: *Journal of Geophysical Research: Atmospheres* 111.D11. ISSN: 2156-2202.

Van den Broeke, M. R., P. G. Duynkerke, and J. Oerlemans (1994). "The observed katabatic flow at the edge of the Greenland ice sheet during GIMEX-91". en. In: *Global and Planetary Change*. Greenland ice margin experiment (GIMEx) 9.1, pp. 3–15. ISSN: 0921-8181.

Van den Broeke, M. et al. (2009). "Partitioning Recent Greenland Mass Loss". In: *Science* 326.5955, pp. 984–986.

Van den Broeke, M. R. (1997). "Spatial and temporal variation of sublimation on Antarctica: Results of a high-resolution general circulation model". en. In: *Journal of Geophysical Research: Atmospheres* 102.D25, pp. 29765–29777. ISSN: 2156-2202.

Van den Broeke, M. R. et al. (2016). "On the recent contribution of the Greenland ice sheet to sea level change". English. In: *The Cryosphere* 10.5, pp. 1933–1946. ISSN: 1994-0416.

Van der Veen, C. J. (1998). "Fracture mechanics approach to penetration of surface crevasses on glaciers". en. In: *Cold Regions Science and Technology* 27.1, pp. 31–47. ISSN: 0165-232X.

— (1999). "Crevasses on glaciers". In: *Polar Geography* 23.3, pp. 213–245. ISSN: 1088-937X.

— (2007). "Fracture propagation as means of rapidly transferring surface meltwater to the base of glaciers". en. In: *Geophysical Research Letters* 34.1. ISSN: 1944-8007.

Van As, D. et al. (2018). "Reconstructing Greenland Ice Sheet meltwater discharge through the Watson River (1949–2017)". In: *Arctic, Antarctic, and Alpine Research* 50.1, S100010. ISSN: 1523-0430.

Vandecrux, B. et al. (2019). "Firn data compilation reveals widespread decrease of firn air content in western Greenland". English. In: *The Cryosphere* 13.3, pp. 845–859. ISSN: 1994-0416.

Van Wessem, J. M. et al. (2014). "Improved representation of East Antarctic surface mass balance in a regional atmospheric climate model". en. In: *Journal of Glaciology* 60.222, pp. 761–770. ISSN: 0022-1430, 1727-5652.

Vaughan, D. G. et al. (1993). "A synthesis of remote sensing data on Wilkins Ice Shelf, Antarctica". en. In: *Annals of Glaciology* 17, pp. 211–218. ISSN: 0260-3055, 1727-5644.

Vaughan, D. et al. (2013). *Observations: Cryosphere. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report*

of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Tech. rep. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Velicogna, I., T. C. Sutterley, and M. R. van den Broeke (2014). "Regional acceleration in ice mass loss from Greenland and Antarctica using GRACE time-variable gravity data". en. In: *Geophysical Research Letters* 41.22, pp. 8130–8137. ISSN: 1944-8007.

Verfaillie, D. et al. (2022). "The circum-Antarctic ice-shelves respond to a more positive Southern Annular Mode with regionally varied melting". en. In: *Communications Earth & Environment* 3.1, pp. 1–12. ISSN: 2662-4435.

Vermeer, M. and S. Rahmstorf (2009). "Global sea level linked to global temperature". In: *Proceedings of the National Academy of Sciences* 106.51, pp. 21527–21532.

Vihma, T., E. Tuovinen, and H. Savijärvi (2011). "Interaction of katabatic winds and near-surface temperatures in the Antarctic". en. In: *Journal of Geophysical Research: Atmospheres* 116.D21. ISSN: 2156-2202.

Villejo, S. J., J. B. Illian, and B. Swallow (2023). "Data fusion in a two-stage spatio-temporal model using the INLA-SPDE approach". In: *Spatial Statistics* 54, p. 100744. ISSN: 2211-6753.

Walker, C. C. and A. S. Gardner (2017). "Rapid drawdown of Antarctica's Wordie Ice Shelf glaciers in response to ENSO/Southern Annular Mode-driven warming in the Southern Ocean". In: *Earth and Planetary Science Letters* 476, pp. 100–110. ISSN: 0012-821X.

Wangchuk, S. and T. Bolch (2020). "Mapping of glacial lakes using Sentinel-1 and Sentinel-2 data and a random forest classifier: Strengths and challenges". en. In: *Science of Remote Sensing* 2, p. 100008. ISSN: 2666-0172.

Watson, C. S. et al. (2018). "Optimising NDWI supraglacial pond classification on Himalayan debris-covered glaciers". en. In: *Remote Sensing of Environment* 217, pp. 414–425. ISSN: 0034-4257.

Weertman, J. (1962). *Catastrophic glacier advances*. en-us. Report. Cold Regions Research and Engineering Laboratory (U.S.)

Wendleder, A. et al. (2021). "Evolution of supraglacial lakes on the Greenland Ice sheet". In: *EUSAR 2021; 13th European Conference on Synthetic Aperture Radar*, pp. 1–4.

