

Wilfrid Laurier University

Scholars Commons @ Laurier

Theses and Dissertations (Comprehensive)

2023

A Machine Learning Approach to Classify Open Water and Ice Cover on Slave River Delta

Ida Moalemi
imoalemi@wlu.ca

Follow this and additional works at: <https://scholars.wlu.ca/etd>



Part of the [Environmental Monitoring Commons](#)

Recommended Citation

Moalemi, Ida, "A Machine Learning Approach to Classify Open Water and Ice Cover on Slave River Delta" (2023). *Theses and Dissertations (Comprehensive)*. 2598.
<https://scholars.wlu.ca/etd/2598>

This Thesis is brought to you for free and open access by Scholars Commons @ Laurier. It has been accepted for inclusion in Theses and Dissertations (Comprehensive) by an authorized administrator of Scholars Commons @ Laurier. For more information, please contact scholarscommons@wlu.ca.

A Machine Learning Approach to Classify Open Water and Ice Cover on Slave River Delta

by

Ida Moalemi

A thesis

submitted to the Faculty of Science

in partial fulfilment of the requirements for

Master of Science in Geography

Wilfrid Laurier University

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

ABSTRACT

Seasonal temperature trend and ice phenology in Great Slave lake (GSL), are strongly influenced by warmer inflow from Slave river. The Slave river flows to GSL through Slave river delta (SRD), bringing a rise in temperature that triggers the ice break-up process of the lake. Slave river discharge is subject to multiple stressors including climate warming and upstream water activities, which in turn, directly affects the GSL break-up process. Consequently, monitoring the break-up process at SRD, where the river connects to the lake, serves as an indicator to better understand the cascading effects on GSL ice break-up. This research aims to develop random forest (RF) models to monitor the SRD ice break-up processes, using a combination of satellite images with optical sensors at high spatial resolution, including Landsat-5, Landsat-8, Sentinel-2a, and Sentinel-2b. The RF models were trained using manually selected training pixels to classify ice, open water, and cloud within the SRD.

The break-up start period is defined by minimum and maximum thresholds of 60% and 90% on ice fraction, which are a trade-off between maximizing the available images and not including images that are taken after the break-up start. The results show high variability in the rate of break-up within delta using images in recent years with better temporal resolution. Furthermore, a statistically significant trend is observed from 1984 to 2023 using the Mann-Kendall test, with the p-value of 0.05.

This study is of great significance to northern and high latitude communities who rely on lake ice for activities such as transportation, and sustenance. Moreover, the break-up of the delta plays a pivotal role in supplying nutrients and sediments, and also in the occurrence of spring flooding. Therefore, the outcomes of this study can be leveraged to shape effective water resource management policies based on the regional characteristics of climate and hydrological patterns.

ACKNOWLEDGEMENT

I would like to express my heartfelt appreciation to Prof. Homa Kheyrollah Pour for her endless support, continuous encouragement, and thoughtful guidance in every single step of this journey. Her mentorship and dedication have not only enriched my academic experience but has also inspired me in my personal life. I would like to extend my appreciation to Prof. Andrea Scott for her constructive criticism and thoughtful discussions which significantly enriched the depth and quality of this research. I am grateful for the opportunity I am provided to learn from them as my supervisors and to have benefitted from their wealth of knowledge. A special thanks to my committee members: Prof. Michael Bitton and Prof. Grant Gunn for their review and thoughtful suggestions.

I would also like to acknowledge ReSEC members, past and present, for providing an environment that fosters learning and growth. Your advice and essential critiques have been pivotal in shaping the outcome of this thesis.

Thank you to my family for their support through screens and messages. Your presence has always been felt in my heart, and our shared values have been a guiding light during this journey. Thank you to my husband, Hamidreza, for being my constant source of encouragement. This achievement is as much yours as it is mine.

Lastly, this research was supported by funding from Global Water Future (GWF) Remotely Sensed Monitoring of Northern Lake Ice Using RADARSAT Constellation Mission and Cloud Computing Processing project, Government of Northwest Territories (GNWT) Environment and Climate Change Cumulative Impact Monitoring Program and Natural Science and Engineering Research Council of Canada (NSERC) Canada Research Chair funding.

Table of Contents

| | |
|--|------|
| AUTHOR’S DECLARATION | ii |
| ABSTRACT..... | iii |
| ACKNOWLEDGEMENT | iv |
| LIST OF TABLES | vii |
| LIST OF FIGURES | viii |
| Chapter 1. General Introduction | 1 |
| 1.1 Canada Sub-Arctic Region | 1 |
| 1.1.1 Climatology..... | 3 |
| 1.2 Lake Ice..... | 5 |
| 1.2.1 Lake Ice Phenology..... | 6 |
| 1.2.2 Lake Ice and Climate Change Interaction..... | 7 |
| 1.3 Lake Ice Monitoring | 9 |
| 1.3.1 In-situ Observations | 9 |
| 1.4 Ice and Open Water Classification..... | 13 |
| 1.4.1 Threshold-based Methods | 13 |
| 1.4.2 Machine Learning Techniques..... | 15 |
| 1.2 Research Objectives..... | 17 |
| Chapter 2. Open Water and Ice Cover Classification on Slave River Delta Using Machine Learning | |
| Techniques | 19 |
| 2.1 Abstract..... | 19 |
| 2.2 Introduction..... | 20 |
| 2.3 Study Area | 23 |
| 2.4 Dataset and Preprocessing | 25 |
| 2.4.1 Landsat Archives | 26 |
| 2.4.2 Sentinel-2 Archives..... | 26 |
| 2.5 Method..... | 27 |
| 2.5.1 Random Forest Modelling | 27 |
| 2.5.2 Models Evaluation | 31 |
| 2.5.3 SRD Trend Analysis | 31 |
| 2.6 Results..... | 35 |
| 2.6.1 RF Models Validation..... | 35 |
| 2.6.2 Seasonal Dynamics of Ice and Water fractions..... | 37 |

| | |
|---|----|
| 2.6.3 Trend of break-up onset | 38 |
| 2.7 Discussion | 40 |
| 2.8 Conclusion | 42 |
| Chapter 3. Conclusion and Future Work | 44 |
| 3.1 Conclusion | 44 |
| 3.2 Future Works | 45 |
| Appendices..... | 47 |
| Appendix A. WICI and Local Average Gradient | 47 |
| Appendix B. Landsat Quality Assessment Layer | 48 |
| Appendix C. Configuration of Models Hyperparameters | 50 |
| Appendix D. Mahalanobis Distance for Feature Selection..... | 51 |
| REFERENCES | 52 |

LIST OF TABLES

| | |
|--|----|
| Table 1: Spatial resolution and central wavelength of Sentinel-2, Landsat-8, and Landsat-5 bands. | 27 |
| Table 2. Sorted bands based on Mahalanobis distance..... | 29 |
| Table 3. Accuracies of Landsat and Sentinel models for the SRD dataset. Landsat with two classes of ice and water outperforms Sentinel model with classes of ice, water, and cloud. | 35 |
| Table 4. Landsat 8 collection-1 and level-1 BQA band attributes and pixel value possibilities (From USGS)..... | 48 |
| Table 5. Landsat 5 collection-1 and level-1 BQA band attributes and pixel value possibilities (From USGS)..... | 49 |

LIST OF FIGURES

Figure 1. Canada’s landcover as of 2015 from 30m gridded data (CCRS and NRCan, 2020). Sub-Arctic boundary is delineated in black. 2

Figure 2. Canada's Terrestrial Ecoregion from 1950 to 2000. Data is provided by Commission for Environmental Cooperation (CEC), 2011. Sub-Arctic boundary is delineated in black..... 2

Figure 3: Mean Annual Temperature from 1950 to 2000. Data is provided by Commission for Environmental Cooperation (CEC), 2011. Sub-Arctic boundary is delineated in black..... 3

Figure 4. Total Annual Precipitation from 1950 to 2000. Data is provided by Commission for Environmental Cooperation (CEC), 2011. Sub-Arctic boundary is delineated in black..... 4

Figure 5. Observed changes in normalized seasonal precipitation (%) between 1948 and 2012 for the four seasons (From Canada Climate Change Report 2019). 5

Figure 6. Observed changes (°C) in seasonal mean temperatures between 1948 and 2016 for the four seasons (From 2019 Canada Climate Change Report). 8

Figure 7. Site map of Slave river delta (SRD) and Peace Athabasca delta (PAD) in Northwest Territories and Alberta, Canada. The a) SRD and b) PAD images are acquired on 2018-05-23, and 2014-05-14, respectively. 25

Figure 8. Examples of manually selecting training areas generated from a) Sentinel image captured on 2014-05-12 b) Landsat image captured on 2020-05-30, and c) Landsat image captured on 2019-05-10. The black, cyan, and blue colors correspond to cloud, ice, and water training areas. Training polygons with nearly equal contributions of ice and water pixels have even distribution over SRD..... 31

Figure 9. Distribution of cloud fractions of Landsat and Sentinel datasets from May 1 – May 30. Each open triangle corresponds to an individual image with a total number of 210. The cloud percentages are generated from the SRD boundary (The images are masked using the SRD shape file to exclude land pixels). 33

Figure 10. Data distribution provides insight into optimizing the thresholds of break-up identification. Data distribution provides insight into optimizing the thresholds of break-up identification. 33

Figure 11. Workflow of RF modelling and Trend analysis. 34

Figure 12. Examples of Landsat model performance with an overall accuracy of 97.8%. (a) RGB scene captured on 2013-05-25. (b) RGB scene captured on 2014-05-28. (c) and (d) Corresponding model classification plots. 35

Figure 13. Examples of Sentinel model performance with an overall accuracy of 91.5%. (a) RGB scene captured on 2019-05-21. (b) RGB scene captured on 2021-05-27. (c) and (d) Corresponding plots. 36

Figure 14. Evaluation of Landsat and Sentinel models performance by comparing their classified images captured on the same date (2018-05-23). (a) RGB plot captured by Landsat-8 satellite. (b) corresponding classified image (Sentinel-2b). (c) corresponding classified image (Landsat-8). The image shown here has mixed pixels of water and ice, mostly around the bottom left corner. Additionally, closer to SRD tributaries at the bottom, ice seems to be on top of water. In these cases, depending on ice portion or thickness and corresponding reflectance, the model determines the classes. In general, the lack of a clear boundary between ice and water areas makes the classification

more complicated and requires more involved method than simple decision trees or threshold-based classifications..... 36

Figure 15. Examples of classification within PAD. RGB scene acquired on (a) 2013-05-11, and (b) 2014-05-14, and corresponding plots in c) and d) respectively..... 37

Figure 16. Rate of break-up occurrence in a) 2018, b)2019, and c) 2020. Years with data from Sentinel-2a and Sentinel-2b indicate high variability in break-up rate. Solid lines are used as a visual aid. 38

Figure 17. Black triangles indicate non-cloudy images with ice fraction of 60% to 90%, while red positive points are corresponding weighted averages or estimated days for the start of break-up. A linear model is fitted to estimated break-up starts with confidence level of 0.99, RSE of 11.98, and slope of -0.84. 40

Figure 18. Images captured during the break-up period with less than 20% cloud fraction. Each colour corresponds to one satellite. Landsat 5 started in 1984 and continued until 2011. Landsat 8 starts afterward and still is active. Furthermore, to fill the temporal gaps in recent years, Sentinel data is used as supplementary data starting from 2015 and 2017. 42

Figure 19. Break-up anomalies from 1984 to 2023 using Landsat archives (right) and combination of Landsat and Sentinel archives (left). Lower temporal resolution resulted in more temporal gaps, however, years of 2015, 2017, 2022, and 2023 could be identified successfully without Sentinel records and have close estimation of break-up to the results with Sentinel records. 42

Figure 20. Performance of a) Landsat and b) Sentinel models in different configurations of Mtry and Ntree. A RF model has been generated for each potential value of hyperparameters, and its training accuracy is calculated using cross-validation method ($k = 10$)...... 50

Chapter 1. General Introduction

This chapter is structured into four main sections, each designed to provide background knowledge for this research. The first section (1.1) discusses the Canada Sub-Arctic region in terms of land cover, terrestrial ecoregions and climate. Section 1.2 reviews the lake ice and its interaction with climate in Canada's North. Then, section 1.3 reviews lake ice monitoring techniques through the utilization of remote sensing and in-situ observations. The subsequent section (1.4) discusses the classification of open water and lake ice, with particular emphasis on threshold-based methods and machine learning techniques. Lastly, section 5 concludes the chapter, where the research gap is addressed, and the objectives of the study are outlined. Understanding the literature presented in these sections assists the approach and direction taken in the research program.

1.1 Canada Sub-Arctic Region

Situated in the northern part of Canada, below the Arctic Circle, the Sub-Arctic spans several provinces and territories. The Sub-Arctic predominantly consists of temperate and sub-polar needleleaf forests (Figure 1), encompassing the taiga and Hudson Plains regions (Figure 2) (Rouse et al., 1997; Quinton et al., 2011). This region encounters distractive challenges due to ongoing environmental change. Consequently, monitoring this region holds immense significance for northern communities, whose lives are intricately connected to the health of the land and its resources (Davidson et al., 2003; Dibike et al., 2011; Barrette and Charlebois et al., 2018).

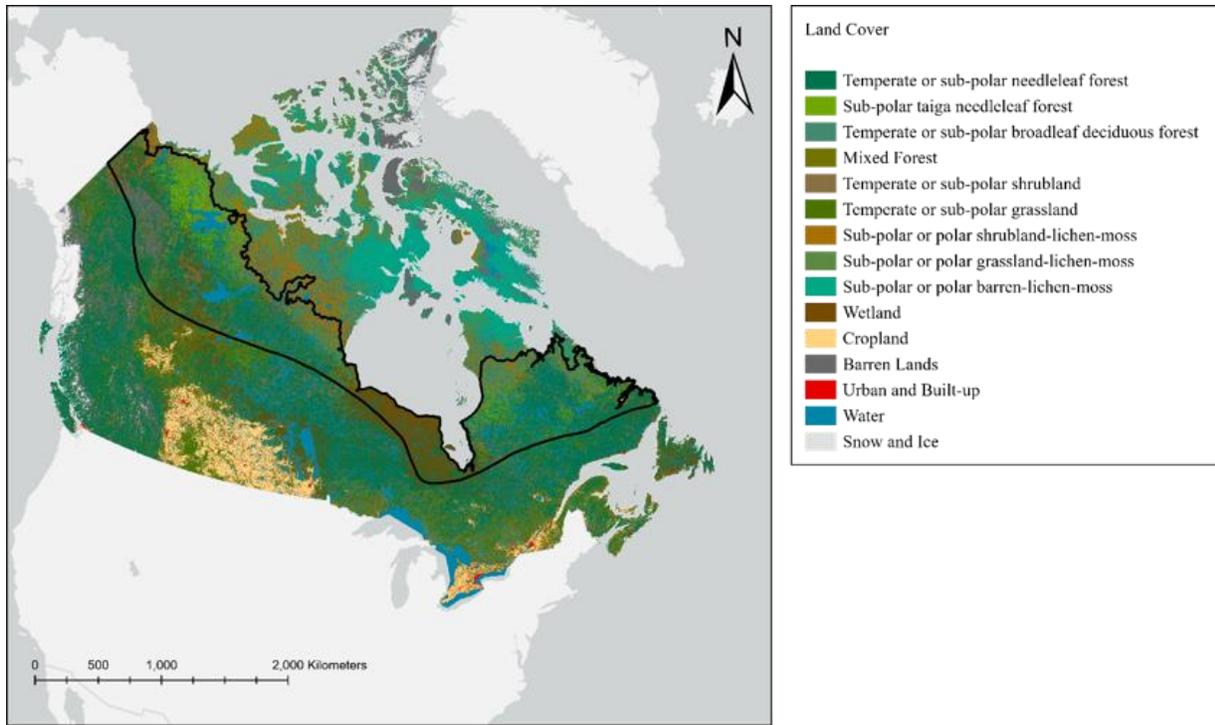


Figure 1. Canada's landcover as of 2015 from 30m gridded data (CCRS and NRCAN, 2020). Sub-Arctic boundary is delineated in black.

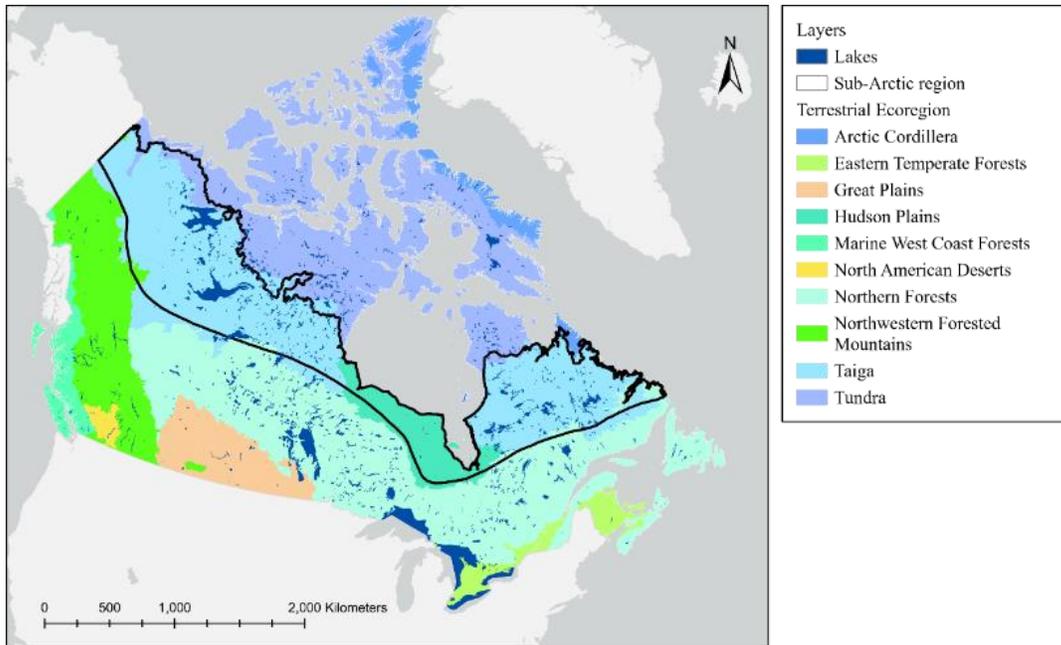


Figure 2. Canada's Terrestrial Ecoregion from 1950 to 2000. Data is provided by Commission for Environmental Cooperation (CEC), 2011. Sub-Arctic boundary is delineated in black.

