ANALYZING LAKE VARIABILITY IN A HIGHLY DYNAMIC AREA OF THE YUKON FLATS, ALASKA USING REMOTE SENSING

by

CATHARINE FLEMING

A THESIS

Presented to the Department of Geography,
the Department of Environmental Studies,
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

May 2024

An Abstract of the Thesis of

Catharine Fleming for the degree of Bachelor of Science
in the Department of Geography and the Department
of Environmental Studies to be taken June 2024

Title: Analyzing Lake Variability in a Highly Dynamic Area of the Yukon Flats, Alaska Using Remote Sensing

Approved: *Sarah Cooley, Ph.D.*
Primary Thesis Advisor

Sentinel-2 optical imagery was used to track changes in lake area in a highly dynamic area of the Yukon Flats, Alaska over 2019-2023. Specific questions this research addresses include: how spatially consistent is the variability in water area? Why are certain lakes far less seasonally variable than others? How consistent is the temporal and spatial variability in the water area from year to year? Seasonal lake area fluctuations follow a similar progression from year to year. The seasonal maxima in lake extent are very pronounced for smaller water bodies in 2021 (20% higher average maximum lake area, 40% lower average minimum lake area) than larger water bodies. On average 2019 had the highest mean lateral change along lake shorelines while 2023 had the lowest.

Acknowledgements

I am extremely grateful to my primary advisor Professor Sarah Cooley for her mentorship and guidance throughout this process.

I would like to extend my sincere thanks Professor Carol Paty and Professor Daphne Gallager for their support and guidance towards the honors college requirements.

Thank you to the graduate students of the Cryo-Hydro Lab for Earth Observation for their advice, input, and support.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Introduction

The Arctic-Boreal region contains the highest density of lakes on earth (Lehner and Doll, 2004, Verpoorter et al., 2014). This region is also home to permafrost, a thick subsurface layer of soil that remains frozen throughout the year. Arctic lakes allow carbon dioxide and methane gases that are trapped in frozen permafrost to be transported through groundwater to the lakes where they are released into the atmosphere. Understanding the variability in surface area of these lakes is essential for quantifying their greenhouse gas emissions. Yet, the distribution and size of these lakes remains an area of uncertainty in calculations of gas exchanges.

Understanding variability in Arctic lakes requires high resolution satellite imagery. Satellite imagery has multiple types of resolution, two of which are of particular importance for this study: spatial resolution and temporal resolution. Spatial resolution refers to the area on the ground that each pixel represents. Temporal resolution refers to the revisit time of the same location on Earth. Both are important in determining what level of analysis is appropriate and useful for a given satellite. Previous studies have used moderate spatial resolution imagery to understand global surface water, surface water in the Arctic, and the Yukon Flats specifically. The satellites used for such analyses lack the spatial resolution to identify small lakes and ponds that make up an important part of surface water. Such analyses also lack the temporal resolution necessary to examine the seasonality of lakes and ponds. New object-based methods for tracking individual lakes in the Arctic allow for the utilization of satellites like CubeSats with higher resolution. Previously unquantified changes in seasonal lake area variability have implications for conceptions of surface water dynamics in the Arctic.

While the Yukon is more extensively studied compared to other portions of the Arctic, we still lack an understanding of the lake-level changes in surface water throughout the season over multiple years. The goal of this project is to quantify subseasonal lake variability over 5 years in the Yukon Flats, Alaska. The primary aim of this research is to determine seasonal variability of lakes in the study region. Specific questions to investigate this include:

1. How spatially consistent is the variability in water area?
2. How consistent is the temporal and spatial variability in the water area from year to year?
3. Why are certain lakes far less seasonally variable than others?

Chapter 2: Background

Long-term lake area changes in the Arctic

Despite the high density of lakes in the Arctic, their dynamics remain poorly understood. Understanding fluctuations in global surface water is important for monitoring water resources, flood events, and human management systems. Yet, there are gaps in knowledge of the spatial and temporal dynamics of surface freshwater discharge and changes in storage globally (Alsdorf et al., 2007). This is in part due to the long standard in-situ gauge measurements being one dimensional and insufficient in complex wetlands and floodplains (Carroll and Loboda, 2017). In contrast, remote sensing data provides spatially explicit and temporally frequent observational data (Huang et al. 2018). Remote sensing has provided an avenue for monitoring surface water that has been integral in parsing dynamics in remote arctic regions. Increases in the quantity and types of remote sensors in the past twenty years have led to a rapid expansion of remote sensing analyses for monitoring surface water and flood inundation (Huang et al., 2018).