Wheel, I., P. Christoffersen, and S. H. Mernild (2020). *Can katabatic winds directly force retreat of Greenland outlet glaciers? Hypothesis test on Helheim Glacier in Sermilik Fjord*. en. preprint. Glaciers/Glaciers.

Williamson, A. G. et al. (2017). "A Fully Automated Supraglacial lake area and volume Tracking ("FAST") algorithm: Development and application using MODIS imagery of West Greenland". en. In: *Remote Sensing of Environment* 196, pp. 113–133. ISSN: 0034-4257.

Williamson, A. G. et al. (2018). "Dual-satellite (Sentinel-2 and Landsat 8) remote sensing of supraglacial lakes in Greenland". English. In: *The Cryosphere* 12.9, pp. 3045–3065. ISSN: 1994-0416.

Wingham, D. J., D. W. Wallis, and A. Shepherd (2009). "Spatial and temporal evolution of Pine Island Glacier thinning, 1995–2006". en. In: *Geophysical Research Letters* 36.17. ISSN: 1944-8007.

Winther, J. G. et al. (1996). "Melting, runoff and the formation of frozen lakes in a mixed snow and blue-ice field in Dronning Maud Land, Antarctica". In: *Journal of Glaciology* 42.141, pp. 271–278. ISSN: 0022-1430.

Xu, H. (2006). "Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery". In: *International Journal of Remote Sensing* 27.14, pp. 3025–3033. ISSN: 0143-1161.

Yang, C. et al. (2023). "Soil organic matter mapping using INLA-SPDE with remote sensing based soil moisture indices and Fourier transforms decomposed variables". In: *Geoderma* 437, p. 116571. ISSN: 0016-7061.

Yang, K. et al. (2014). "River Delineation from Remotely Sensed Imagery Using a Multi-Scale Classification Approach". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7.12, pp. 4726–4737. ISSN: 2151-1535.

Yang, K. and L. C. Smith (2013). "Supraglacial Streams on the Greenland Ice Sheet Delineated From Combined Spectral–Shape Information in High-Resolution Satellite Imagery". In: *IEEE Geoscience and Remote Sensing Letters* 10.4, pp. 801–805. ISSN: 1558-0571.

— (2016). "Internally drained catchments dominate supraglacial hydrology of the southwest Greenland Ice Sheet". In: *Journal of Geophysical Research: Earth Surface* 121.10, pp. 1891–1910. ISSN: 2169-9003.

Yang, K. et al. (2017). "Automated High-Resolution Satellite Image Registration Using Supraglacial Rivers on the Greenland Ice Sheet". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10.3, pp. 845–856. ISSN: 2151-1535.

Yang, K. et al. (2019a). "Supraglacial rivers on the northwest Greenland Ice Sheet, Devon Ice Cap, and Barnes Ice Cap mapped using Sentinel-2 imagery". en. In: *International Journal of Applied Earth Observation and Geoinformation* 78, pp. 1–13. ISSN: 1569-8432.

Yang, K. et al. (2019b). "Surface meltwater runoff on the Greenland ice sheet estimated from remotely sensed supraglacial lake infilling rate". In: *Remote Sensing of Environment* 234, p. 111459. ISSN: 0034-4257.

Yang, K. et al. (2021). "Seasonal evolution of supraglacial lakes and rivers on the southwest Greenland Ice Sheet". en. In: *Journal of Glaciology* 67.264, pp. 592–602. ISSN: 0022-1430, 1727-5652.

Yin, F., P. E. Lewis, and J. L. Gómez-Dans (2022). "Bayesian atmospheric correction over land: Sentinel-2/MSI and Landsat 8/OLI". English. In: *Geoscientific Model Development* 15.21, pp. 7933–7976. ISSN: 1991-959X.

Yuan, J. et al. (2020). "Automatic Extraction of Supraglacial Lakes in Southwest Greenland during the 2014–2018 Melt Seasons Based on Convolutional Neural Network". en. In: *Water* 12.3, p. 891.

Zhang, L. et al. (2022). "The relative role of the subsurface Southern Ocean in driving negative Antarctic Sea ice extent anomalies in 2016–2021". en. In: *Communications Earth & Environment* 3.1, pp. 1–9. ISSN: 2662-4435.

Zhang, R. et al. (2022). "Tracking annual dynamics of mangrove forests in mangrove National Nature Reserves of China based on time series Sentinel-2 imagery during 2016–2020". In: *International Journal of Applied Earth Observation and Geoinformation* 112, p. 102918. ISSN: 1569-8432.

Zhu, D. et al. (2022). "Evolution of supraglacial lakes on Sermeq Avannarleq glacier, Greenland using Google Earth Engine". In: *Journal of Hydrology: Regional Studies* 44, p. 101246. ISSN: 2214-5818.

Zilong, T., S. Yubing, and D. Xiaowei (2022). "Spatial-temporal wave height forecast using deep learning and public reanalysis dataset". In: *Applied Energy* 326, p. 120027. ISSN: 0306-2619.

Zwally, H. J. et al. (2012). "Antarctic and Greenland drainage systems". In: *GSFC Cryospheric Sciences Laboratory* 270.