1.1.1 Climatology

Influenced by its high latitude and proximity to the Arctic Circle, the Sub-Arctic region is characterized by extremely cold temperatures and limited daylight hours in winter, as well as relatively cool temperatures in its short summers (Figure 3) (Vincent et al., 2015). According to the updated Köppen-Geiger climate classification (Peel et al., 2007), the Sub-Arctic region is predominately covered by *Sub-Arctic climate* zone, which features year-round precipitation. Additionally, it includes the *continental Sub-Arctic* zone with dry summers, as well as the *ice cap climate* zone. Within the Sub-Arctic climate zone, mean temperatures rise above 10°C for one to three months of the year, while the coldest month averages below 0°C (Beck et al., 2018). While precipitation has increased across Canada in response to climate change, the Sub-Arctic region experiences relatively low amounts of precipitation compared to other parts of the country (Figure 4) (Vincent et al., 2018).

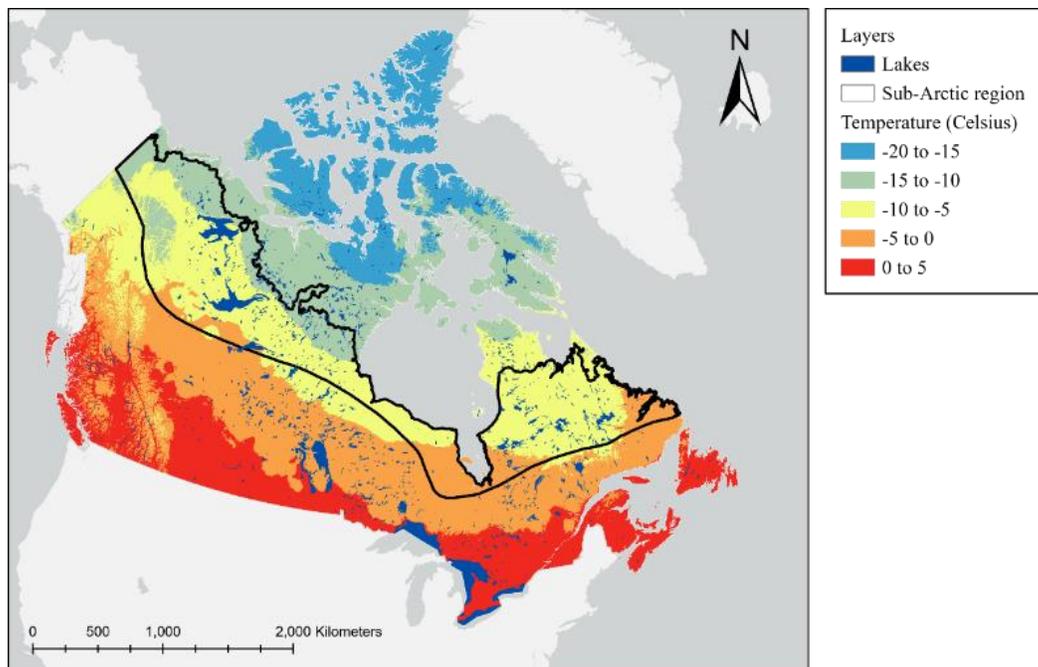


Figure 3: Mean Annual Temperature from 1950 to 2000. Data is provided by Commission for Environmental Cooperation (CEC), 2011. Sub-Arctic boundary is delineated in black.

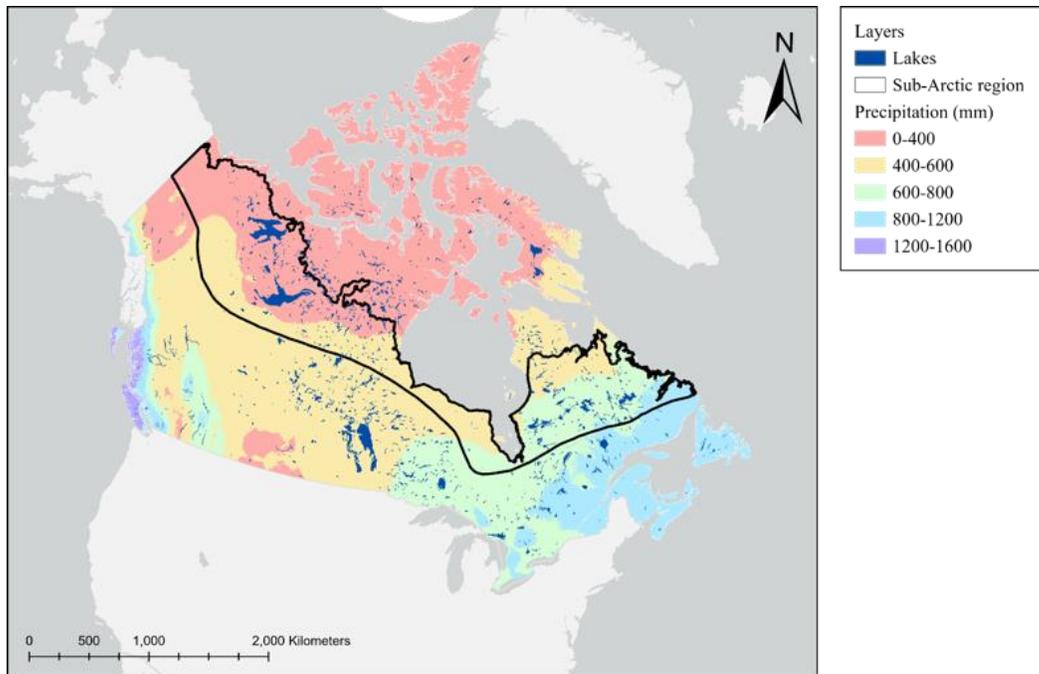


Figure 4. Total Annual Precipitation from 1950 to 2000. Data is provided by Commission for Environmental Cooperation (CEC), 2011. Sub-Arctic boundary is delineated in black.

The summer months (Figure 5) bring about precipitation that supports vegetation growth and contributes to the lush green landscapes of the area. During the winter, colder temperatures lead to precipitation primarily falling as snow. This snowfall can be substantial, resulting in accumulating thick snow cover that persists for several months (Estilow et al., 2015). These snowy conditions contribute to the formation of ice and frozen lakes, characteristic features of the Sub-Arctic environment (Duguay et al., 2003). It's worth noting that the Sub-Arctic encompasses various provinces and territories including the Yukon, Northwest Territories, Nunavut, parts of northern British Columbia, Alberta, Saskatchewan, Manitoba, Ontario, and Quebec. Precipitation patterns within this region can vary due to factors such as latitude, elevation, proximity to water bodies, and local geography. Coastal areas might experience more precipitation due to the influence of moist air from the ocean, while inland areas tend to be drier. Although the Sub-Arctic receives less precipitation

compared to other regions, the precipitation it receives plays a crucial role in shaping the unique ecosystems and climate conditions (Groisman et al., 1994).

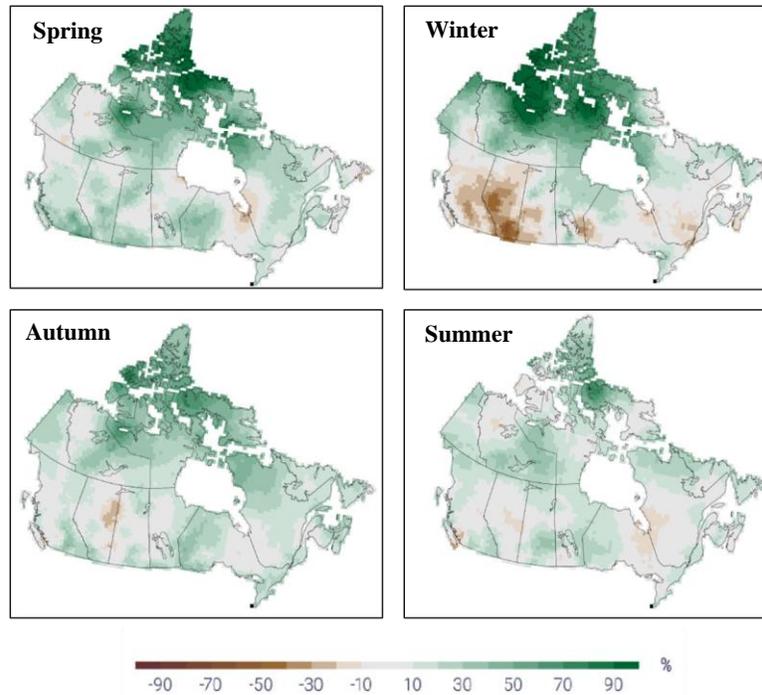


Figure 5. Observed changes in normalized seasonal precipitation (%) between 1948 and 2012 for the four seasons (From Canada Climate Change Report 2019).

1.2 Lake Ice

Lakes, reservoirs, streams, and rivers collectively cover approximately 3% of the Earth's land surface, with over half of the area covered by these waterbodies at latitudes above 44°N (Pekel et al., 2016). Within the Arctic and Sub-Arctic regions, lakes play a significant role, covering around 15% to 40% of the landscape with variations in depth and surface area (Duguay et al., 2003b). Lakes located in high latitude regions, including the Sub-Arctic, are characterized by both perennial and annual ice cover, exhibiting a diverse range of ice phenology such as break-up, freeze-up, and ice duration and ice composition varies across these regions. (Brown and Cote, 1992; Latifovic and Pouliot, 2007).

1.2.1 Lake Ice Phenology

Lake ice phenology refers to the distinct stages of ice formation, decay, and the duration of ice cover on lakes. Freeze-up defines the period between freeze onset (initial onset of ice formation) and the complete freeze-up (complete establishment of ice-cover), while break-up defines the period between melt onset and the point at which the water becomes free of ice cover (complete disappearance of ice). The process of lake ice formation, growth, and decay is significantly influenced by the energy from three sources: heat exchange between the lake and the atmosphere, the thermal energy stored within the lake, and heat transfer from inflowing water sources (William, 1965). Lake and atmosphere interactions are largely governed by factors such as air temperature, the presence of snow, solar radiation, and wind, whereas the heat stored and imported to the lake is governed by lake morphology and the discharge of rivers, respectively (Palecki and Barry, 1986; Vavrus et al., 1996; Schertzer et al., 2008).

Air temperature plays a dominant role in the energy balance of lake ice, consequently impacting the timing of both break-up and freeze-up (William, 1965). Greater heat storage within the lake during ice-free months can lead to delayed ice formation and vice versa (Bonsal et al., 2006). As autumn brings a decline in air temperatures, the lake's surface experiences heat loss, prompting vertical movement between the colder upper layer and the warmer beneath. The convective overturning continues until the entire water column reaches its highest density at 4°C. Ice formation initiates once the surface of water reaches 0°C, the freezing point of fresh water (Brown and Duguay, 2010). Ice freeze-up is relatively less influenced by changes in air temperature, compared to ice break-up. A study by Duguay et al. (2006) indicates significantly earlier break-up occurrences over Canadian lakes due to recent climate warming; however no clear pattern of changes in freeze-up dates was discerned.

Another climatic factor that can alter the energy balance of lake ice is the presence of snow. Snow has lower thermal conductivity ($2.24 \text{ Wm}^{-1}\text{k}^{-1}$) compared to ice ($0.08\text{-}0.54 \text{ Wm}^{-1}\text{k}^{-1}$), as well as high albedo (Oke, 1978; Sturm et al., 1997). Snowfall substantially delays the freeze-up process by acting as an insulation layer for lake ice (Adams, 1976; Adams and Roulet, 1980). Furthermore, wind characteristics, particularly wind velocity, play a vital role in changing the lake ice formation and decay by drifting and redistributing the snow (Bengtsson, 1986). Higher wind speeds result in less snow accumulation on the lake-ice surface, thus limiting the formation of snow-ice layers (Liston and Hall, 1995; Huang et al., 2017).

1.2.2 Lake Ice and Climate Change Interaction

The warming trend in northern Canada exceeds the global average, with approximately twice the magnitude of warming, particularly noticeable during the winter (Figure 6). The impacts of this widespread warming are evident in various Canadian regions and are expected to intensify in the future (Bush et al., 2019). These effects include more extreme heat events, reduced extreme cold spells, prolonged growing seasons, shorter period of snow and ice cover seasons, earlier peak spring streamflow, glaciers thinning, permafrost thaw, and rising sea level (Adrian et al., 2009; Jansen et al., 2007).

The response of lake ice to climate change depends on lake-specific characteristics. For instance, lakes with shallow depths are more affected by changes in heat storage than deeper lakes, and accordingly, more sensitive to climate change (Duguay et al., 2003b). Therefore, there is considerable regional variation in lake ice response to current and future climate warming. Despite these regional variabilities, literature shows a general trend toward earlier ice break-up, later freeze-up, and consequently shorter ice duration across Canada (Magnuson et al., 2000). Break-up and freeze-up dates have been extensively studied as an indicator of climate change since they can be influenced by

only a few degrees in air temperature fluctuation (Liston and Hall 1995; Livingstone et al., 2010). Analysis of 12-year (2002-2015) passive microwave ice record indicates an increasingly shorter ice duration for roughly 60.6% of Northern Hemisphere lakes (Du et al., 2017). According to 2019 Canada Climate Change Report, break-up is projected to occur 10 to 25 days earlier by mid-century and freeze-up dates to delay by 5 to 15 days, depending on lake-specific characteristics (Bush et al., 2019). Similarly, Dibike et al., (2011) predicted that in 2040-2079, lake-ice freeze-up dates will be delayed by 5-20 days, and break-up will be advanced by 10-30 days, resulting in an overall decrease in lake-ice cover duration by about 15–50 days.

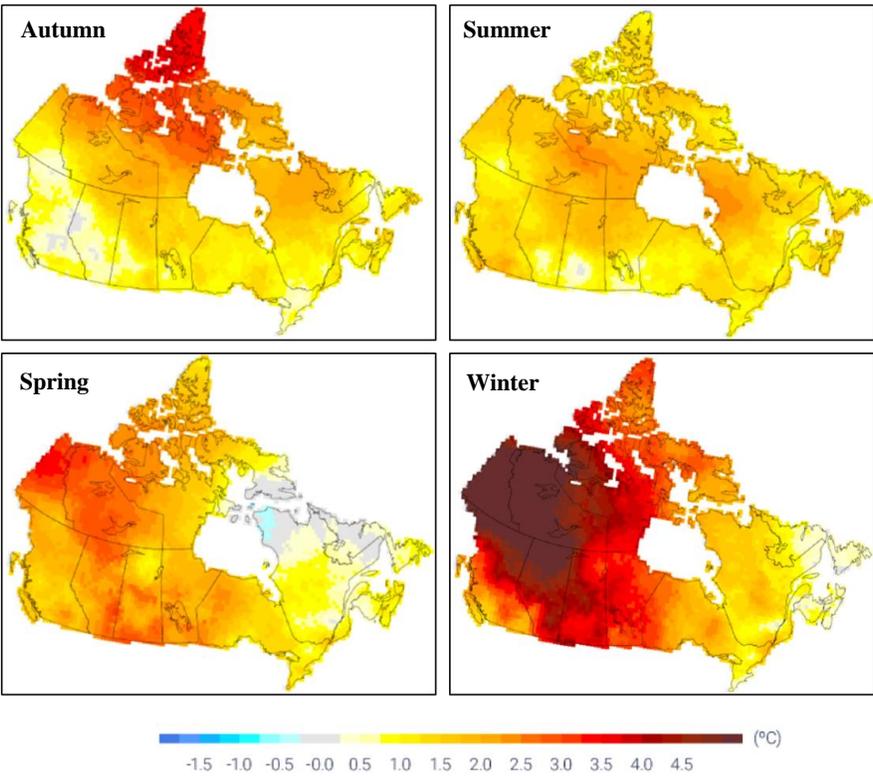


Figure 6. Observed changes (°C) in seasonal mean temperatures between 1948 and 2016 for the four seasons (From 2019 Canada Climate Change Report).

1.3 Lake Ice Monitoring

This section explores the available in-situ datasets and remote sensing observations, discussing their strengths and limitations in the context of ice monitoring. In-situ datasets refer to ground-based measurements and observations acquired directly from the study area. These data can include break-up and freeze-up dates estimation, localized ice thickness profile, ice properties, e.g., ice surface roughness and structure, snow cover thickness, and meteorological data. (Kheyrollah Pour et al., 2017; Lenormand et al., 2002). Remote sensing, on the other hand, refers to the process of collecting data without being in direct physical contact. Remote sensing data offers large-scale data, ranging from individual lakes to regional and global scale. Such observations typically rely on instruments such as satellites, aircraft, drones, or ground-based sensors (Campbell and Wynne, 2011).

1.3.1 In-situ Observations

For many years, Canadian government agencies have been actively collecting in-situ data on freshwater bodies, including lakes. Three departments within Environment Canada have been compiling records separately on ice conditions: the Meteorological Service of Canada Headquarters (MSC-HQ), the Canadian Ice Service (CIS), and the Water Survey of Canada (WSC) (Lenormand et al., 2002). Given that these records were not easily accessible and were dispersed across different government departments, a significant effort was undertaken to gather all available observations into a common database known as the Canadian Ice Database (CID). This repository houses 63546 indexed records from 757 sites, with 562 originating from the MSC-HQ and 195 from the CIS. The MSCHQ data mainly relate to freeze-up and break-up dates, whereas the CIS data provide information on ice thickness and snow on ice. Unfortunately, a decline in the ground-based observation networks that formed the basis for documenting changes in ice cover has occurred since the 1980s. The freeze-up and break-up network of 2000-01 only represents 4% of what it

encompassed in 1985-86. Similarly, a drastic decline of the ice thickness and the snow on ice network is also observable, accounting for just 10% of its 1984-85 counterpart by 1997-98. The severe reduction in ice observation networks can be attributed, in part, to significant budget cuts within Canadian government agencies during the late 1980s and the 1990s. Currently, the CIS monitors ice cover on a weekly basis for over 130 lakes across North America using a combination of SAR and optical imagery. The acquired scenes are subject to visual interpretation, and ice cover is reported as a fraction out of ten on a weekly basis for each lake. This data is used in a numerical weather prediction system operated by Environment and Climate Change (ECCC) (Lenormand et al., 2002).