A key question motivating surface water research today is how lake area is changing at the decadal scale. Much of the work analyzing decadal-scale changes in lake area uses moderate spatial resolution satellite imagery from Landsat (30m) with lower temporal resolution (16-day repeat period). Such methods have been used to quantify changes in global surface water using the entirety of the Landsat archive 1985- (Pekel et al., 2016). Others have implemented Landsat-based methods of identifying lake area change in the Yukon Flats specifically (Chen et al., 2012, Jepsen et al., 2013). Jepsen et al., 2013 use 23 images from 1952-2010 and find that many lakes have undergone substantial changes due to potential mechanisms of changes in lateral groundwater flow, loss of lake water to permeable subpermafrost, and reduced snow meltwater inputs. Chen et al., 2012 use 17 images spanning 1984-2009 and find that most of the variation in

lake areas (80.7%) was explained by local water balance and mean air temperature since snowmelt. Carroll and Loboda, 2017 find that small water bodies have been shrinking whereas moderate to large water bodies have been increasing in size over the past 31 years.

More recent analysis finds that permafrost thaw is leading to widespread surface water decline, sooner than most models predict (Webb et al., 2022). This corroborates findings that models considering only gradual permafrost thaw are substantially underestimating carbon emissions from thawing permafrost (Turetsky et al., 2020). Lake districts located along the transition zone between continuous and discontinuous permafrost (such as the Yukon Flats) are among the most dynamic regions of lake change with dominating lake area loss (Nitze et al., 2018).

While analyses of long-term estimates of surface water extent are important, there are some limitations of the data used for them. Decadal water maps based on Landsat do not have enough temporal detail to determine if they represent actual long-term trends in water cycles (Caroll and Loboda, 2017). In addition, such analyses lack the spatial resolutions needed to detect ponds and track the seasonality of their areal extent (Mullen et al., 2023). Previous research at large scales have overlooked small lakes and ponds in favor of large lakes. Downing 2010 identifies that smaller water bodies dominate the areal extent of continental waters.

Remote sensing technologies have been implemented to monitor long-term changes in lake areas at regional and global scales. The satellites used for such analyses lack the spatial resolution to identify small lakes and ponds that make up an important part of surface water. Such analyses also lack the temporal resolution necessary to examine the seasonality of lakes and ponds. In the Arctic, these can be highly dynamic and integral in understanding lake area changes.

Seasonal lake dynamics

Patterns of lake and pond occurrence and fluctuations throughout the season are key parts of freshwater storage and dynamics. Previous work has been limited to the available lower spatial resolution and higher temporal resolution using data from the MODIS sensor to monitor larger lakes with high frequency (McCullough et al., 2012). This sensor has a much higher temporal resolution (average 2-day repeat period) but lower spatial resolution (250-1000m). Finger-Higgens 2022 identifies that both temporal and spatial scales impact the magnitude and direction of change in dynamics of Arctic surface hydrology, underscoring the importance of high-resolution data. Further complicating the study of small lakes and ponds is the variability of seasonal water surfaces with strong variability interannually (moving between wet and dry or shifting geographically) (Pekel et al., 2016). Recent technological developments have led to satellites with higher spatial and temporal resolution (ex. Sentinel-2, CubeSat). Such satellites have recently been utilized in studies of seasonal lake dynamics.

The implementation of high temporal, high spatial resolution satellite imagery in mapping Arctic surface water has revealed previously unknown patterns. For example, Cooley et al., 2017 uses the Planet CubeSat constellation to analyze seasonal lake variability in 2016, finding drivers in surface water extent in the Yukon Flats to include the seasonal runoff cycle, connection with the Yukon River, and localized physiographic conditions. Further works have expanded upon the methods used in this paper to look at lakes across the Arctic (Cooley et al. 2019, Mullen et al. 2023). There is a general seasonal decline in lake area, but with localized exceptions (due to wetland flooding and vegetation on lake surface) (Cooley et al., 2019). Seasonal changes in pond area are exceptionally pronounced (Mullen et al., 2023). These

previously unquantified changes in seasonal lake area variability have implications for model estimations of greenhouse gas emissions in the Arctic.