Another main source of in-situ datasets comes from traditional knowledge and also volunteer-based monitoring programs, such as Nature Watch. Nature Watch is initiated by ECCC in partnership with the Canadian Nature Federation as part of lake and river break-up and freeze-up network. While volunteers in the volunteer-based program are provided with training and guidelines, the dataset may not be as consistent as remote sensing data since it can be influenced by variations in observer scope and experience. (Futter, 2003). Traditional knowledge is often held by indigenous and local communities. Rooted in the depth of cultural heritage, this form of knowledge offers profound insights and perspectives that have evolved through long-term, intergenerational experiences. Nevertheless, similar to any data source, traditional knowledge does present certain challenges, notably the lack of consistency in current form of documentation and its limited applicability beyond specific local contexts.

In conclusion, while in-situ datasets are well-suited to the purposes of validation and calibration, they face spatial and temporal limitations as the main source of data for lake ice monitoring. Remote Sensing Observation

In contrast to in-situ datasets, remote sensing observations offer a comprehensive and continuous perspective of lake and river ice cover at regional and global scales. Remote sensing employs sensors

such as optical, microwave, and thermal infrared to provide a wide range of temporal, spatial, and radiometric resolutions for lake monitoring (Chaouch et al., 2014; Dörnhöfer and Oppelt, 2016; Murfitt et al., 2021). Optical sensors, such as Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) aboard Landsat satellites and Moderate Resolution Imaging Spectroradiometer (MODIS) on board the Terra and Aqua satellites, have been used extensively in lake monitoring due to their high to medium spatial and temporal resolution (Paltan et al., 2015; Wu et al., 2021). MODIS as part of NASA's Earth Observing System (EOS) is designed to provide near-daily coverage (Sirguey et al., 2008; Chaouch et al., 2014; Zhang et al., 2021). Despite its moderate spatial resolution ranging from 250 meters to 1 kilometer, numerous studies, as reviewed by Dörnhöfer and Oppelt (2016), and Talukdar et al., (2021), have benefit from its global coverage and frequent revisits in land cover and land use change, and sea and lake monitoring (García-Mora et al., 2012; Dörnhöfer and Oppelt, 2016; Talukdar et al., 2020).

Furthermore, Landsat data, jointly managed by NASA and the United States Geological Survey (USGS), has been collecting Earth observation data since 1972, offering high spatial resolution of 30 meters for most spectral bands (Sirguey et al., 2008). Several studies suggest that combining Landsat data with other sources may be necessary to achieve higher temporal resolution and thus, more precise monitoring of lake and river ice phenology (Yang et al., 2021; Li et al., 2023).

In lake ice mapping, optical data suffers from three limitations: First, their utility is constrained in polar darkness due to the absence of solar radiation. Second, spectral overlap between ice and cloud leads to significant misclassification errors. Third, a study by Heinilä et al., (2021) indicates that the optical data may not discern between open water pixels and dark ice pixels with a water layer on top (Heinilä et al., 2021). Conversely, passively radiated microwave radiation from the Earth's surface can penetrate clouds at low frequencies, providing all weather observation with a significant

advantage in regions with persistent cloud cover (Girard et al., 2002; Murfitt and Duguay, 2021). Despite this advantage, passive sensors face challenges arising from their coarse spatial resolution, which can limit their ability to discern finer details on the Earth's surface (Cai et al., 2020). Additionally, active sensors have a limitation in their historical time-series, making them more adept at capturing recent changes rather than extensive long-term observations (Hoekstra et al., 2020).

Airborne observation involves collecting data using sensors and instruments mounted on aircraft or remotely piloted aircraft systems (RPAS). RPAS, also known as drones, have gained popularity in recent years for monitoring due to their agility and low-altitude capabilities (Alfredsen et al., 2014; Kalke and Loewen, 2018). They can carry lightweight sensors like optical cameras, thermal cameras, and LiDAR. While satellites typically have coarser spatial resolution, which limits the ability to capture fine-scale details of the lake, airborne platforms can offer more localized and detailed observations (Everaerts, 2008). However, localized observation can restrict the large scale context, and also airborne observation faces constraints in terms of lower temporal coverage due to logistical challenges and higher operational costs than satellite-based data access. To maximize the benefits, a combination of satellite and airborne observations can be employed (Koenig et al., 2010; Chu and Lindenschmidt, 2016). Integrating data from both sources provides detailed local analysis within the context of broader-scale observations and also short-term airborne data can be used in the evaluation process. A study by Chu and Lindenschmidt (2016) integrated space-borne and air-borne data in monitoring river ice processes in the Slave river. They developed an automated approach using MODIS and RADARSAT-2 imagery to monitor the freeze-up process, and characterize ice and water (e.g., intact ice, smooth rubble ice, rough rubble ice and open water) during break-up process, respectively. They compared the results obtained from the analysis of MODIS and RADARSAT-2 with time-lapse photos and aerial surveys along the Slave river (Chu and Lindenschmidt, 2016).

1.4 Ice and Open Water Classification

Several methods have been proposed for mapping sea ice and lake ice to address human biases associated with visual interpretation and to enhance analysis efficiency (Xu et al., 2017; Wu et al., 2021; Saberi et al., 2022). As one of various popular proposed methods, classification algorithms range from parametric to non-parametric approaches, depending on data distribution. In parametric classification, data is assumed to follow a specific probability distribution, and the classification model's parameters are estimated based on the data's statistical characteristics. Conversely, non-parametric approaches do not assume any particular data distribution. Classification trees (Breiman et al., 1984; Friedl and Brodley, 1997), neural networks (Atkinson and Tatnall, 1997), support vector machines (SVM), RF (Hoekstra et al., 2020) are some of the more popular non-parametric data driven models.

This chapter focuses on the classification methods employed to discriminate between lake open water and ice cover, with a particular emphasis on two approaches: threshold-based methods and machine learning techniques.

1.4.1 Threshold-based Methods

Early efforts in lake open water and ice cover classification initially revolved around the utilization of single-band thresholding techniques. Single-band methods relied on the stability of water and ice reflectance in one single band, mostly on water's low reflectance in longer wavelengths (Bennett, 1987). Subsequent studies demonstrated the potential of multi-band processing and algebraic-operation-based indexes, such as Normalized Difference Water Index (NDWI) (McFeeters, 1996; English et al., 1999; Frazier et al., 2000; Xu, 2006; Hollstein et al., 2016; Barbieux et al., 2021). The reflectance from lake open water and ice cover may change depending on water clarity and presence of algae or other organisms. In general, lake materials can be characterized into four categories. 1-

Deep water with the lowest reflectance compared to other classes. 2- Shallow water with the common property of stable ultra-blue values and low shortwave infrared values. 3- Opaque ice with high reflectance in the visible and the near-infrared bands. 4- Clear ice with similar reflectance of water below it (Barbieux et al., 2018). Water has notable low reflectance in both the NIR and the shortwave infrared (SWIR) and the main difference between shallow and deep water is the visibly brighter reflectance of deep water and stable ultra blue and very low shortwave infrared properties of shallow water. Ice, on the other hand, shows lower reflection in the NIR, SWIR, and a high degree of reflection at visible wavelengths. Also, the surface texture and morphology of ice differ from that of open water and can be projected to texture-based indexes. Despite the unique features of ice and open water in general, different reflectance of ice, and water across each lake depending on variables, such as depth turbidity and also the existence of mixed pixels of ice and water makes it challenging for threshold-based method to provide robust classification (Wu et al., 2021). Furthermore, the high solar zenith angles in high-latitude regions results in lower TOA reflectance over lakes, and therefore threshold-based retrieval algorithms using TOA satellite data do not perform well under such a condition (Šmejkalová et al., 2016). However, the latter mentioned challenge also can be addressed by using SAR data in the studies focused on recent years. A study by Howell et al. (2009) utilized Ku-band QuikSCAT data to assess ice phenology events, such as melt onset, water clear of ice, and freeze onset dates for Great Bear and GSL from 2000 to 2006. The stable backscatter pattern observed during winter months is used to establish thresholds for estimating the timing of freeze-up, break-up, and water becoming clear of ice (Howell et al., 2009). A subsequent study by Surdu et al., (2015) employed the same threshold-based approach to determine ice phenology for lakes on the North Slope of Alaska, combining ASAR and RADARSAT-2 data from 2005 to 2011 (Surdu et al., 2015).

1.4.2 Machine Learning Techniques

Machine learning (ML), a subset of data-driven approaches, has gained significant popularity within remote sensing applications (Tom et al., 2018; Wu et al., 2021; De Coste et al., 2022). This is primarily attributed to its capacity to comprehend the intrinsic features of data and extract valuable information, leading to notable improvements in classification accuracy.

The type of ML technique used in a given task, whether supervised or unsupervised, depends on the specific objectives, such as available training data, and the target accuracy. Although supervised classification can be influenced by biases present in training datasets, the predefined training allows for a more targeted and precise classification process. Unsupervised classification, on the other hand, autonomously identifies patterns and clusters within the dataset without prior training or labeling, which may require additional post-processing and validation steps to interpret and assign meaningful labels to the identified clusters. Semi-supervised classification combines elements of both supervised and unsupervised approaches, leveraging a smaller set of labeled data in conjunction with a larger unlabeled dataset to achieve improved accuracy and efficiency in the classification process (Verpoorter et al., 2012; Yan et al., 2020).

Several ML models have been developed for ice retrieval from optical imagery (Kalke and Loewen, 2018; Yang et al., 2021; De Coste et al., 2022). A study by Wu et al. (2021) investigates the capability of four machine learning classifiers (i.e. multinomial logistic regression, MLR; SVM; RF; gradient boosting trees, GBT) for mapping lake ice cover, water and cloud cover using the MODIS/Terra L1B TOA (MOD02) product. Their results show RF and GBT provided overall and class-specific accuracies above 98% and the two tree-based classifiers (SVM and RF) offered the most robust spatial transferability over the 17 lakes and performed consistently well across ice seasons (Wu et al., 2021).

ML techniques also can be integrated with other classification or segmentation methods (Leigh et al., 2013; Hoekstra et al., 2020). A study by Hoekstra et al. (2020) presents an approach to classify lake open water and ice cover that integrates the iterative region growing using semantics (IRGS) algorithm with supervised RF, and SVM, and manual classification. IRGS first locally segments homogeneous regions in an image, then merges similar regions into classes across the entire scene. Accuracies calculated based on author-generated reference data and reported ice fraction from CIS indicate that the RF classification, with an average accuracy of 95.8%, outperformed SVM and the manual approach (Hoekstra et al., 2020).

RF is known as a “random” forest because it introduces randomness during the construction of individual decision trees and combines their predictions to make the final classification. Specifically, the algorithm randomly selects a subset of the training data (with replacement) for each decision tree in the forest. This process is known as bootstrapping or bagging. At each node of the decision tree, instead of considering all features, RF randomly selects a subset of features to make a split. This helps reduce the correlation between trees and promotes diversity. Each decision tree is constructed by recursively splitting the data based on selected features, typically using metrics like Gini impurity. The process continues until a stopping criterion is met, such as reaching a maximum depth or minimum number of samples per leaf. Once all the decision trees are built, the final prediction is determined by majority voting, i.e., the class that receives the most votes from the individual trees is selected as the predicted class.

RF classifier, like other ensemble learning algorithms, can handle complex classification scenarios by producing multiple classifications, and ensuring a convergent approach to pixel labeling. Additionally, ensemble learning algorithms outweigh the error of a single classification. Another advantage of RF classifier is its ability to reduce the chance of overfitting issue (Yang et al., 2021).

Overfitting is a major challenge in implementing ML classification; it occurs when the accuracy from a training dataset is much higher than that of the validation dataset. Overfitting happens when the training dataset is either too small or biased to represent the actual data distribution and variation. RF hyperparameters configuration including the number of trees, denoted as Ntree and the number of variables available to a split, denoted as Mtry, also can be used to control the overfitting issue (Belgiu and Drăguț, 2016).

In conclusion, after careful consideration of various classification methods, the RF classification has been chosen. The decision to opt for RF is based on its non-sensitivity to outliers in dataset, its proven ability to effectively control the overfitting issue, handle complex classification scenarios, provide robust predictions, and exhibit model generalization capabilities (Breiman, 2001; Belgiu and Drăguț, 2016). The above-mentioned features make RF classifiers well-suited for analyzing the dynamic break-up process across rivers and lakes.

1.2 Research Objectives

The general objective of this research is to use ML techniques to improve our understanding of GSL ice break-up process by monitoring SRD where the lake break-up starts. The aim of this research will be achieved through the following specific objectives.

- 1- Develop two RF models to classify ice, open water, and cloud on SRD using Landsat and Sentinel observations respectively.
- 2- Identify the start of break-up process, following the results of RF models, and analyze statistically from 1984 to present.

Chapter 2, written in manuscript form, details the above-mentioned research objectives. The final chapter (Chapter 3) of this thesis provides a general summary discussing contributions made to the

field. The chapter also talks about the next steps and the following objectives need to be investigated in future works.

Chapter 2. Open Water and Ice Cover Classification on Slave River Delta Using Machine Learning Techniques

2.1 Abstract

Great Slave lake (GSL) experiences seasonal variation in long-term temperature trends and ice phenology, strongly influenced by climate change, and cascading effects of upstream water management, and water extraction on Slave river inflow. The Slave river flows through Lake Athabasca and the Slave river delta (SRD) before reaching GSL, bringing a rise in temperature that triggers the ice break-up of the lake. Therefore, monitoring the break-up process at SRD, where the river connects to the lake, serves as an indicator to better understand the effects of changes in Slave river inflow on GSL ice break-up. This research aims to use machine learning techniques for mapping the SRD ice break-up processes. To achieve this goal, a combination of satellite images with optical sensors at high spatial resolution, including Landsat-5, Landsat-8, Sentinel-2a, and Sentinel-2b, are used to train random forest (RF) models. The Landsat and Sentinel models have accuracies of 91.5% and 97.8%, respectively, by testing with independent scenes.

The break-up start is estimated based on the proportion of the ice versus water pixels in images with less than 20% cloud coverage. The start of break-up period is defined by minimum and maximum thresholds of 60% and 90% on ice fraction, which are a trade-off between maximizing the available images and not including images that are taken after the break-up start. The results show high variability in the rate of break-up within delta using images in recent years with better temporal resolution. Furthermore, a statistically significant trend is observed from 1984 to 2023 using the Mann-Kendall test, with the p-value of 0.05. These findings aim to improve our understanding of GSL break-up in response to upstream water activities and climate change.

2.2 Introduction

Seasonal ice cover across rivers and lakes plays a critical role in climatology, closely intertwined with both global and regional variability (Prowse and Beltaos 2002; Brown and Duguay, 2010). Lake ice has significant implications for ecological, hydrological, and chemical processes (Prowse, 2001; Prowse and Brown, 2010; Hampton et al., 2017; Denfeld et al., 2018) and has socio-economic importance in regions where it supports sustenance and transportation (Lynch et al., 2010; Barrette and Charlebois, 2018). Despite the regional variability of lake ice characteristics, climate warming has caused a general trend toward later freeze-up and earlier break-up processes in northern lakes, shortening the duration of the ice season (Robertson et al., 1992; Duguay et al., 2006; Dauginis and Brown, 2021; Rafat et al., 2023). Changes in seasonal ice duration can significantly affect the lake by impacting lake circulation, solar radiation inputs, the exchange of gases between the atmosphere and water, and the overall heat budget (Fujisaki et al., 2013; Cavaliere et al., 2021).

Previous studies have primarily focused on the characteristics of river and lake ice separately, with river ice exhibiting multifaceted drivers of change due to dynamic flow, and lake ice being mainly influenced by climate factors (Palecki et al., 1986; Williams et al., 2006; Prowse et al., 2007; Mishra et al., 2011). However, the factors influencing the ice cover of rivers and lakes are intricately linked within a unified system. As the connection of rivers and lakes, the delta is a pivotal node in studying the historical behavior of these interconnected components.

To fully understand this topic and examine long-term trends of changes, the availability of a long data record helps in capturing the complex extent of changes. Since in-situ observations are spatially and temporally limited especially in the Arctic and Sub-Arctic regions, lake and river ice monitoring largely relies on the use of satellite observations. Various satellite sensors have been used to monitor seasonal ice cover (Latifovic et al., 2007; Surdu et al., 2016; Scott et al., 2020). Optical sensors, including Landsat series, offer a valuable historical data series back to 1970s, despite limitations

posed by relatively low temporal resolution, cloud cover, and low light conditions during the polar night at high latitudes.

Accurately monitoring and classifying seasonal ice cover necessitates precise exclusion of optical characteristics from lake components, including open water, snow/ice, and cloud cover. Previous studies have demonstrated that the optical reflectance of cloud pixels, especially in visible and thermal bands, closely resembles that of a surface covered with ice/snow. Furthermore, the reflectance of the water exhibits variation due to turbidity and the presence of suspended materials. Monitoring of water in shallow areas also can vary in visible and NIR range by the reflectance originating from the underlying surface (Barbieux et al., 2018). Additionally, the reflectance of ice experiences fluctuations depending on ice types and composition (Doxaran et al., 2002; Hall and Riggs, 2007).

Given the persistent and projected climate warming, it is essential to generate data on the lake and river ice phenology, utilizing contemporary and precise methods. Various classification methods from traditional threshold-based decision trees to ML and deep learning methods have been used in river and lake ice classification (Frazier et al., 2000; Xu, 2006; Hollstein et al., 2016; Scott et al., 2019; Wu et al., 2021). Traditional decision trees, mostly rely on NIR and SWIR in optical imagery (Barbieux et al., 2018; Zhang et al., 2019) and temporal evolution of C-band backscatter in synthetic aperture radar (SAR) imagery (Murfit et al., 2018). A study by Kang et al. (2012) developed a threshold-based approach by using daily time series of brightness temperature from the Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E) and showed that 18.7 GHz H-pol is the most suitable channel for detecting ice phenological events on GBL and GSL (Kang et al., 2012). While threshold-based techniques have made a significant contribution to the classification of river and lake ice, recent studies are increasingly shifting towards learning techniques, ranging from parametric to non-parametric and supervised to un-supervised classification. This is because of ability

of learning techniques to capture intricate patterns of data, model generalizability across different study areas, and capacity to address limitations posed by threshold-based approaches (Hoekstra et al., 2020; Yang et al., 2021).

Among un-supervised approaches, K-means has been frequently used for mapping river ice types. A study by Sobiech and Dierking (2013) evaluated the performance of the k-means classification on lakes and river channels of the central Lena Delta and showed that it is comparable to that of a fixed-threshold approach (Sobiech and Dierking, 2013). Neural networks, support vector machines (SVM), and RF classifiers are popular examples of supervised learning method used in river and lake ice classification (Singh et al., 2020; Kirsikka Heinil et al., 2021; Sola and Scott, 2022). A review study by Belgiu and Drăguț, (2016) reports that RF classifier outperforms artificial neural network classifiers in terms of classification accuracy, and provide slightly better results than SVM for high dimensional input data such as hyperspectral imagery (Ham et al., 2005). A subsequent study by Wu et al., 2021 investigates the capability of four learning classifiers (i.e., multinomial logistic regression, SVM, RF, gradient boosting trees) using the MODIS/Terra L1B TOA (MOD02) product. Their results show RF and GBT provided overall and class-specific accuracies above 98% and the two tree-based classifiers (SVM and RF) offered the most robust spatial transferability over the 17 lakes and performed consistently well across ice seasons (Wu et al., 2021).

In this study, we used Landsat-5, Landsat-8, Sentinel-2a, and Sentinel-2b datasets to develop RF models for classifying open water, ice and cloud pixels in the SRD, Northwest Territories, Canada. The RF classification results are used to estimate the start of the SRD break-up process from 1984 to 2023, highlighting a statistically significant trend from 1984 to 2023 using two tailed Mann-Kendall test.

The subsequent sections of this paper are structured as follows: The next section provides background information on the study area and datasets used. The method section describes the RF modelling, feature selection, evaluation, and statistical analysis. Moving forward, the Results and Discussion section presents the modelling outputs and results of analysis. Finally, the paper concludes with a summary of the key findings and paper contribution.

2.3 Study Area

The research study area is SRD where the Slave river meets the GSL, located in the Mackenzie River Basin in the Northwest Territories, Canada (Figure 7). GSL (61°40'N, 114°W) with a surface area of 28.6×10^3 km² and an average depth of 76 m feeds several rivers including the Slave river (Rouse et al., 2008). The Slave river flows from Peace Athabasca delta (PAD) in Alberta and provides 74% of GSL inflow (English et al., 1997) and 82% of the GSL outflow into the Mackenzie River (Rouse et al., 2008).

SRD has a surface area of approximately 400 km² and comprises hundreds of shallow (<4 m deep) and small (<3 km²) lakes (Brock et al., 2009) with side channels of varying depths and widths, including Old Steamboat Channel, Resdelta Channel, Middle Channel, Nagle Channel, and the Jean River (English et al., 1997).

Slave river has a drainage area of 616 400 km², including Peace river, Athabasca river, lake Athabasca and the GSL sub-basins in Alberta, British Columbia, Saskatchewan, and the NWT (English et al., 1997). Notably, Peace river provides ~65% of the Slave river annual flow and joins the Athabasca river at lower Peace Athabasca delta (PAD). PAD is one of the world's largest inland freshwater deltas, and situated within Wood Buffalo National Park, a UNESCO World Heritage Site. The PAD is closely tied to the hydrological system of the region. Observations by various agencies (Alberta Environment and Parks, BC Hydro, Parks Canada, ECCC) have indicated that mechanical

breakup is initiated far upstream of the PAD and often triggered by the ice run of the Smoky river, a major tributary at ~800 km above the mouth of Peace river; it then essentially progresses in a sequence of intermittent break-ups all the way to Slave river and ultimately to GSL. While thermal breakups are preceded by advanced thermal decay of the ice, above mentioned break-up is mostly governed by mechanical break-up. Mechanical break-up occurs when the physical forces exerted by water flow break apart the ice cover (Beltaos and Bonsal, 2021)

Slave river flow has significant influence on GSL long term temperature data, water quality, and ice phenology as it is transferred to GSL via the mouth of SRD (Ménard et al., 2002). Slave river discharge is one of main drivers of GSL water lower transparency and has significant influence on the optical properties of GSL water, due to the high concentration of suspended particles and dissolved organic matter carried by the river (Evans and Muir, 2016). Moreover, the Slave river flow plays a significant role in transmitting heat and energy via SRD, thereby influencing convective heat fluxes within GSL. This phenomenon leads to an increase in lake temperature, triggering the break-up process of the lake (Ménard et al., 2002; Rouse et al., 2008; Schertzer et al., 2008; Kang et al., 2012). Accordingly, SRD break-up process can be served as an indicator of GSL overall break-up trend. Slave river flow is subject to several stressors, including upstream water management, water extraction, and climate change. These changes results in Slave river regime alteration, such as lower flood peaks during the summer and higher discharges in the winter due to the construction of the W.A.C. Bennett Dam in 1968 (English et al., 1997; Dubé and Wilson, 2013; Thuan Chu et al., 2016). Therefore, monitoring the SRD is vital to improve our understanding of the cascading effects of Slave river upstream activities on GSL ice phenology.

Additionally, as previously discussed, the discharge of the SRD plays a crucial role in shaping the circulation patterns and heat fluxes within GSL, consequently influencing the fundamental

characteristics of water and ice in the region. The implications of these changes extend beyond the physical environment. The water and ice characteristics influenced by SRD discharge, such as temperature, turbidity, and nutrient availability shape the habitats and life cycles of organisms, ultimately, impacting the ecological balance and biodiversity of the lake (English et al., 1997; Brock et al., 2007). Also, the changes impact the lives of the local communities as they rely on lake ice for transportation and crossing via ice roads, and sustenance through fishing (Rafat et al., 2021). Given the impacts of SRD discharge on human communities, ecosystems, and regional climate, the delta region where SRD flows into GSL emerges as a critical geographic location for monitoring.

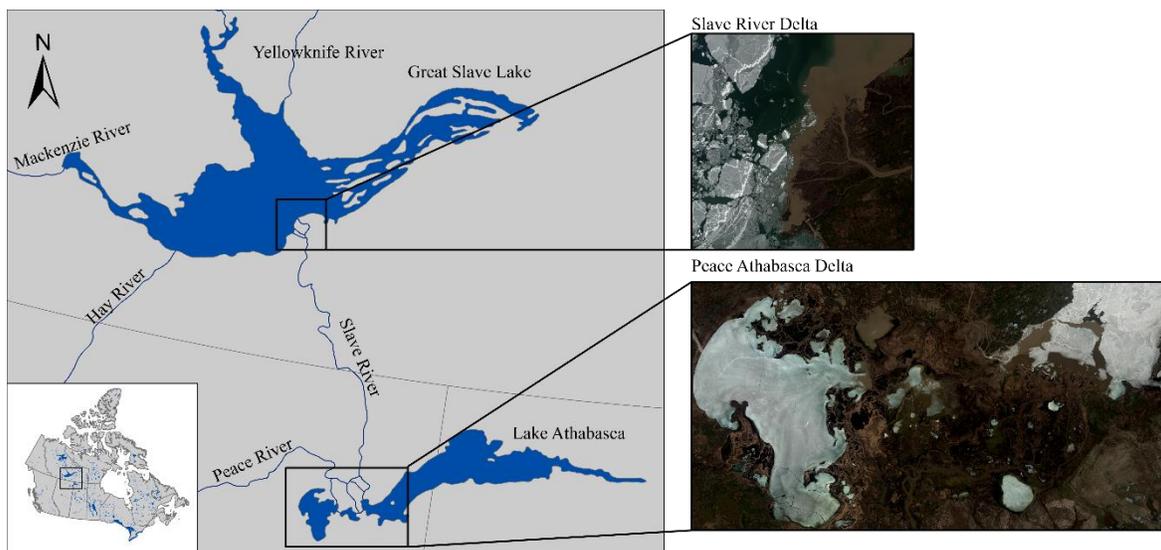


Figure 7. Site map of Slave river delta (SRD) and Peace Athabasca delta (PAD) in Northwest Territories and Alberta, Canada. The a) SRD and b) PAD images are acquired on 2018-05-23, and 2014-05-14, respectively.

2.4 Dataset and Preprocessing

Data from Landsat and Sentinel-2 satellites are used to develop RF models. Landsat data are selected as the main source of data starting from 1984 to 2013 with Landsat 5 and from 2013 to present with Landsat 8. Furthermore, Sentinel-2 imagery, including Sentinel-2a and Sentinel-2b are used to increase the temporal resolution from 2015 and 2017, respectively. All images from Landsat

and Sentinel have been cropped using a bounding box covering the SRD area and then masked using a shape file to remove the land pixels.

2.4.1 Landsat Archives

In this study, we used Landsat series of Earth observation data, including Landsat-5 with Thematic Mapper (TM) sensor (1984–2013) and Landsat-8 with Operational Land Imager (OLI) and Thermal Infrared (TM) Sensors (2013–present). Landsat archives have been downloaded from the United States Geological Survey (USGS). Although Landsat images have a temporal resolution of 16 days, due to the convergence of Landsat orbit at higher latitudes and the overlap of lateral swaths, the temporal resolution has been reduced to an 8-day repeat cycle at the SRD. Landsat (Collection 1 Level-1) bands are used to train the model and discriminate ice and water pixels, including visible, NIR bands, and SWIR bands, as well as the Quality Band (BQA) for cloud masking (Table 1). BQA provides cloud and cloud shadow classes at different confidence levels: high (67-100 %), medium (34-66 %) and low (0-33 %) (Appendix B). In this study, based on visual inspection, high-confidence level is selected.

2.4.2 Sentinel-2 Archives

Sentinel-2 data from Copernicus Open Access Hub, acquired by Multi-Spectral Instrument aboard the Sentinel-2a (2015–present) and Sentinel-2b (2017–present) satellites are used. Sentinel-2 images consist of 13 spectral bands (Table 1) and provide a revisit time of 10 days with one satellite (2015-present) and 5 days with two satellites (from 2017 onwards). The Sentinel RF model utilizes the Level1C bands that have been resampled to 20 meters to discriminate water, ice, and cloud pixels.

Table 1: Spatial resolution and central wavelength of Sentinel-2, Landsat-8, and Landsat-5 bands.

| Sentinel-2 | | | Landsat-8 | | | Landsat-5 | | |
|------------|--------------------------|-----------------------------|-----------|--------------------------|-----------------------------|-----------|--------------------------|-----------------------------|
| Band | Wavelength (micrometers) | Spatial Resolution (meters) | Band | Wavelength (micrometers) | Spatial Resolution (meters) | Band | Wavelength (micrometers) | Spatial Resolution (meters) |
| B1 | 0.443 | 60 | B1 | 0.43-0.45 | 30 | B1 | 0.45-0.52 | 30 |
| B2 | 0.490 | 10 | B2 | 0.45-0.51 | 30 | B2 | 0.52-0.60 | 30 |
| B3 | 0.560 | 10 | B3 | 0.53-0.59 | 30 | B3 | 0.63-0.69 | 30 |
| B4 | 0.665 | 10 | B4 | 0.64-0.67 | 30 | B4 | 0.76-0.90 | 30 |
| B5 | 0.705 | 20 | B5 | 0.85-0.88 | 30 | B5 | 1.55-1.75 | 30 |
| B6 | 0.740 | 20 | B6 | 1.57-1.65 | 30 | B6 | 10.40-12.5 | 120 |
| B7 | 0.783 | 20 | B7 | 2.11-2.29 | 30 | B7 | 2.08-2.35 | 30 |
| B8 | 0.842 | 10 | B8 | 0.50-0.68 | 15 | | | |
| B8a | 0.865 | 20 | B9 | 1.36-1.38 | 30 | | | |
| B9 | 0.940 | 60 | B10 | 10.6-11.19 | 100 | | | |
| B10 | 0.137 | 60 | B11 | 11.50-12.51 | 100 | | | |
| B11 | 0.161 | 20 | | | | | | |
| B12 | 0.219 | 20 | | | | | | |

2.5 Method

2.5.1 Random Forest Modelling

The RF method creates multiple decision trees, each trained on a bootstrapped sample of training data, and searches across a randomly selected subset of features to split nodes of trees. (Leo Breiman, 2001). By introducing randomness to the construction of each individual tree, the model gains the ability to address overfitting issues and remain resilient to outliers within the training data. The later one is especially critical as the training data is manually collected from optical imagery since light layers of cloud may not be seen by visualization (Yang et al., 2021). This differentiates RF classifier from other learning models, such as SVM, and gradient boosting trees for the purpose of this study. In terms of classification output, each tree has a vote, and the result is determined based on the majority votes of trees. Like other ensemble learning algorithms, RF can handle complex classification scenarios since multiple classifications ensures a convergence approach to pixel labeling and outweigh the error of a single classification (Belgiu and Drăguț, 2016). Therefore, RF

models are a suitable choice for the purpose of this study and are expected to provide accurate and reliable results.

2.5.1.1 Feature Selection

One of the fundamental outcomes of RF models is feature importance measures, including mean decrease accuracy or mean decrease Gini values. RF calculates feature rankings by removing each feature from the model in turn and comparing accuracies for the model trained with and without the feature. The more the model loses accuracy, the more important the feature is. However, it has been shown that the RF feature importance measurement can be affected and biased by within-predictor correlation (Kristin K. Nicodemus., 2011). Therefore, to assess the capability of Landsat and Sentinel bands in discriminating the classes, the average and standard deviation (STD) values for each band and each class was computed using training pixels. A greater difference of Average \pm STD values for a given band across the classes indicates its efficacy in distinguishing between the classes. All bands, excluding those with negligible discriminative potential, including SWIR1, SWIR2, and cirrus from Landsat bands, and water vapor from Sentinel bands have been selected. Additionally, Mahalanobis distance is calculated with a multiple class approach for the remaining features (Scott et al., 2013; Masnan et al., 2015) (Appendix D). Mahalanobis distance sorts the features based on their importance (Table 2).

The initial set of features is augmented by incorporating two indices to improve model performance, including the water ice classification index (WICI) and a texture-based feature, the local average gradient of the red band (Kévin Barbieux et al., 2018). The local average gradient of each pixel is defined as the mean of the gradient image in the $n \times n$ neighborhood pixels. The parameter 'n' is an odd positive integer, which has been set to 5. Although increasing the 'n' value improves class identification by including more spatial context, it can also introduce confusion at the borders

between ice and water. The local average gradient relies on the difference in intensity changes between water and ice. According to Barbieux et al. (2018), water pixels exhibit smoother changes in intensity, resulting in a lower local average gradient, whereas ice pixels exhibit high changes in intensity resulting in a high local average gradient. The red band for the local average gradient feature is selected since its STD values, derived from training pixels, were one of the highest in the ice class and the lowest in water class, compared to the other Landsat and Sentinel bands. In essence, while some other bands demonstrated distinction in the intensity changes between ice and water classes, the red band could more effectively represent this for our study region. WICI and the texture-based feature have been shown to improve classification results when included in feature vector (Kévin Barbieux et al., 2018). WICI plays a vital role in enabling the model to differentiate between shallow water and ice, which is accomplished by leveraging the uniformity of shallow water reflectance in the near and shortwave infrared spectral. Ice, on the other hand, shows noticeable deviations in reflectance between these two spectral ranges (Appendix A).

The final set of features selected for this study are summarized in (Table 3). Variance inflation factor (VIF) test with a threshold of 10 has been used to remove features that are correlated (Chatterjee and Simonoff, 2013; O'brien 2007).

Table 2. Sorted bands based on Mahalanobis distance.

| Sentinel bands | D_A^2 | Landsat bands | D_A^2 |
|-----------------------|---------------------------|----------------------|---------------------------|
| SWIR3 | 0.861 | ultra blue | 0.852 |
| ultra blue | 0.860 | blue | 0.851 |
| green | 0.853 | green | 0.843 |
| blue | 0.851 | NIR | 0.823 |
| red edge3 | 0.850 | red | 0.816 |
| NIR | 0.849 | thermal 1 | 0.779 |
| red edge2 | 0.842 | thermal 2 | 0.744 |
| red | 0.838 | | |
| red edge1 | 0.834 | | |
| narrow NIR | 0.830 | | |

Table 3: Sentinel and Landsat Model's Features. As BQA band is used to discriminate cloud, the Landsat model needs less features than Sentinel one.

| Sentinel model's Features | Landsat Model's Features |
|-------------------------------|-------------------------------|
| Local Average Gradient of Red | Local Average Gradient of Red |
| ultra blue | ultra blue |
| red | red |
| NIR | NIR |
| narrow NIR | WICI |
| WICI | thermal 1 |
| SWIR1 | |
| SWIR2 | |

2.5.1.2 Training and Modeling

RF models are sensitive to hyperparameters, such as the number of trees (Ntree) and the number of randomly selected variables to split the nodes (Mtry). Here, the Ntree and Mtry are tuned using the cross-validation method. The cross-validation method divides the training set into k folds (here k = 10). The model is trained on k-1 folds and tested on the remaining fold. The iteration continues until the model is tested at least once with each fold. Through this approach, an RF model was generated for each potential value of the hyperparameters, allowing us to identify the configurations that resulted in the highest accuracy (Appendix C).

In terms of training data, a hand-picked set of ten images, approximately 600,000 pixels, have been selected, from early to late break-up process to include different types of ice cover. All the images were first visually checked and those with less cloud cover are selected. To achieve the optimum sets of images, the training sets start with two images for each model and continue by increasing the number of images until increased images no longer yield improvements in accuracies. The models trained with five images had the same accuracies as six and seven. Then, training areas (pixels) are extracted by drawing polygons on the selected images (Figure 8). The cloud and cloud

shadow on training images of Landsat archives are masked using BQA layer before drawing the polygons to reduce the chance of including non-clear cloud pixels.