Connection to carbon cycle

Inland waters make up a significant part of the global carbon cycle and the abundance and spatial distribution of lakes will continue to alter with global climate change (Tranvik et al., 2009). As permafrost thaws, it releases carbon dioxide and methane. Arctic lakes play a key role in this as the gases are transported through groundwater to the lakes from which they are released to the atmosphere. These processes are important for modeling climate change. Much work has been done to understand how lake dynamics will change with increasing permafrost thaw. Smith et al., 2005 identifies that high-latitude warming of permafrost triggers an initial but transitory phase of lake and wetland expansion, followed by their widespread disappearance. Increasing surface water and groundwater connectivity from thaw will influence exchange of moisture with the atmosphere (Bring et al., 2016). Thermokarst lakes are widespread in arctic permafrost lowlands with ice-rich sediments and the formation of them is a dominant mode of permafrost degradation (Grosse 2013). Such lakes are one of the landforms that result from the process of the thawing of ice-rich permafrost or the melting of massive ice (van Everdingen, 1998). Thermokarst lakes are important in understanding the permafrost-related carbon budget including the potential release of carbon via lake expansion or sequestration as peat in drained lake basins. (Jones et al., 2011).

Small lakes and ponds contribute to the carbon cycle. Holgerson and Raymond, 2016 find that very small ponds are only 8.6% of lakes and ponds by area globally but account for 15.1% of carbon dioxide emissions and 40.6% of diffusive methane emissions. Arctic and Boral regions have experienced warming rates twice the global average which has implications for increasing

methane emissions from permafrost degradation (Elder et al., 2020). This warming is expected to change the distribution of permafrost, soil moisture, connectivity of inland waters, and streamflow seasonality (Walvoord et al., 2016).

While the role small lakes and ponds play in the carbon cycle is considerable, global scale methane emissions from these lakes remain poorly understood. The global size distributions of lakes and ponds are a source of uncertainty in calculations of gas exchange (Bastviken et al., 2004, Holgerson and Raymond, 2016). Despite the limitations of knowledge in terms of lake distribution, open water methane emissions are predictable and emissions per lake is mainly related to lake area (Bastviken et al., 2004). The Yukon Flats is among the regions in Alaska with estimates of carbon fluxes the most vulnerable to future climate change, because of the heightened sensitivity of arctic and boreal ecosystems to intensified warming (Stackpoole et al., 2017). The dynamics of small lakes and ponds in Arctic systems are an important input to the of modeling methane emissions.

Chapter 3: Study Area

Located in the north-central part of Alaska, the Yukon Flats are a dynamic wetland system. The region is covered by discontinuous permafrost. The climate in this region is characterized by extremely cold winters and warm dry summers with a mean annual air temperature of ~6C. Annual precipitation is low (mean of ~250mm/year) (Walvoord 2012). The study area of this project is a small sub-area (roughly 4,850 km^2) within the Yukon Flats, bounded to the south by the Yukon River, to the west by the Chandalar River, and to the east by the track of the satellite used. The bounds of the area of interest (AOI) include a southwest corner: 66°31'32.2"N 146°36'40.3"W and a Northeast corner: 67°03'08.0"N 144°44'10.9"W (Figure 1). The study area includes approximately 2176 lakes ranging in area from 0.001 km^2 to 6.14 km^2.