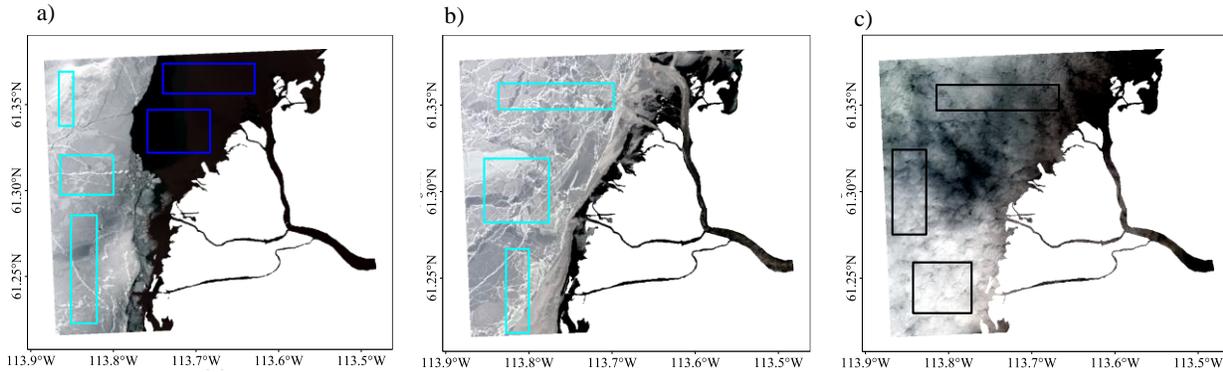


Figure 8. Examples of manually selecting training areas generated from a) Sentinel image captured on 2014-05-12 b) Landsat image captured on 2020-05-30, and c) Landsat image captured on 2019-05-10. The black, cyan, and blue colors correspond to cloud, ice, and water training areas. Training polygons with nearly equal contributions of ice and water pixels have even distribution over SRD.

2.5.2 Models Evaluation

In terms of the pixel-based evaluation, training and testing accuracies were calculated based on 30% of the data chosen as the training sets, which has not been seen by the model, and the remaining 70% as the testing dataset. In this regard, the balance between training and testing accuracies is monitored to avoid overfitting, which occurs when a model performs extremely well on the training data but poorly on the testing data. Image-based evaluation is performed by visual inspection between classified and RGB plots, and also calculated ice, water, and cloud percentages were cross-checked against the RGB plots to confirm their congruity. To assess generalizability of the model, it is further evaluated with independent scenes captured from the PAD break-up process.

2.5.3 SRD Trend Analysis

To identify the break-up dates, cloudy images with a cloud percentage of more than 20% are not considered. As mentioned earlier, while cloud pixels in Landsat imagery were masked using BQA layer, the cloud pixels in Sentinel imagery were distinguished during RF classification. The data-

driven threshold of 20% on cloud coverage is chosen based on the distribution of cloud percentages across all the images. As shown in Figure 9, most of the images have low (0 – 20%) or high (80 – 100%) cloud coverage. In other words, the data is sparse between the thresholds, hence increasing the threshold would not impact the analysis. The average and STD of cloud coverage across the remaining images, with less than 20% cloud pixels, are 2.17% and 3.44%, respectively.

The ice and water percentages for the remaining images (cloud cover of 0 – 20%) are calculated from non-cloud pixels (ice portion + water portion = 1). To define the optimum values of thresholds for break-up start, the images are divided into ten categories based on the fraction of ice cover for each image (Figure 10). Therefore, distribution of images could be effectively tracked by applying different thresholds on ice fraction. Given that Landsat images have a temporal resolution of 8 days on SRD, the ice minimum threshold is critical to exclude images captured after the break-up start period. Consequently, minimum and maximum ice cover thresholds of 60% and 90% were chosen as a trade-off to maximize data utilization while excluding images taken after the break-up start period. During the break-up period, the images have ice fraction of 60% to 90%, and the number of images during this period vary depending on temporal resolution of data. When multiple images are available for a given year, originating from either different or the same satellites, the average day of break-up is calculated. To enhance the accuracy of this estimation, the average values are weighted based on ice fractions. The weighted average compensates the effects of different temporal resolution of data and justifies the date to better represent the start of break-up. The overall workflow has been highlighted in Figure 11.

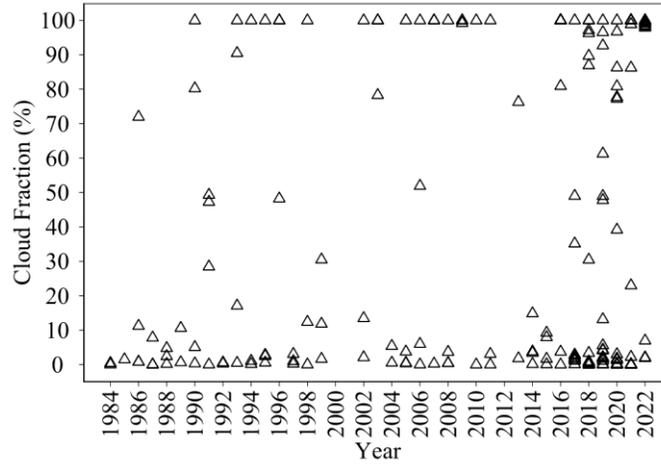


Figure 9. Distribution of cloud fractions of Landsat and Sentinel datasets from May 1 – May 30. Each open triangle corresponds to an individual image with a total number of 210. The cloud percentages are generated from the SRD boundary (The images are masked using the SRD shape file to exclude land pixels).

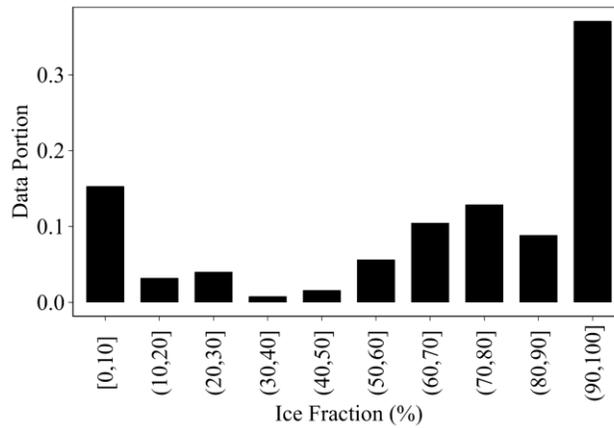


Figure 10. Data distribution provides insight into optimizing the thresholds of break-up identification. Data distribution provides insight into optimizing the thresholds of break-up identification.

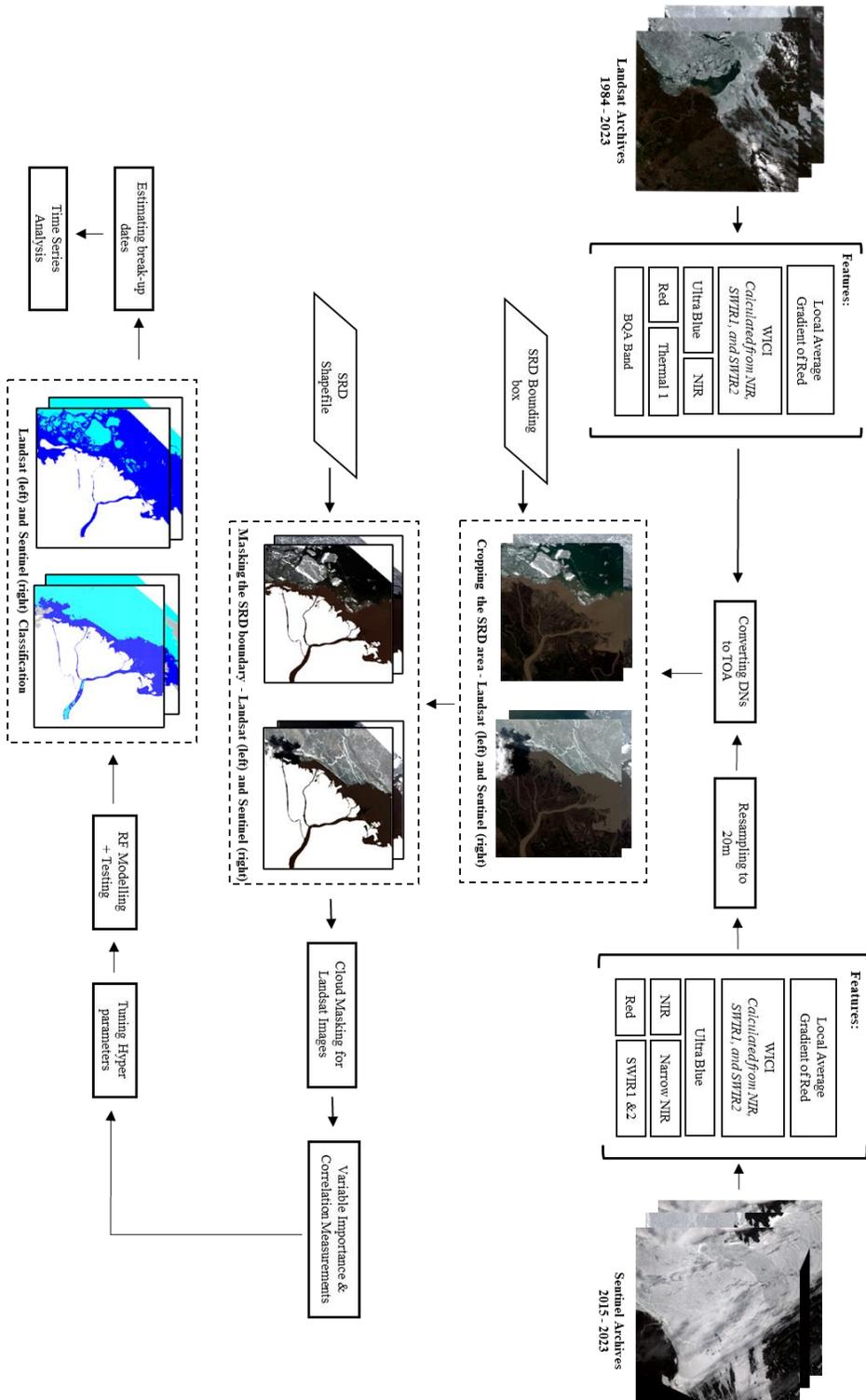


Figure 11. Workflow of RF modelling and Trend analysis.

2.6 Results

2.6.1 RF Models Validation

Table 3 presents the accuracies of the Landsat and Sentinel models. To address the concern of overfitting, the Landsat and Sentinel RF models are closely monitored by comparing the accuracies obtained from testing and training data. In terms of visual assessment, Figure 12 and Figure 13 provide examples of the classification results from Landsat and Sentinel models, respectively. Additionally, the performance of both models demonstrates strong agreement when images acquired on the same dates are evaluated (Figure 14).

Table 3. Accuracies of Landsat and Sentinel models for the SRD dataset. Landsat with two classes of ice and water outperforms Sentinel model with classes of ice, water, and cloud.

| Evaluation type | Landsat model | Sentinel model |
|-------------------|---------------|----------------|
| Training accuracy | 99.71% | 97.62% |
| Testing accuracy | 97.8% | 91.53% |

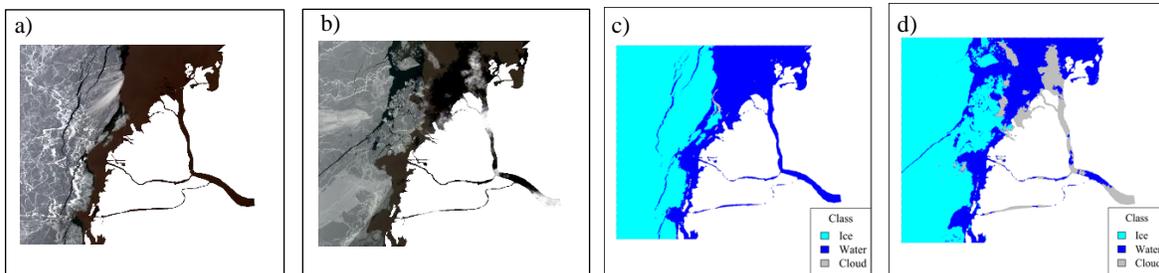


Figure 12. Examples of Landsat model performance with an overall accuracy of 97.8%. (a) RGB scene captured on 2013-05-25. (b) RGB scene captured on 2014-05-28. (c) and (d) Corresponding model classification plots.

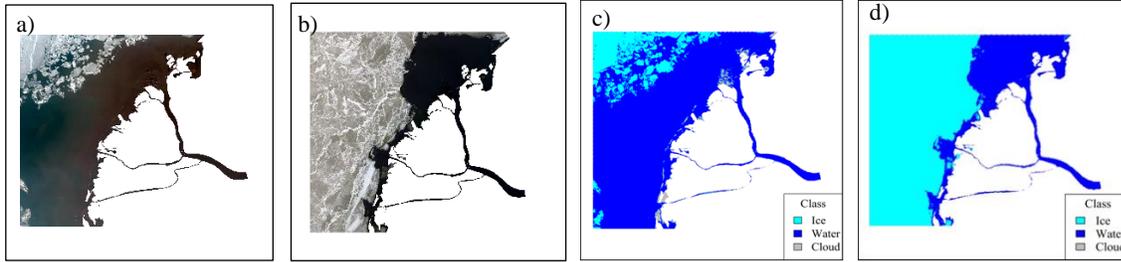


Figure 13. Examples of Sentinel model performance with an overall accuracy of 91.5%. (a) RGB scene captured on 2019-05-21. (b) RGB scene captured on 2021-05-27. (c) and (d) Corresponding plots.

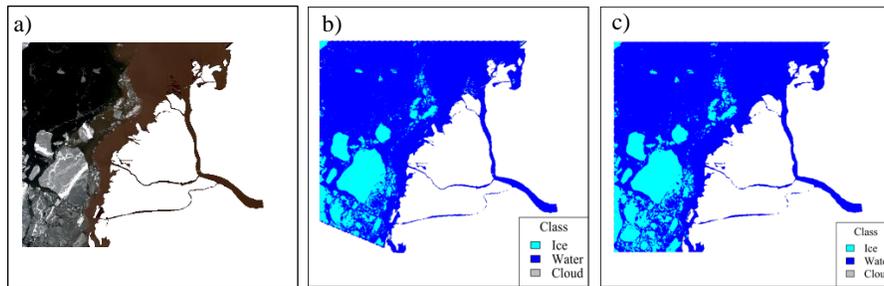


Figure 14. Evaluation of Landsat and Sentinel models performance by comparing their classified images captured on the same date (2018-05-23). (a) RGB plot captured by Landsat-8 satellite. (b) corresponding classified image (Sentinel-2b). (c) corresponding classified image (Landsat-8). The image shown here has mixed pixels of water and ice, mostly around the bottom left corner. Additionally, closer to SRD tributaries at the bottom, ice seems to be on top of water. In these cases, depending on ice portion or thickness and corresponding reflectance, the model determines the classes. In general, the lack of a clear boundary between ice and water areas makes the classification more complicated and requires more involved method than simple decision trees or threshold-based classifications.

2.6.1.1 Performance of model in a different area: Peace Athabasca Delta

The generalizability of the RF models was assessed and tested by evaluating their performance in a region different from where they were trained, chosen as the PAD, located in Alberta, Canada (58°42'N 111°30'W). PAD has three major geographic regions 1) Open-drainage lake sector composed of Lake Athabasca, Lake Claire, Mamawi Lake and Richardson Lake 2) Peace Delta and 3) Athabasca Delta. Figure 7 shows the location of Peace Athabasca Delta.

Both Landsat and Sentinel models, trained using SRD images, exhibited similar accuracies when applied to PAD. The testing images and corresponding polygons were manually selected using the same criteria as for SRD image selection process. The Landsat model achieved an accuracy of

95.34%, while the Sentinel model achieved 89.01% accuracy within the PAD. Consequently, the presented Landsat and Sentinel models can be effectively adapted to different study areas, providing robust accuracies. Figure 15 depicts the classification results within (PAD) using Landsat images.

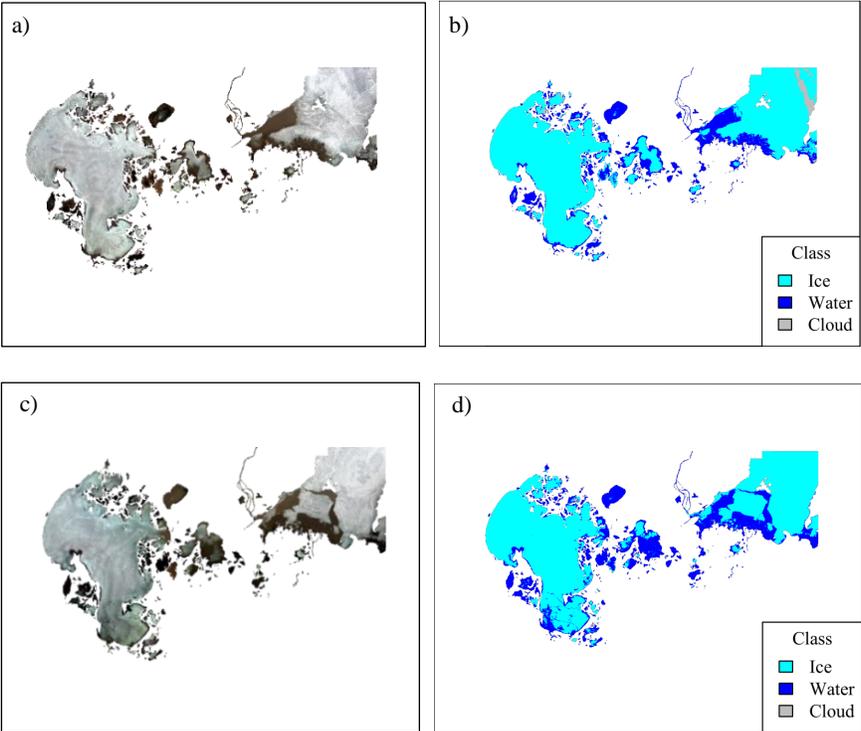


Figure 15. Examples of classification within PAD. RGB scene acquired on (a) 2013-05-11, and (b) 2014-05-14, and corresponding plots in (c) and (d) respectively.

2.6.2 Seasonal Dynamics of Ice and Water fractions

Analysis of non-cloudy data from 2018 to 2021 with higher temporal resolution than previous years provides insights into the varying rates of the break-up process, following the results of RF models. Figure 16 presents the water and ice fraction of non-cloudy images from 100% ice cover to 0%. A piecewise linear regression model has been fitted to water and ice fractions to better observe rate of changes (Tomé and Miranda, 2004). The data indicates fluctuation and non-monotonic patterns in the rate of break-up process. Figure 18 shows that a significant portion of ice break-up in 2018 and 2019 occurred rapidly. Based on RF classification results, transition from approximately 30% to 70%

water fraction took only two days. In contrast, the same transition took 15 days in 2021. Break-up rates can be influenced by changes in air temperature and Slave river discharge and faster rates reduce the chance of satellites to record the whole process of break-up.

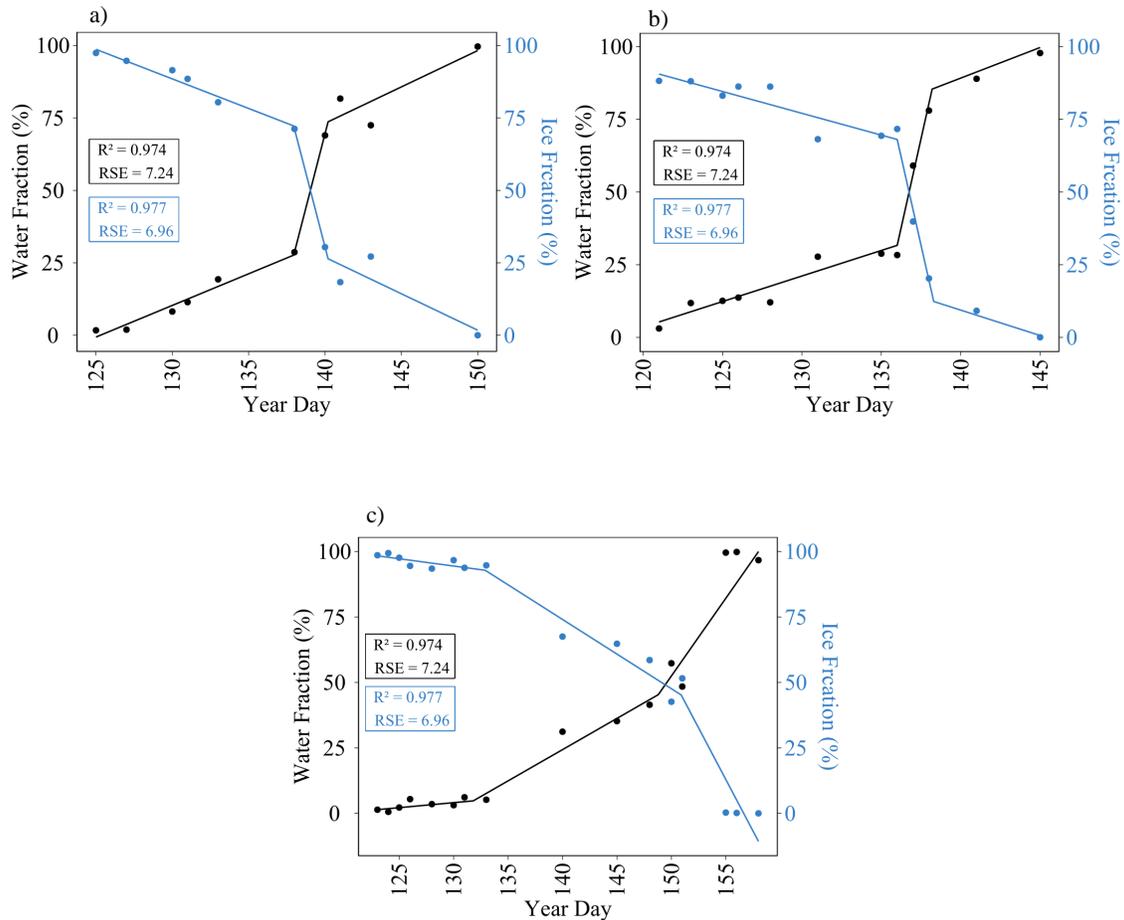


Figure 16. Rate of break-up occurrence in a) 2018, b)2019, and c) 2020. Years with data from Sentinel-2a and Sentinel-2b indicate high variability in break-up rate. Solid lines are used as a visual aid.

2.6.3 Trend of break-up onset

The earliest break-up event occurred on YD 126 in 2015, captured by an image with an ice fraction of 80.9%. Conversely, the latest break-up was observed on YD 152 in 1992. Table presents the break-up start dates from 1984 to 2023. Despite annual variations, a consistent declining trend in the break-

up process was identified. The Mann-Kendall test conducted from 1984 to 2023 yielded a p-value of 0.05, indicating a significant trend towards earlier break-up.

The observed trend aligns with other studies that have examined the break-up of GSL. In a study by Duguay et al. (2006), the temporal and spatial patterns of freeze-up and break-up were analyzed using in-situ data from the CID, covering three 30-year climatological periods (1951–1980, 1961–1990, and 1971–2000). The third period has 16 years overlap with the timeframe of our analysis and shows a significant earlier break-up trend from two stations and non-significant earlier trend from the other one. In general, the paper noted that the trends transitioned from significant and non-significant later break-up trend to significant and non-significant earlier trend. Similarly, our data from 1984 to 2000 showed a non-significant earlier break-up trend.

Table 5. Estimated the start of break-ups 1984 to 2023. Days in column two represents the start of break-up process for each year in column one.

| | | | | | | | | | | | | | | | | | | | | | | | |
|----------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Year | 1984 | 1986 | 1988 | 1989 | 1992 | 1993 | 1994 | 1995 | 1997 | 2005 | 2008 | 2011 | 2012 | 2013 | 2014 | 2015 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 |
| Year day | 137 | 146 | 144 | 143 | 152 | 141 | 148 | 143 | 149 | 138 | 148 | 140 | 145 | 145 | 148 | 126 | 132 | 134 | 129 | 142 | 145 | 144 | 129 |

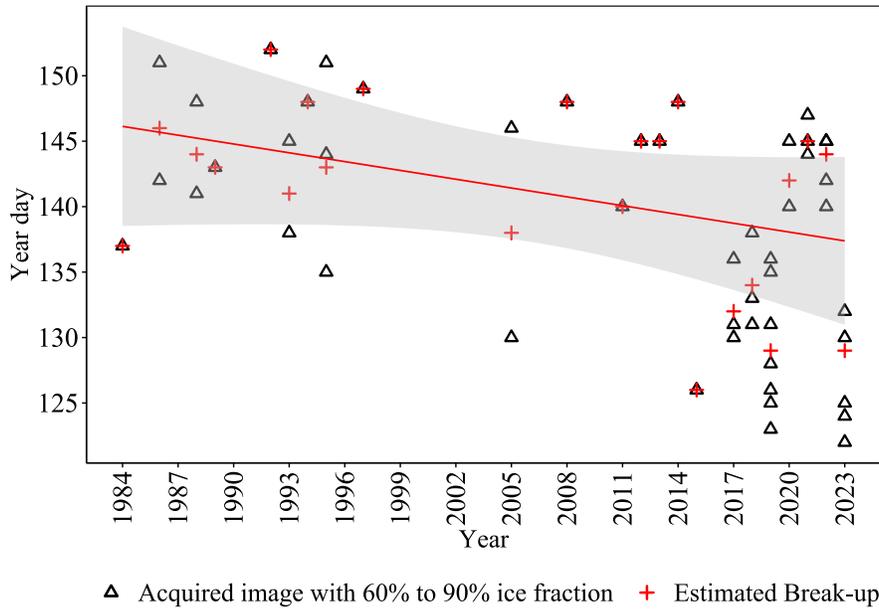


Figure 17. Black triangles indicate non-cloudy images with ice fraction of 60% to 90%, while red positive points are corresponding weighted averages or estimated days for the start of break-up. A linear model is fitted to estimated break-up starts with confidence level of 0.99, RSE of 11.98, and slope of -0.84.

2.7 Discussion

Figure 18 illustrates the temporal resolution of non-cloudy images acquired by Landsat and Sentinel satellites during the month of May. The data collection begins in 1984 with Landsat 5 and 8 satellites and continues to present with a better temporal resolution in recent years by Sentinel images. Lower temporal resolution in early years results in more challenges to estimate the start of break-up, and therefore the values of thresholds on ice fraction. Data unavailability is primarily attributed to cloud coverage over the SRD area and inevitable gaps between recorded data due to the satellites' temporal resolution. Consequently, while the data provides valuable insights into the break-up trend, the analysis is not conducted on a yearly basis (Table 5).

Different temporal resolution may affect the precision of break-up estimation. The better temporal resolution in recent years may result in a more precise estimation of the break-up start than earlier years. However, the difference in precision can fluctuate only within the duration between minimum

and maximum thresholds, accounting for 30% of the overall period. Furthermore, the estimation can be affected by break-up rate since it determines the number of images could be acquired within the transition of 60% to 90% of ice fraction.

To justify the estimated break-up start dates and consequently reduce the effects of different temporal resolution, the weighted average date within the transition of 60% to 90% of ice fraction is calculated based on ice portion. However, a significant earlier trend is also observed by considering the break-up date as the date of first acquired image during the time that ice fraction is between 60% and 90%.

In order to comprehensively assess the impact of varying temporal resolution on trend analysis, a separate trend analysis was conducted using a dataset comprised solely of Landsat images. Figure 19 illustrates the anomaly plots derived from two datasets: one incorporating Sentinel records and the other excluding them. The results obtained without Sentinel data exhibit a higher number of temporal gaps in break-up identification. Nevertheless, the break-up onset in 2015, 2017, 2022, 2023 was successfully identified, yielding the same estimation of break-up onset process in 2015, one day earlier in 2017, 2023, and one day later in 2022.

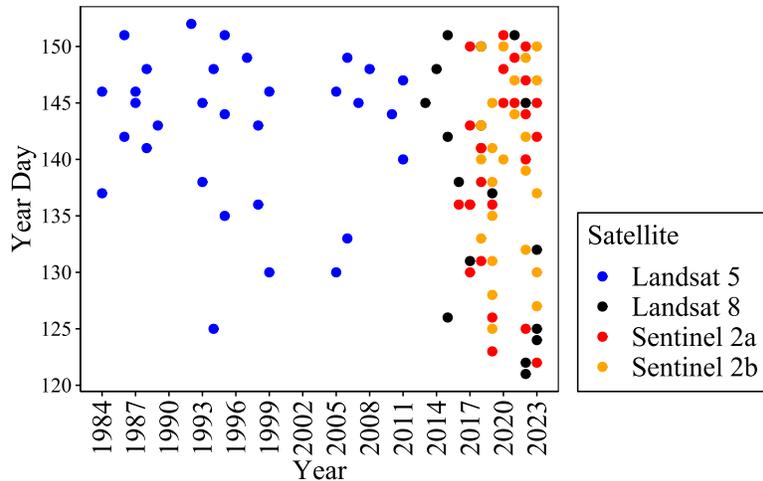


Figure 18. Images captured during the break-up period with less than 20% cloud fraction. Each colour corresponds to one satellite. Landsat 5 started in 1984 and continued until 2011. Landsat 8 starts afterward and still is active. Furthermore, to fill the temporal gaps in recent years, Sentinel data is used as supplementary data starting from 2015 and 2017.

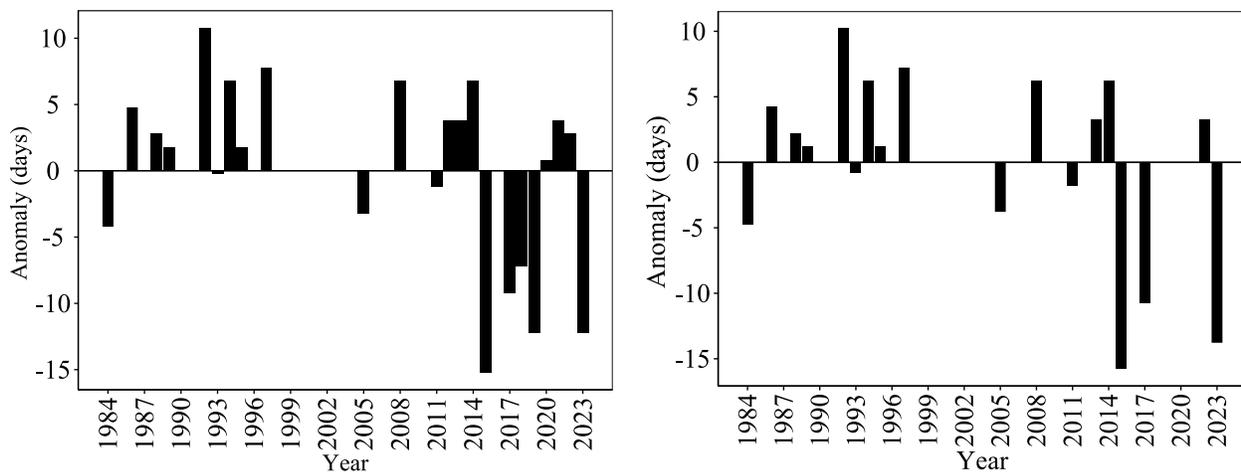


Figure 19. Break-up anomalies from 1984 to 2023 using Landsat archives (right) and combination of Landsat and Sentinel archives (left). Lower temporal resolution resulted in more temporal gaps, however, years of 2015, 2017, 2022, and 2023 could be identified successfully without Sentinel records and have close estimation of break-up to the results with Sentinel records.

2.8 Conclusion

In this study, the start of break-up process within SRD has been monitored by two presented RF models, using Landsat and Sentinel imagery. SRD served as a representative of GSL break-up process since the Slave river flow delivers energy and initiates the break-up process of lake at the SRD. Furthermore, GSL ice break-up is subject to upstream water activities in addition to climate change

indirectly by changes in Slave river discharge. Therefore, SRD is the first area of lake that can be studied to better understand the effects of upstream water management and water extraction on GSL ice break-up.

Landsat and Sentinel RF models have accuracies of 97.8% and 91.5% using independent scenes, separately, and are tested by images acquired from PAD, showing robustness for the models to be used in different study areas. The initiation of break-up process at SRD was identified based on ice and water portion, calculated from the result of RF classification. The results showed high variability in rate of break-up and significant trend toward earlier break-up from 1984 to present, following the result of Mann-Kendall test with p-value of 0.05.

The presented research is of great significance to northern communities as lake ice supports activities such as transportation, and fishing. Therefore, it is essential to better understand lake ice behavior against climate change and human activities. The findings can help the policymakers and resource managers in implementing adaptive strategies to ensure the preservation of reliable ice conditions and the sustainable use of natural resources.

Chapter 3. Conclusion and Future Work

This chapter summarizes the objectives of presented research and key findings, following by contribution made to the field. Furthermore, limitations encountered during the study, and potential directions for future research will be discussed.

3.1 Conclusion

The research presented in this thesis has monitored the SRD break-up trend over the last four decades, using ML techniques. Slave river flows to GSL and initiates the break-up process of GSL at SRD. The changes in Slave river flow, as a result of climate and non-climate factors, such as upstream water activities, strongly influences GSL long term temperature trend and ice phenology . Therefore, the SRD break-up trend can be served as an indicator to better understand GSL break-up process, and also the impact of upstream water activities on GSL ice break-up. In this study, two RF models have been developed to monitor break-up process in SRD, using Landsat-8, Landsat-5, Sentinel-2a, and Sentinel-2b imagery. Landsat imagery is the main source of data, starting from 1984 to present, and Sentinel-2 imagery is a supplementary data to cover the temporal gaps and increase the break-up retrieval accuracy. The RF models, tested by independent scenes, showed accuracies of 91.5% and 97.8%. Additionally, the generalizability of the RF models was assessed by evaluating their performance in PAD. The testing images and corresponding polygons in PAD were manually selected using the same criteria as for SRD image selection process. The Landsat model achieved an accuracy of 95.34%, while the Sentinel model achieved 89.01% accuracy within the PAD. Consequently, the presented Landsat and Sentinel models can be effectively adapted to different study areas, providing robust accuracies.

The start of break-up in each year has been identified using the ice and water portion, calculated from RF models. The threshold values on ice/water portion, which define the break-up time frame, are a compromise to maximize data utilization while effectively excluding images taken after break-up. The latter criteria play a critical role as the early years have only been covered by Landsat-5 with 8 days temporal resolution on SRD. In other words, depending on SRD break-up rate, the satellite may record the SRD state, when the break-up has already happened. Therefore, the maximum threshold on ice excludes those images and consequently those years that the break-up has been missed by the satellites. Accordingly, the monitoring is not yearly based, but presents a valuable perspective on historical trend of SRD break-up. The results show while SRD showed high variability in the rate of break-up, the result of two-tailed Mann-Kendall test shows a significant trend toward earlier break-up (p-value of 0.05). This research is of great significance for the northern communities as the earlier break-up on SRD and consequently on GSL can affect communities who rely on lake ice for transportation via ice roads, sustenance e.g. fish harvesting.

3.2 Future Works

As mentioned earlier, lakes and rivers are intricately linked in a unified system, and previous studies have focused on lakes and rivers separately (Palecki et al., 1986; Williams et al., 2006; Prowse et al., 2007; Mishra et al., 2011). A study by Ménard et al., (2003) simulates the lake ice phenology on GSL, using CLIMo model and validates the model against in-situ break-up/ freeze-up, ice thickness, on-ice snow depth measurements at three sites on GSL. They compared the observed break-up and freeze-up on Hay River site with and without considering river inflow. Simulated break-up dates show an almost constant difference from in-situ observations. This indicates that an external factor not well captured by the lake ice model influences break-up dates at the Hay River site. A plausible explanation is that the ice observation site is located at the mouth of Hay River. Indeed, a

river flowing into a lake increases the heat flux, which can lead to accelerated ice break-up during spring thaw (Bengtsson, 1986). This discussion opens a way for further improvement to develop regional and global lake-river modeling in GSL that presents lake and its main rivers, including Slave River in a unified system.

Additionally, a promising further investigation lies in tracking the time between the break-up of the PAD and the SRD, and monitoring how it has changed in recent years (Figure 7). Unfortunately, due to temporal gaps in existing data, it was not feasible to accurately track this time lag. Missing years in either delta's break-up identification results in NA (Not Available) rows in the time lag calculation, leading to an incomplete understanding. However, with the emergence of advanced satellite technology and the availability of more extended time series data from Sentinel-2, this monitoring task becomes increasingly achievable in the future.