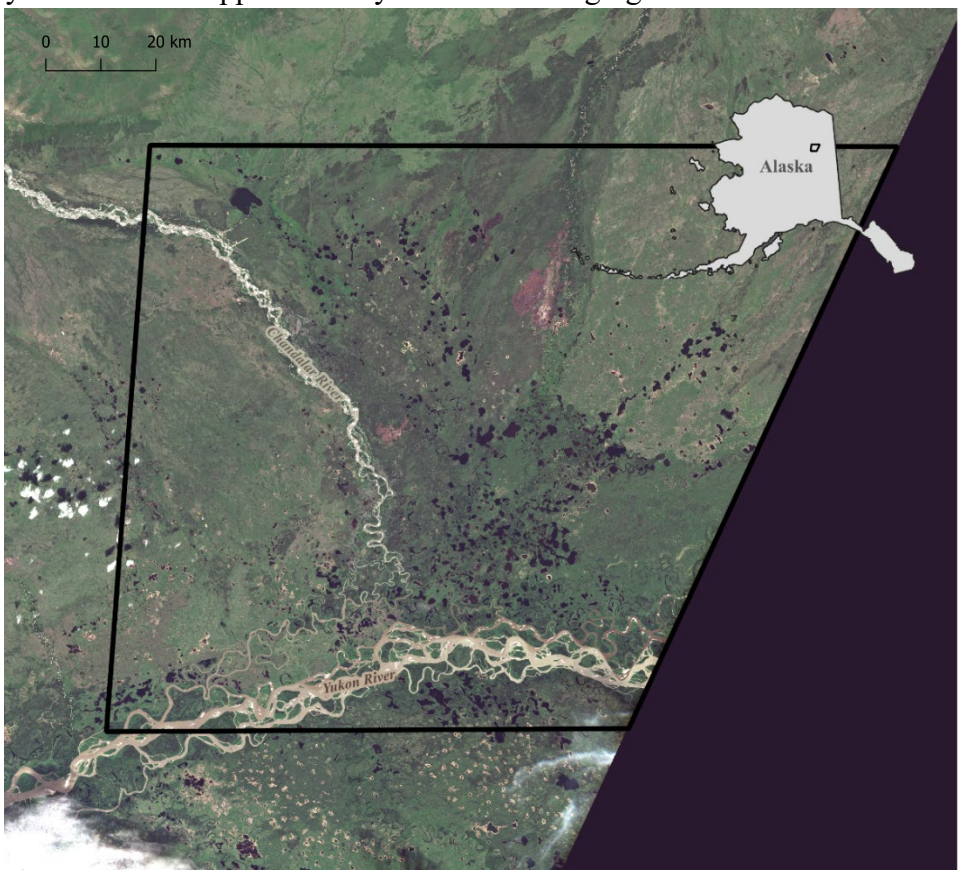

Figure 1. Map of Study Area. This map shows the study area outlined in black. The inset map of the state shows an outline of the site location. The Chandalar River is on the left bound of the study area and the Yukon River is to the south.

Chapter 4: Methods

Datasets

The satellite imagery used in this analysis is from the European Space Agency's Copernicus program's Sentinel-2 mission. This mission is a constellation of two satellites (Sentinel-2A and Sentinel-2B). The combination of these two satellites results in a temporal resolution of approximately five days (though the true observation frequency is often reduced by cloud cover). The sensors are pushbroom instruments with a 290km swath width that operate at an orbit altitude of 786km. These satellites are multispectral instruments with 13 spectral bands that capture different parts of the electromagnetic spectrum. Four of these bands have 10-meter spatial resolution and are most useful to this analysis—the red, blue, green, and near-infrared (NIR) bands (Table 1).

Band	*Central Wavelength (nm)*
Band 2 (Blue)	493
Band 3 (Green)	560
Band 4 (Red)	665
Band 8 (Near Infrared)	833

Table 1. 10-meter Sentinel-2 bands and their central wavelengths. Each band corresponds to a part of the electromagnetic spectrum and a central wavelength in nanometers is given. https://sentinels.copernicus.eu/web/sentinel/user-guides/sentine2-msi/resolutions/spectral

Sentinel-2 imagery comes in different levels of processing for different purposes. For this analysis, I use the Level-2A product that is atmospherically corrected. The Level-2A product, generated by Copernicus, is an atmospherically corrected surface reflectance computed from the Level-1C top-of-atmosphere reflectance. This process is crucial because top-of-atmosphere reflectance is impacted by aerosols and other particles in the atmosphere changing the values. There are atmospheric corrections applied to account for interference from particles in the

atmosphere. To access this data, I used the ESA's Copernicus dataspace browser. In the browser I downloaded all imagery from ice break up (mid-May) through refreezing (late September) for the years 2019-2023 that had at least some portion of the region of interest visible. Dates with high cloud coverage (over 70% of the scene) or those determined unusable based on visual inspection were excluded. The Level-2A data also comes with a 20-meter resolution cloud quality indicator created as a part of the Level-2A product generation.

Lake shapefile

The lakes included in this analysis are from the new Alaska Lake and Pond Occurrence Dataset (ALPOD) (Levenson, 2024). ALPOD was created through training a neural network model to identify lake areas (Mullen et al., 2023) in Sentinel-2 imagery. For this study, I clipped the shapefile to the AOI. To prepare the lake shapefile data I first buffered the lakes to 60 meters to allow for tracking of lake variability and then combine overlapping buffered lakes into single lake features to avoid double counting of lake area.

Data preprocessing

To prepare the satellite data for analysis, first I translated the necessary bands from .jp2 format to a GeoTIFF format. Additionally, I mosaiced image scenes from the same date into composites that covered the study area. To do this I used Rasterio's merge function taking the higher band values of overlapping portions. The Level-2A imagery is stored as a digital number with the conversion to surface reflectance values:

$$SR = \frac{DN + BOA\_offset}{10{,}000}$$

Because of extremely low values in the NIR band, dividing by 10,000 led to values that did not make sense. I determined that the most likely cause of this was the datatype the values were being stored in. Because I use only a ratio between values, scaling the surface reflectance

by 10,000 does not alter the relationship. Accordingly, I did not divide the values by 10,000. The BOA_offset value was added to this conversion in Jan. 2022, and I applied it to the 2022 and 2023 data.

Cloud masking

Cloud cover in satellite imagery can block the study area making the data unusable. Images dominated by cloud cover were excluded from this analysis entirely, but due to the relatively frequent occurrence of clouds in this region some imagery with clouds was used. When an image has some cloud cover it is important to mask out any pixels containing cloud.

To apply the cloud quality indicator to the imagery, I first resampled the Level-2A cloud mask product from 20 meters to 10 meters to match the pixel size of the bands that are used in this analysis. This was done using a bilinear resampling method. The cloud quality indicator product contains pixels ranging from 0 and 100 representing the probability of the pixel being cloud. Based on visual inspection of several thresholds for cloud probability I determined that the mask was quite conservative and any pixel with a cloud probability greater than 0% should be considered cloud. I reclassified this indicator to a binary raster in which 0 values represent cloud free pixels and values of 1 are cloud pixels. I then applied the cloud mask to the bands that would be used in the analysis, making pixel values in the bands that were cloud in the cloud mask null.

Water classification

In satellite imagery, distinguishing water from land is not always easy. To make these contrasts clearer, the differences in the spectral reflectance curves between water and other things including soil and vegetation can be enhanced. Water peaks in its reflectance in green wavelengths and has almost no reflectance in the near-infrared wavelengths (Figure 2.). To enhance these differences, I used a normalized difference water index (NDWI) (McFeeters, 1996). This index uses a ratio of the green band and near-infrared bands in the imagery to make water stand out. The specific calculations for this index are:

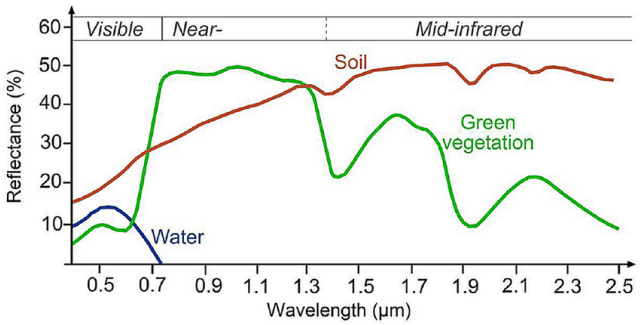

Figure 2. Spectral Reflectance Curves. Spectral reflectance curves of water (blue), soil (red), vegetation (green) across the electromagnetic spectrum. Source: Lillesand T, Kiefer RW, Chipman J (2015) Remote sensing and image interpretation. Wiley, New Jersey (https://www.researchgate.net/figure/Spectral-signatures-for-water-soil-and-vegetation-Lillesand-etal-2015_fig1_342827236)

$$\frac{band\ 3\ (green) - band\ 8\ (NIR)}{band\ 3\ (green) - band\ 8\ (NIR)}$$

All pixels are then turned into a value from -1 to 1 where higher values depict water. After calculating the NDWI I determined a threshold value above which pixels are considered water. The histogram of NDWI values generally depicts two peaks—one of land values that is lower and one of water values that is higher (Figure 3).

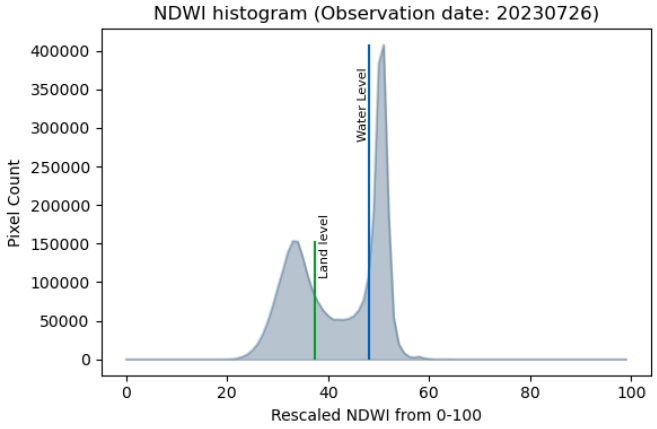

Figure 3. Example NDWI histogram of July 26th, 2023. Values are rescaled from -1 to 1 to 0 to 100. The lower peak on the left is land pixels. The high peak on the right is water pixels. The land level is shown in green. The water level is shown in blue.