Appendices

Appendix A. WICI and Local Average Gradient

A1. Water and Ice Classification Index (WICI)

WICI is one of features used in developing the RF models in this study and has been proposed by Barbieux et al. (2018). WICI, presented in equation 1, relies on the uniformity of water pixels in NIR and SWIR reflectance.

$$\text{WICI} = \frac{\sigma_{\rho_{\text{SWIR}1}} + \sigma_{\rho_{\text{SWIR}2}}}{2\sigma_{\rho_{\text{NIR}}}} \quad (1)$$

Where $\sigma_{\rho_{\text{SWIR}1}}$ and $\sigma_{\rho_{\text{NIR}}}$ are calculated standard deviations of SWIR and NIR values in a 5 by 5 window for each pixel. Barbieux et al., 2018 showed that WICI can help the models to achieve higher accuracy by enabling the model to differentiate between shallow water and ice, which is accomplished by leveraging the uniformity of shallow water reflectance in the near and shortwave infrared spectral. Ice, on the other hand, shows noticeable deviations in reflectance between these two spectral ranges.

A2. Local Average Gradient

The local average gradient is a texture-based feature used in developing the RF models in this study and has been proposed by Barbieux et al. (2018). The local average gradient presented in equation 2, is defined as the mean of the gradient image in the $n \times n$ neighborhood pixels. The parameter 'n' is an odd positive integer, which has been set to 5. As mentioned earlier, the local average gradient relies on the difference in textures of water and ice. According to Barbieux et al. (2018), water pixels exhibit smoother changes in intensity, resulting in a lower local average gradient, whereas ice pixels exhibit high changes in intensity resulting in a high local average gradient.

$$\text{Local average gradient} = \sqrt{\left(\frac{\delta \text{TOA}(i,j)}{\delta i}\right)^2 + \left(\frac{\delta \text{TOA}(i,j)}{\delta j}\right)^2} \quad (2)$$

Where $\frac{\delta TOA(i,j)}{\delta i}$ and $\frac{\delta TOA(i,j)}{\delta j}$ are the changes of TOA values in direction of x and y in a localized window.

Appendix B. Landsat Quality Assessment Layer

The presented Landsat RF model has two classes of water and ice, and the cloud pixels are masked out using Quality Assessment layer (BQA) of Landsat. BQA provides cloud and cloud shadow classes at different confidence levels: high (67-100 %), medium (34-66 %) and low (0-33 %). Table 4 and Table 5 display the attributes and pixel values from BQA and interpretations for Landsat 8 OLI and OLI/TIRS and Landsat 5 TM, respectively.

Table 4. Landsat 8 collection-1 and level-1 BQA band attributes and pixel value possibilities (From USGS).

| Attribute | Pixel values |
|------------------------------------|---|
| Fill | 1 |
| Dropped Pixel | 2, 674 |
| Clear | 672, 676, 680, 684 |
| Radiometric Saturation – 1-2 bands | 676, 708, 756, 932, 964, 1700, 1732 |
| Radiometric Saturation – 3-4 bands | 680, 712, 760, 936, 968, 1704, 1736 |
| Radiometric Saturation – 5+ bands | 684, 716, 764, 940, 972, 1708, 1740 |
| Cloud confidence-low | 672, 674, 676, 680, 684, 928, 932, 936, 940, 1696, 1700, 1704, 1708 |
| Cloud confidence - Medium | 704, 708, 712, 716, 960, 964, 968, 972, 1728, 1732, 1736, 1740 |
| Cloud confidence - High | 752, 756, 760, 764 |
| Cloud shadow - High | 928, 932, 936, 940, 960, 964, 968, 972 |
| Snow\ice - High | 1696, 1700, 1704, 1708, 1728, 1732, 1736, 1740 |

Table 5. Landsat 5 collection-1 and level-1 BQA band attributes and pixel value possibilities (From USGS).

| Attribute | Pixel values |
|------------------------------------|--|
| Fill | 1 |
| Terrain Occlusion | 2, 2722 |
| Clear | 2720, 2724, 2728, 2732 |
| Radiometric Saturation – 1-2 bands | 2724, 2756, 2804, 2980, 3012, 3748, 3780, 6820, 6852, 6900, 7076, 7108, 7844, 7876 |
| Radiometric Saturation – 3-4 bands | 2728, 2760, 2808, 2984, 3016, 3752, 3784, 6824, 6856, 6904, 7080, 7112, 7848, 7880 |
| Radiometric Saturation – 5+ bands | 2732, 2764, 2812, 2988, 3020, 3756, 3788, 6828, 6860, 6908, 7084, 7116, 7852, 7884 |
| Cloud confidence - Low | 2720, 2722, 2724, 2728, 2732, 2976, 2980, 2984, 2988, 3744, 3748, 3752, 3756, 6816, 6820, 6824, 6828, 7072, 7076, 7080, 7084, 7840, 7844, 7848, 7852 |
| Cloud confidence - Medium | 2752, 2756, 2760, 2764, 3008, 3012, 5016, 3020, 3776, 3780, 3784, 3788, 6848, 6852, 6856, 6860, 7104, 7108, 7112, 7116, 7872, 7876, 7880, 7884 |
| Cloud confidence - High | 2800, 2804, 2808, 2812, 6896, 6900, 6904, 6908 |
| Cloud shadow - High | 2976, 2980, 2984, 2988, 3008, 3012, 3016, 3020, 7072, 7076, 7080, 7084, 7104, 7108, 7112, 7116 |
| Snow\ice – High | 3744, 3748, 3752, 3756, 3776, 8760, 3784, 3788, 7840, 7844, 7848, 7852, 7872, 7876, 7880, 7884 |
| Cirrus Confidence – Low | 2720, 2722, 2724, 2728, 2732, 2752, 2756, 2760, 2764, 2800, 2804, 2804, 2808, 2812, 2976, 2980, 2984, 2988, 3008, 3012, 3016, 3020, 3744, 3748, 3752, 3756, 3780, 3784, 3788 |
| Cirrus Confidence – High | 6816, 6820, 6824, 6828, 6848, 6852, 6856, 6860, 6896, 6900, 6904, 6908, 7072, 7076, 7080, 7084, 7104, 7108, 7112, 7116, 7840, 7844, 7848, 7852, 7872, 7876, 7880, 7884 |

Appendix C. Configuration of Models Hyperparameters

To tune the hyperparameters of Mtry and Ntree, a RF model has been generated for each potential value of the hyperparameters. The configurations that resulted in the highest accuracies and stability have been chosen. presents the Landsat and Sentinel models accuracies with Ntree values ranging from 5 to 100, and Mtry values ranging from 2 to 6 for Sentinel and 2 to 4 for Landsat model.

The Landsat model tends to be stable at the Ntree of 50, and by increasing Mtry to 4, the model achieved ~ %0.5 more accuracy (Figure 20 a). Therefore, the final configuration is set to Ntree of 50 and Mtry of 4. The sentinel model, on the other hand, has one more class of cloud and a few more features, resulting in more fluctuations in accuracy across different configurations (Figure 20 b). The Sentinel model with higher value of Mtry tends to get stable in lower Ntree, except the Mtry of 6. Mtry of 6 and 8 shows almost same performance at Ntree of 50 and 100. However, the model with Mtry of 8 is more stable than the one with Mtry of 6. Therefore, to maximize accuracy and stability the model with Mtry of 8 and Ntree of 50 has been chosen.

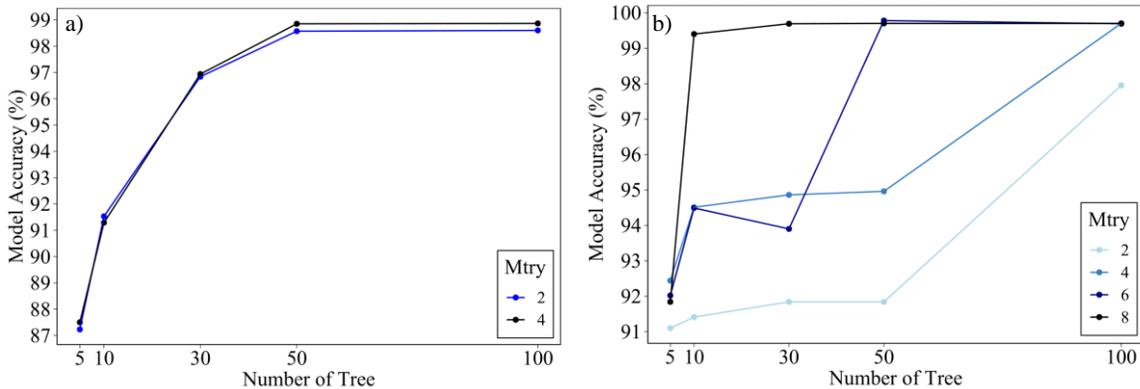


Figure 20. Performance of a) Landsat and b) Sentinel models in different configurations of Mtry and Ntree. A RF model has been generated for each potential value of hyperparameters, and its training accuracy is calculated using cross-validation method ($k = 10$).

Appendix D. Mahalanobis Distance for Feature Selection

Equation 3 presents Mahalanobis distance (Scott et al., 2013; Masnan et al., 2015).

$$D^2 = (\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2) \quad (1)$$

$$\bar{X}_1 = \frac{1}{N_1} \sum_{k=1}^{N_1} X_{1k} \quad (2)$$

$$\bar{X}_2 = \frac{1}{N_2} \sum_{k=1}^{N_2} X_{2k} \quad (3)$$

$$S = \frac{\sum_{k=1}^{N_1} (X_{1k} - \bar{X}_1)' (X_{1k} - \bar{X}_1) + \sum_{k=1}^{N_2} (X_{2k} - \bar{X}_2)' (X_{2k} - \bar{X}_2)}{N_1 + N_2 + 2} \quad (4)$$

$(X_{11}, X_{12}, X_{13}, \dots)$ and $(X_{21}, X_{22}, X_{23}, \dots)$ are pixel values of class C_i and C_j that each has N_1 and

N_2 training pixels, respectively. D_A^2 criterion for the class pair of (C_i, C_j) is given by:

$$D_{Aij}^2 = \frac{D_{ij}^2}{D_{ij}^2 + 4} \quad (5)$$

$$D_A^2 = \frac{1}{C_{(2,m)}} \sum_{i=1}^{m-1} \sum_{j=1}^m D_{ij}^2 \quad (6)$$

Where m is the number of classes, and $C_{(2,m)}$ is the combination calculator to choose 2 classes from m classes.

REFERENCES

- Abdi, H. (2010). Coefficient of variation. *Encyclopedia of research design*, 1(5).
- Adams, W. P. (1976). Field determination of the densities of lake ice sheets. *Limnology and Oceanography*, 21(4), 602-608.
- Adams, W. P., & Roulet, N. T. (1980). Illustration of the roles of snow in the evolution of the winter cover of a lake. *Arctic*, 100-116.
- Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., ... & Winder, M. (2009). Lakes as sentinels of climate change. *Limnology and oceanography*, 54(6part2), 2283-2297.
- Alfredsen, K., Haas, C., Tuhtan, J. A., & Zinke, P. (2018). Brief Communication: Mapping river ice using drones and structure from motion. *The Cryosphere*, 12(2), 627-633.
- Barbieux, K., Charitsi, A., & Merminod, B. (2018). Icy lakes extraction and water-ice classification using Landsat 8 OLI multispectral data. *International journal of remote sensing*, 39(11), 3646-3678.
- Barrette, P. D., & Charlebois, L. (2018). Winter roads and climate adaptation: prospective solutions through R&D. In *TAC 2018: Innovation and Technology: Evolving Transportation-2018 Conference and Exhibition of the Transportation Association of Canada*.
- Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, 5(1).
Doi:10.1038/sdata.2018.214
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing*, 114, 24-31.
- Beltaos, S., & Bonsal, B. (2021). Climate change impacts on Peace River ice thickness and implications to ice-jam flooding of Peace-Athabasca Delta, Canada. *Cold Regions Science and Technology*, 186, 103279.
- Bengtsson, L. (1986). Spatial variability of lake ice covers. *Geografiska Annaler: Series A, Physical Geography*, 68(1-2), 113-121.

Bonsal, B. R., Prowse, T. D., Duguay, C. R., & Lacroix, M. P. (2006). Impacts of large-scale teleconnections on freshwater-ice break/freeze-up dates over Canada. *Journal of Hydrology*, 330(1-2), 340-353.

Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32.

Brock, B. E., Yi, Y., Clogg-Wright, K. P., Edwards, T. W., & Wolfe, B. B. (2009). Multi-year landscape-scale assessment of lakewater balances in the Slave River Delta, NWT, using water isotope tracers. *Journal of Hydrology*, 379(1-2), 81-91.

Brown, L. C., & Duguay, C. R. (2010). The response and role of ice cover in lake-climate interactions. *Progress in physical geography*, 34(5), 671-704.

Brown, L. C., & Duguay, C. R. (2011). The fate of lake ice in the North American Arctic. *The Cryosphere*, 5(4), 869-892.

Brown, R. D., & Cote, P. (1992). Interannual variability of landfast ice thickness in the Canadian High Arctic, 1950-89. *Arctic*, 273-284.

Brown, R. D., Smith, C., Derksen, C., & Mudryk, L. (2021). Canadian in-situ snow cover trends for 1955–2017 including an assessment of the impact of automation. *Atmosphere-Ocean*, 59(2) 77–92. Doi: 10.1080/07055900.2021.1911781

Bush, E. and Lemmen, D.S., editors (2019): *Canada's Changing Climate Report*; Government of Canada, Ottawa, ON. 444 p.

Cai, Y., Duguay, C. R., & Ke, C. Q. (2022). A 41-year (1979–2019) passive-microwave-derived lake ice phenology data record of the Northern Hemisphere. *Earth System Science Data*, 14(7), 3329-3347.

Campbell, J. B., & Wynne, R. H. (2011). *Introduction to remote sensing*. Guilford press.

Cavaliere, E., Fournier, I. B., Hazuková, V., Rue, G. P., Sadro, S., Berger, S. A., ... & O'Reilly, C. M. (2021). The Lake Ice Continuum Concept: Influence of winter conditions on energy and ecosystem dynamics. *Journal of Geophysical Research: Biogeosciences*, 126(11), e2020JG006165.

Chaouch, N., Temimi, M., Romanov, P., Cabrera, R., et al. (2012). An automated algorithm for river ice monitoring over the Susquehanna River using the MODIS data. *Hydrological Processes*, doi:10.1002/hyp.9548.

Chaouch, N., Temimi, M., Romanov, P., Cabrera, R., McKillop, G., & Khanbilvardi, R. (2014). An automated algorithm for river ice monitoring over the Susquehanna River using the MODIS data. *Hydrological Processes*, 28(1), 62-73.

Chu, T., & Lindenschmidt, K. E. (2016). Integration of space-borne and air-borne data in monitoring river ice processes in the Slave River, Canada. *Remote Sensing of Environment*, 181, 65-81.

Chu, T., & Lindenschmidt, K. E. (2016). Integration of space-borne and air-borne data in monitoring river ice processes in the Slave River, Canada. *Remote Sensing of Environment*, 181, 65-81.

Commission for Environmental Cooperation (CEC). 2011. "North America Climate – Total Annual Precipitation". Museum of Vertebrate Zoology, University of California, Berkeley. Ed. 1.0, Raster digital data [1-km]. Available at <http://www.cec.org/north-american-environmental-atlas/precipitation-1950-2000-annual/>

Commission for Environmental Cooperation (CEC). 2011. "North America Climate – Mean Annual Temperature". Museum of Vertebrate Zoology, University of California, Berkeley. Ed. 1.0, Raster digital data [1-km]. Available at <http://www.cec.org/north-american-environmental-atlas/temperature-1950-2000-annual-mean/>

Cui, D., Liang, S., & Wang, D. (2021). Observed and projected changes in global climate zones based on Köppen Climate Classification. *WIREs Climate Change*, 12(3). Doi:10.1002/wcc.701

Dauginis, A. A., & Brown, L. C. (2021). Recent changes in pan-Arctic sea ice, lake ice, and snow-on/off timing. *The Cryosphere*, 15(10), 4781-4805.

Davidson, D. J., Williamson, T., & Parkins, J. R. (2003). Understanding climate change risk and vulnerability in northern forest-based communities. *Canadian Journal of Forest Research*, 33(11), 2252-2261.

De Coste, M., Li, Z., & Dibike, Y. (2022). Machine-learning approach for predicting the occurrence and timing of mid-winter ice breakups on canadian rivers. *Environmental Modelling & Software*, 152, 105402.

Denfeld, B. A., Baulch, H. M., del Giorgio, P. A., Hampton, S. E., & Karlsson, J. (2018). A synthesis of carbon dioxide and methane dynamics during the ice-covered period of northern lakes. *Limnology and Oceanography Letters*, 3(3), 117-131.

Dibike, Y., Prowse, T., Saloranta, T., & Ahmed, R. (2011). Response of Northern Hemisphere lake-ice cover and lake-water thermal structure patterns to a changing climate. *Hydrological Processes*, 25(19), 2942-2953.

Dörnhöfer, K., & Oppelt, N. (2016). Remote sensing for lake research and monitoring—Recent advances. *Ecological Indicators*, 64, 105-122.

Doxaran, D., Froidefond, J. M., Lavender, S., & Castaing, P. (2002). Spectral signature of highly turbid waters: Application with SPOT data to quantify suspended particulate matter concentrations. *Remote sensing of Environment*, 81(1), 149-161.

Du, J., Kimball, J. S., Duguay, C., Kim, Y., & Watts, J. D. (2017). Satellite microwave assessment of Northern Hemisphere lake ice phenology from 2002 to 2015. *The Cryosphere*, 11(1), 47-63.

Dubé, M. G., & Wilson, J. E. (2013). Accumulated state assessment of the Peace-Athabasca-Slave River system. *Integrated environmental assessment and management*, 9(3), 405-425.

Duguay, C. R., Flato, G. M., Jeffries, M. O., Ménard, P., Morris, K., & Rouse, W. R. (2003). Ice-cover variability on shallow lakes at high latitudes: model simulations and observations. *Hydrological Processes*, 17(17), 3465-3483.

Duguay, C. R., Flato, G. M., Jeffries, M. O., Ménard, P., Morris, K., & Rouse, W. R. (2003). Ice-cover variability on shallow lakes at high latitudes: Model simulations and observations. *Hydrological Processes*, 17(17), 3465–3483. Doi:10.1002/hyp.1394

Duguay, C. R., Prowse, T. D., Bonsal, B. R., Brown, R. D., Lacroix, M. P., & Ménard, P. (2006). Recent trends in Canadian lake ice cover. *Hydrological Processes: An International Journal*, 20(4), 781-801.