For water classification I followed the water fraction approach of Cooley 2017 (originally adapted from Olthof et al., 2015). This

non-binary water classification is well-suited for the complex floodplains of the Yukon Flats with high amounts of inundated vegetation. This threshold uses the characteristic bimodal water histogram to determine a land level (LL) and a water level (WL) (Shown on Figure 3). To identify peaks, I used Sci.Py package and the find_peaks, peak_prominences, and peak_widths modules. To delineate the land level and water levels I found the 70% peak prominence of each peak. The 70% prominence of the right side of the land peak is the land level and the 70% prominence of the water peak on the left side is the water level. Initially prominence values of 90% were used following Cooley et al. 2017, but due to the extreme pointiness of the peaks, I shrunk it to 70%. This makes the classification more conservative. The 70% threshold worked for almost all NDWI histograms.

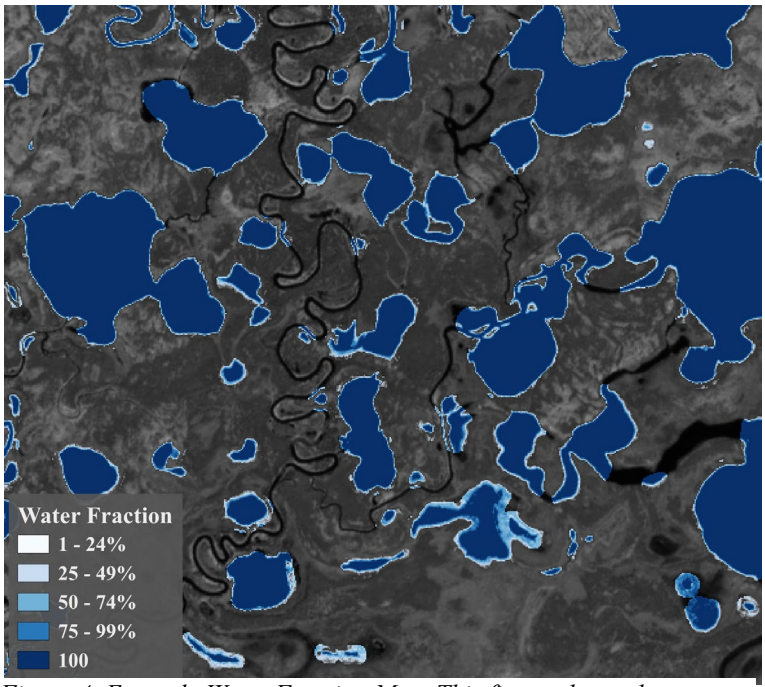

Figure 4. Example Water Fraction Map. This figure shows the water fraction classification for June 24th, 2023. Dark blue values are classified as 100% water. Values classified as land (0% water) are excluded from the map.

One area of complication with this analysis was images that did not show a characteristic two peak NDWI histogram. Such anomalies had clear land peaks but a much more sprawling water area without a clear threshold. Such images were excluded from further analyses.

Values that fall below the land level are classified as 100 percent land. Values above the water level are classified as 100 percent water. For values that fall above the land level and

below the water level I determined a water fraction. The water fraction (WF) values were determined using the equation:

$$WF = 100 \times \frac{NDWI - LL}{WL - LL}$$

I then reclassified the water fraction maps, so that water fraction values of 70% or greater were considered water and values below that threshold were considered not water (Figure 4). Then for each lake, I calculated the sum of the count of water pixels for each observation date. To do this I used the rasterstats zonal_stats module. The pixel sum for each lake was multiplied by 100 (due to the 10x10 meter pixel size) to get the area of each lake that is water in square meters.

Filtering

I then performed additional filtering to remove lakes for which there were not enough observations or that had falsely reduced area due to masked out cloud pixels. While cloud pixels were filtered out and excluded from NDWI calculations, the gaps left in portions of lakes covered by clouds could lead to undercounting the lake area. Therefore, I needed a process to eliminate lakes that would be underestimated in area. I determined a total pixel count of each lake and a sum of cloud pixels within the lake area. I then divided the number of cloud pixels by the count of pixels in the lake to determine the portion of the lake covered by cloud. If this exceeded ten percent the lake feature was removed. Additionally, for each year, lakes missing more than 10 percent of observations were removed from the dataset.

Change metrics

To identify patterns in changes in lakes I needed to standardize them by area. Larger lakes are going to have larger changes in area. To account for this, I looked at metrics of relative change in terms of lake area throughout the year. I calculated the percent difference from the

mean area for each lake for each observation. To calculate this metric, I first found the mean area value for each lake. I then divided the area value of each observation by this value to get a percent difference from the mean. To smooth the data, I performed a median filtering using a 3-observation window on all lakes. To visualize differences from the mean area between lakes of different sizes I first sorted all lakes within each year by mean area then split the lakes into two bins with equal numbers of lakes. The area of the median lake for each year ranged from 50,000m^2 to 54,000m^2.

Because percent dynamism is negatively correlated with size (Cooley et al. 2019), I also calculated lake dynamism using the RΔ metric for each of the lakes for each year. RΔ represents the mean lateral change along lake shorelines. This metric is independent of lake area because lateral shrinkage generally does not depend on lake size, and both large and small lakes experience a wide range of RΔ values (Cooley 2019). To do this I first identified the max extent of a lake. I then resampled this max extent from 10-meter pixel size to 1x1m. Then I eroded the max extent in all directions 1 meter at a time using Skiimage.morphology binary erosion until the value was approximately equal to the minimum extent. RΔ is the number of erosion increments in meters needed for this.

Quality assessment

To assess the quality of my water classification methods, I performed a manual validation analysis. First, I generated a random sample of observations. Then for each of the observations I generated a random lake that had valid observations for the year of the observation. I then looked at the true color image that came with the Sentinel 2 data and overlayed the buffered lake area polygon as an outline. Next to this image I pulled up the 75 percent water fraction classification

overlayed on the true color image. This allowed me to visually compare my classification to the imagery and determine whether it was accurate.

Chapter 5: Results

Over the five years of data included in this study, on average 1,050 lakes were observed each year (Table 2). The date of the first observation ranges from May 17th – May 25th depending on cloudiness and ice breakup date. Final observations each year occur between September 25th and October 14th. The number of valid observations each year ranges from 9 – 16 images.

Of the lakes in the study area, 405 unique lakes had valid observations spanning the entire study (2019-2023). While this subset of lakes is much smaller than the lakes of years individually, it provides an opportunity to track individual lake seasonality across 5 years. Results for the n = 405 subset are included below.

Year	Number of Lakes	Number of Observations	First Observation Date	Last Observation Date
2019	779	9	May 25th	September 25th
2020	1327	14	May 17th	October 9th
2021	1077	10	May 25th	September 25th
2022	1082	9	May 19th	October 14th
2023	980	16	May 19th	October 4th

Table 2. Observations by year.

Percent difference from mean area by year

Three distinct events occur in the seasonal progression of lake area in all years: an early season peak, a midseason low, and end of season recovery and drawdown (Figure 5).

Early Season Peak in Area

In 2019, smaller lakes and ponds had lower initial values (1.05x the 2019 mean) compared to larger lakes (more than 1.25x the 2019 mean). 2019 is the only year where initial values are larger for larger lakes. Smaller lakes and ponds increase in area rapidly reaching a max extent of 1.4x the 2019 mean, although this peak is short-lived and area values decline sharply by early June. Larger lakes reach their max extent of 1.32x the 2019 mean in early June. In 2020 initial values are slightly higher (1.2x 2020 mean for larger lakes and 1.3x 2020 mean for smaller lakes and ponds). Both larger and smaller lakes increase slightly at the same time (each increasing approximately 0.1x from initial values. In 2021 there is a drastic difference