Ehrman, J., Clark, S., & Wall, A. (2021). Ice roughness estimation via remotely piloted aircraft and photogrammetry. *The Cryosphere*, 15(8), 4031-4046.

English, M. C., Hill, R. B., Stone, M. A., & Ormson, R. (1997). Geomorphological and botanical change on the outer Slave River Delta, NWT, before and after impoundment of the Peace River. *Hydrological Processes*, 11(13), 1707-1724.

English, S. J., Eyre, J. R., & Smith, J. A. (1999). A cloud-detection scheme for use with satellite sounding radiances in the context of data assimilation for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 125(559), 2359-2378.

Estilow, T. W., Young, A. H., & Robinson, D. A. (2015). A long-term northern hemisphere snow cover extent Data Record for climate studies and Monitoring. *Earth System Science Data*, 7(1), 137–142. Doi:10.5194/essd-7-137-2015

Evans, M. S., & Muir, D. C. (2016). Persistent organic contaminants in sediments and biota of Great Slave Lake, Canada: Slave River and long-range atmospheric source influences. *Journal of Great Lakes Research*, 42(2), 233-247.

Everaerts, J. (2008). The use of unmanned aerial vehicles (UAVs) for remote sensing and mapping. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(2008), 1187-1192.

Frazier, P. S., & Page, K. J. (2000). Water body detection and delineation with Landsat TM data. *Photogrammetric engineering and remote sensing*, 66(12), 1461-1468.

Fujisaki, A., Wang, J., Bai, X., Leshkevich, G., & Lofgren, B. (2013). Model-simulated interannual variability of Lake Erie ice cover, circulation, and thermal structure in response to atmospheric forcing, 2003–2012. *Journal of Geophysical Research: Oceans*, 118(9), 4286-4304.

Futter, M. N. (2003). Patterns and trends in Southern Ontario Lake ice phenology. *Environmental Monitoring and Assessment*, 88, 431-444.

García-Mora, T. J., Mas, J. F., & Hinkley, E. A. (2012). Land cover mapping applications with MODIS: a literature review. *International Journal of Digital Earth*, 5(1), 63-87.

Ghimire, B., Rogan, J., Galiano, V. R., Panday, P., & Neeti, N. (2012). An evaluation of bagging, boosting, and random forests for land-cover classification in Cape Cod, Massachusetts, USA. *GIScience & Remote Sensing*, 49(5), 623-643.

Girard, R., Lee, P. F., & James, K. (2002, June). The RADARSAT-2&3 topographic mission: an overview. In *IEEE International Geoscience and Remote Sensing Symposium* (Vol. 3, pp. 1477-1479). IEEE.

Government of Northwest Territories (GNWT). (2022). Winter Roads Average Open/Close Dates. Infrastructure, GNWT, May 11, 2022. Data Available: <https://www.inf.gov.nt.ca/en/services/highways-ferries-and-winter-roads/winter-roads-average-open-close-dates>

Groisman, P. Y., & Easterling, D. R. (1994). Variability and trends of total precipitation and snowfall over the United States and Canada. *Journal of Climate*, 7(1), 184–205.

Hall, D. K., & Riggs, G. A. (2007). Accuracy assessment of the MODIS snow products. *Hydrological Processes: An International Journal*, 21(12), 1534-1547.

Ham, J., Chen, Y., Crawford, M. M., & Ghosh, J. (2005). Investigation of the random forest framework for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3), 492-501.

Hampton, S. E., Galloway, A. W., Powers, S. M., Ozersky, T., Woo, K. H., Batt, R. D., ... & Xenopoulos, M. A. (2017). Ecology under lake ice. *Ecology letters*, 20(1), 98-111.

Heinilä, K., Mattila, O. P., Metsämäki, S., Väkevä, S., Luojus, K., Schwaizer, G., & Koponen, S. (2021). A novel method for detecting lake ice cover using optical satellite data. *International Journal of Applied Earth Observation and Geoinformation*, 104, 102566.

Hoekstra, M., Jiang, M., Clausi, D. A., & Duguay, C. (2020). Lake ice-water classification of RADARSAT-2 images by integrating IRGS Segmentation with pixel-based random forest labeling. *Remote Sensing*, 12(9), 1425.

Hollstein, A., Segl, K., Guanter, L., Brell, M., & Enesco, M. (2016). Ready-to-use methods for the detection of clouds, cirrus, snow, shadow, water and clear sky pixels in Sentinel-2 MSI images. *Remote Sensing*, 8(8), 666.

Hollstein, A., Segl, K., Guanter, L., Brell, M., & Enesco, M. (2016). Ready-to-use methods for the detection of clouds, cirrus, snow, shadow, water and clear sky pixels in Sentinel-2 MSI images. *Remote Sensing*, 8(8), 666.

Hollstein, A., Segl, K., Guanter, L., Brell, M., & Enesco, M. (2016). Ready-to-use methods for the detection of clouds, cirrus, snow, shadow, water and clear sky pixels in Sentinel-2 MSI images. *Remote Sensing*, 8(8), 666.

Huang, Y., Liu, H., Hinkel, K., Yu, B., Beck, R., & Wu, J. (2017). Analysis of thermal structure of arctic lakes at local and regional scales using in-situ and multirate Landsat-8 data. *Water Resources Research*, 53(11), 9642-9658.

Jansen, E., Overpeck, J., Briffa, K. R., Duplessy, J. C., Joos, F., Masson-Delmotte, V., ... & Miller, H. L. (2007). *Palaeoclimate*.

Jiang, W., Ni, Y., Pang, Z., Li, X., Ju, H., He, G., ... & Qin, X. (2021). An effective water body extraction method with new water index for sentinel-2 imagery. *Water*, 13(12), 1647.

Kalke, H., & Loewen, M. (2018). Support vector machine learning applied to digital images of river ice conditions. *Cold Regions Science and Technology*, 155, 225-236.

Kang, K. K., Duguay, C. R., & Howell, S. E. L. (2012). Estimating ice phenology on large northern lakes from AMSR-E: algorithm development and application to Great Bear Lake and Great Slave Lake, Canada. *The Cryosphere*, 6(2), 235-254.

Kheyrollah Pour, H., Duguay, C. R., Martynov, A., & Brown, L. C. (2012). Simulation of surface temperature and ice cover of large northern lakes with 1-D models: A comparison with MODIS satellite data and in-situ measurements. *Tellus A: Dynamic Meteorology and Oceanography*, 64(1), 17614.

Kheyrollah Pour, H., Duguay, C. R., Scott, K. A. and Kang, K. (2017). Improvement of lake ice thickness retrieval from MODIS satellite data using a thermodynamic model. *IEEE Transactions on Geoscience and Remote Sensing*, 55(10), 5956-5965 doi:10.1109/tgrs.2017.2718533.

Koenig, L., Martin, S., Studinger, M., & Sonntag, J. (2010). Polar airborne observations fill gap in satellite data. *Eos, Transactions American Geophysical Union*, 91(38), 333-334.

Latifovic, R., & Pouliot, D. (2007). Analysis of climate change impacts on lake ice phenology in Canada using the historical satellite data record. *Remote Sensing of Environment*, 106(4), 492-507.

Leigh, S., Wang, Z., & Clausi, D. A. (2013). Automated ice–water classification using dual polarization SAR satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(9), 5529-5539.

Lenormand, F., Duguay, C. R., & Gauthier, R. (2002). Development of a historical ice database for the study of climate change in Canada. *Hydrological Processes*, 16(18), 3707-3722.

Li, H., Li, H., Wang, J., & Hao, X. (2023). Revealing the river ice phenology on the Tibetan Plateau using Sentinel-2 and Landsat 8 overlapping orbit imagery. *Journal of Hydrology*, 619, 129285.

Liston, G. E., & Hall, D. K. (1995). An energy-balance model of lake-ice evolution. *Journal of Glaciology*, 41(138), 373-382.

Livingstone, D. M., Adrian, R., Arvola, L., Blenckner, T., Dokulil, M. T., Hari, R. E., ... & Weyhenmeyer, G. A. (2010). Regional and supra-regional coherence in limnological variables (pp. 311-337). Springer Netherlands.

Lynch, A. J., Taylor, W. W., & Smith, K. D. (2010). The influence of changing climate on the ecology and management of selected Laurentian Great Lakes fisheries. *Journal of Fish Biology*, 77(8), 1764-1782.

Magnuson, J. J., Robertson, D. M., Benson, B. J., Wynne, R. H., Livingstone, D. M., Arai, T., ... & Vuglinski, V. S. (2000). Historical trends in lake and river ice cover in the Northern Hemisphere. *Science*, 289(5485), 1743-1746.

Masnan, M. J., Mahat, N. I., Shakaff, A. Y. M., Abdullah, A. H., Zakaria, N. Z. I., Yusuf, N., ... & Aziz, A. H. A. (2015, May). Understanding Mahalanobis distance criterion for feature selection. In *AIP Conference Proceedings* (Vol. 1660, No. 1). AIP Publishing.

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International journal of remote sensing*, 17(7), 1425-1432.

Ménard, P., Duguay, C. R., Flato, G. M., & Rouse, W. R. (2002). Simulation of ice phenology on Great Slave Lake, Northwest Territories, Canada. *Hydrological Processes*, 16(18), 3691-3706.

Ménard, P., Duguay, C. R., Flato, G. M., & Rouse, W. R. (2002). Simulation of ice phenology on Great Slave Lake, Northwest Territories, Canada. *Hydrological Processes*, 16(18), 3691-3706.

Mishra, V., Cherkauer, K. A., Bowling, L. C., & Huber, M. (2011). Lake ice phenology of small lakes: Impacts of climate variability in the Great Lakes region. *Global and Planetary Change*, 76(3-4), 166-185.

Murfitt, J., & Duguay, C. R. (2021). 50 years of lake ice research from active microwave remote sensing: Progress and prospects. *Remote Sensing of Environment*, 264, 112616.

Murfitt, J., Brown, L. C., & Howell, S. E. (2018). Evaluating RADARSAT-2 for the monitoring of lake ice phenology events in mid-latitudes. *Remote Sensing*, 10(10), 1641.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & quantity*, 41, 673-690.

Oke, T. R. (2002). *Boundary layer climates*. Routledge.

Palecki, M. A., & Barry, R. G. (1986). Freeze-up and break-up of lakes as an index of temperature changes during the transition seasons: a case study for Finland. *Journal of Applied Meteorology and Climatology*, 25(7), 893-902.

Palecki, M. A., & Barry, R. G. (1986). Freeze-up and break-up of lakes as an index of temperature changes during the transition seasons: a case study for Finland. *Journal of Applied Meteorology and Climatology*, 25(7), 893-902.

Paltan, H., Dash, J., & Edwards, M. (2015). A refined mapping of Arctic lakes using Landsat imagery. *International Journal of Remote Sensing*, 36(23), 5970-5982.

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and earth system sciences*, 11(5), 1633-1644.

Pekel, J. F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633), 418-422.

Prowse, T. D. (2001). River-ice ecology. I: Hydrologic, geomorphic, and water-quality aspects. *Journal of Cold Regions Engineering*, 15(1), 1-16.

Prowse, T. D., & Beltaos, S. (2002). Climatic control of river-ice hydrology: a review. *Hydrological processes*, 16(4), 805-822.

Prowse, T. D., & Brown, K. (2010). Hydro-ecological effects of changing Arctic River and lake ice covers: a review. *Hydrology Research*, 41(6), 454-461.

Prowse, T. D., Bonsal, B. R., Duguay, C. R., & Lacroix, M. P. (2007). River-ice break-up/freeze-up: a review of climatic drivers, historical trends and future predictions. *Annals of Glaciology*, 46, 443-451.

Quinton, W. L., Hayashi, M., & Chasmer, L. E. (2011). Permafrost-thaw-induced land-cover change in the Canadian subarctic: Implications for water resources. *Hydrological processes*, 25(1), 152-158.

Rafat, A., Pour, H. K., Spence, C., Palmer, M. J., & MacLean, A. (2023). An analysis of ice growth and temperature dynamics in two Canadian subarctic lakes. *Cold Regions Science and Technology*, 210, 103808.

Robertson, D. M., Ragotzkie, R. A., & Magnuson, J. J. (1992). Lake ice records used to detect historical and future climatic changes. *Climatic Change*, 21, 407-427.

Rouse, W. R., Blanken, P. D., Bussi eres, N., Walker, A. E., Oswald, C. J., Schertzer, W. M., & Spence, C. (2008). An investigation of the thermal and energy balance regimes of Great Slave and Great Bear Lakes. *Journal of Hydrometeorology*, 9(6), 1318-1333.

Rouse, W. R., Douglas, M. S., Hecky, R. E., Hershey, A. E., Kling, G. W., Lesack, L., ... & Smol, J. P. (1997). Effects of climate change on the freshwaters of arctic and subarctic North America. *Hydrological processes*, 11(8), 873-902.

Saberi, N., Scott, K. A., & Duguay, C. (2022). Incorporating aleatoric uncertainties in lake ice mapping using RADARSAT-2 SAR images and CNNs. *Remote Sensing*, 14(3), 644.

Schertzer, W. M., Rouse, W. R., Blanken, P. D., Walker, A. E., Lam, D. C., & Le on, L. (2008). Interannual variability of the thermal components and bulk heat exchange of Great Slave Lake. *Cold Region Atmospheric and Hydrologic Studies. The Mackenzie GEWEX Experience: Volume 2: Hydrologic Processes*, 197-219.

Scott, K. A., Ashouri, Z., Buehner, M., Pogson, L., & Carrieres, T. (2013, April). Assimilation of SAR data in the marginal ice zone. In 2013 IEEE Radar Conference (RadarCon13) (pp. 1-5). IEEE.

Scott, K. A., Xu, L., & Pour, H. K. (2020). Retrieval of ice/water observations from synthetic aperture radar imagery for use in lake ice data assimilation. *Journal of Great Lakes Research*, 46(6), 1521-1532.

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American statistical association*, 63(324), 1379-1389.

Singh, A., Kalke, H., Loewen, M., & Ray, N. (2020). River ice segmentation with deep learning. *IEEE Transactions on Geoscience and Remote Sensing*, 58(11), 7570-7579.

Sirguey, P., Mathieu, R., Arnaud, Y., Khan, M. M., & Chanussot, J. (2008). Improving MODIS spatial resolution for snow mapping using wavelet fusion and ARSIS concept. *IEEE Geoscience and Remote Sensing Letters*, 5(1), 78-82.

Sola, D., & Scott, K. A. (2022). Efficient Shallow Network for River Ice Segmentation. *Remote Sensing*, 14(10), 2378.

Sturm, M., Holmgren, J., König, M., & Morris, K. (1997). The thermal conductivity of seasonal snow. *Journal of Glaciology*, 43(143), 26-41.

Surdu, C. M., Duguay, C. R., & Fernández Prieto, D. (2016). Evidence of recent changes in the ice regime of lakes in the Canadian High Arctic from spaceborne satellite observations. *The Cryosphere*, 10(3), 941-960.

Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y. A., & Rahman, A. (2020). Land-use land-cover classification by machine learning classifiers for satellite observations—A review. *Remote Sensing*, 12(7), 1135.

Tom, M., Kälin, U., Sütterlin, M., Baltsavias, E., & Schindler, K. (2018). Lake ice detection in low-resolution optical satellite images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4, 279-286.

Tomé, A. R., & Miranda, P. M. A. (2004). Piecewise linear fitting and trend changing points of climate parameters. *Geophysical Research Letters*, 31(2).

Vavrus, S. J., Wynne, R. H., & Foley, J. A. (1996). Measuring the sensitivity of southern Wisconsin lake ice to climate variations and lake depth using a numerical model. *Limnology and Oceanography*, 41(5), 822-831.

Verpoorter, C., Kutser, T., & Tranvik, L. (2012). Automated mapping of water bodies using Landsat multispectral data. *Limnology and Oceanography: Methods*, 10(12), 1037-1050.

Vincent, L. A., Zhang, X., Brown, R. D., Feng, Y., Mekis, E., Milewska, E. J., Wan, H., & Wang, X. L. (2015). Observed trends in Canada's climate and influence of low-frequency variability modes. *Journal of Climate*, 28(11), 4545–4560. Doi:10.1175/jcli-d-14-00697.1

Vincent, L.A., Zhang, X., Mekis, E., Wan, H. & Bush, E.J. (2018). Monitoring changes in Canada's climate: Trends in temperature and precipitation indices based on daily monitoring data; *Atmosphere-Ocean*, doi:10.1080/07055900.2018.1514579

Williams, G. P. (1965). Correlating freeze-up and break-up with weather conditions. *Canadian Geotechnical Journal*, 2(4), 313-326. Williams, S. G., & Stefan, H. G. (2006). Modeling of lake ice characteristics in North America using climate, geography, and lake bathymetry. *Journal of Cold Regions Engineering*, 20(4), 140-167.

Williamson, A. G., Banwell, A. F., Willis, I. C., & Arnold, N. S. (2018). Dual-satellite (Sentinel-2 and Landsat 8) remote sensing of supraglacial lakes in Greenland. *The Cryosphere*, 12(9), 3045-3065.

Wu, Y., Duguay, C. R., & Xu, L. (2021). Assessment of machine learning classifiers for global lake ice cover mapping from MODIS TOA reflectance data. *Remote Sensing of Environment*, 253, 112206.

Xu, H. (2006). Modification of 63ormalized difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International journal of remote sensing*, 27(14), 3025-3033.

Xu, Y., & Scott, K. A. (2017). Sea ice and open water classification of SAR imagery using CNN-based transfer learning. In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 3262-3265). IEEE.

Yan, P., He, F., Yang, Y., & Hu, F. (2020). Semi-supervised representation learning for remote sensing image classification based on generative adversarial networks. *IEEE Access*, 8, 54135-54144.

Yang, X., Pavelsky, T. M., Bendezu, L. P., & Zhang, S. (2021). Simple method to extract lake ice condition from landsat images. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-10.

Zhang, S., & Pavelsky, T. M. (2019). Remote sensing of lake ice phenology across a range of lakes sizes, ME, USA. *Remote Sensing*, 11(14), 1718.

Zhang, X., Wang, K., & Kirillin, G. (2021). An automatic method to detect lake ice phenology using MODIS daily temperature imagery. *Remote Sensing*, 13(14), 2711.