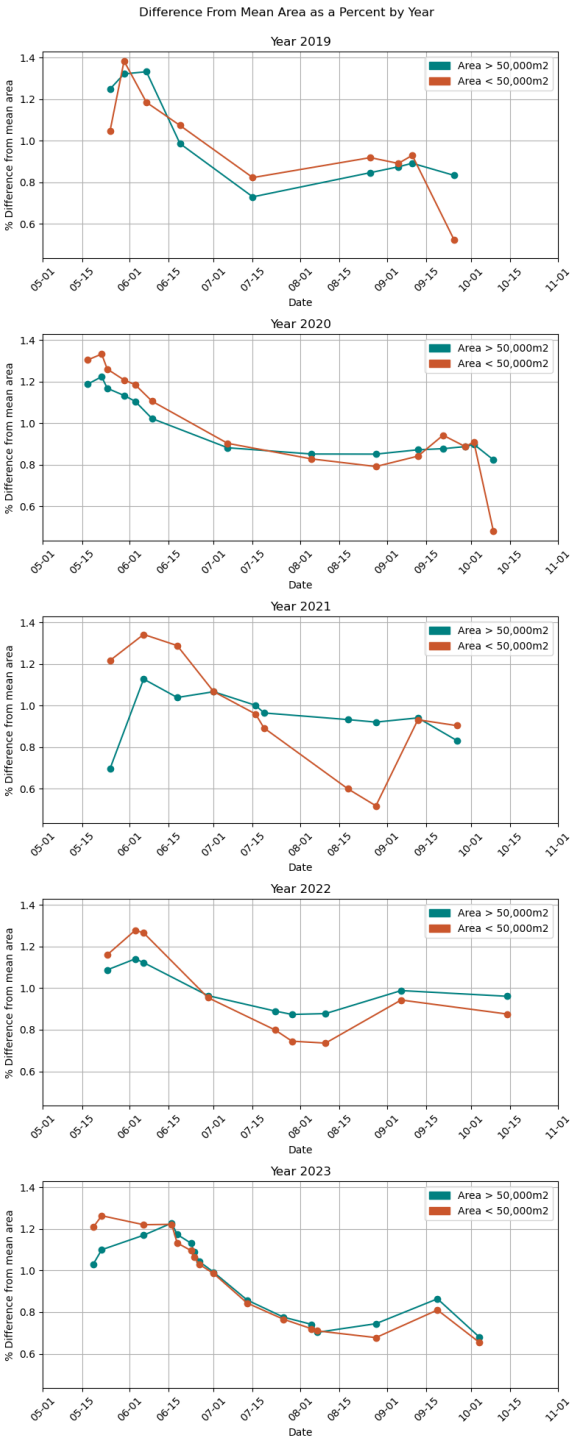

Figure 5. Percent difference from mean area by year. Here lakes for each year are grouped by area into 2 groups. Teal represents larger lakes and orange represents smaller lakes. Observation dates range from May 15th – October 15th.

between larger and smaller lakes. Smaller lakes have characteristic initial values (1.2x 2021 mean), but larger lakes have initial values of only 0.7x the 2021 mean. 2022 has a similar pattern in early season peak to 2020 (initial values of 1.1x and 1.15x 2022 means). In 2023 larger lakes have initial values closer to 1x the 2023 mean and area increases through June 16th whereas smaller lakes and ponds have a more characteristic starting value of 1.2x the 2023 mean and reach their peak values earlier on May 22nd.

Midseason Low in Area

In 2019 all lakes have the same midseason low of July 17th (although the exact timing of this is limited by a lack of data in July and August). Larger lakes have a slightly lower midseason low value (0.75x 2019 mean) compared to larger lakes (0.81x 2019 mean). In 2020 lakes of all sizes decline gradually until early July, reaching a low of 0.85x the 2020 mean. This low holds through early September. 2021 has a noticeable distinction in the progression of larger versus smaller lakes, although they both reach midseason lows at the same time on August 28th. Larger lakes decline gradually reaching a midseason low of 0.92x the 2021 mean, but smaller lakes and ponds have a precipitous decline reaching a low of 0.5x the 2021 mean value. In 2022 smaller lakes and ponds decline more rapidly than larger lakes, but both reach midseason lows in early August. Larger lakes reach a low of about 0.87x the 2022 mean, whereas smaller lakes and ponds reach a low 0.75x the 2022 mean. In 2023 smaller and larger lakes follow almost the exact pattern after June 15th, declining steadily before reaching a midseason low in early August of 0.7x the 2023 mean.

End of season recovery and drawdown

In 2019 both larger and smaller lakes grow in area after the midseason low reaching extent of 0.9x the 2019 mean. There is a large late season drawdown of small lakes and ponds

(reaching a low of 0.5x the 2019 mean). The drawdown of larger lakes is less pronounced (staying above 0.8x the 2019 mean). In 2020 area values hold steady after the midseason low (with marginal increases in area of 0.2x for smaller lakes and ponds in late September). Smaller lakes and ponds then decline precipitously in area, similarly to in 2019 although more rapid, reaching a low of 0.45x the 2020 mean. Larger lakes decline only 0.2x. In 2021 smaller lakes and ponds recover rapidly in area from the drastic midseason low and reach values of 0.9x the 2021 mean in two weeks (by mid-September). Unlike any other year, in 2021 smaller water bodies decline less at the end of the season compared to larger ones. In 2022 both larger and smaller water bodies increase in area after the midseason low to values (0.98x and 0.94x their respective 2022 means) before declining slightly though mid-October. In 2023 both larger and smaller water bodies recover to only 0.8x the 2023 mean (reaching a high around September 15th. All lakes then drawdown to 0.67x the 2023 mean in early October.

All years had broadly similar patterns with peak values in early June, seasonal declines in area in July-August, and late-season increases in area before drawdown in late-September to early-October. In all years, smaller lakes and ponds were more seasonally variable, reaching higher early season peak values and lower midseason lows (excluding 2019 where larger lakes reached lower values). The timing of midseason lows varied from year to year. In 2020 lake extents reached midseason lows earliest (beginning of July) and held steady through the rest of summer. In 2019 the midseason low occurred earlier (mid-July) than in 2022 and 2023, when it occurred in mid-August. In 2021 the midseason low occurred the latest at the end of August. The early season peak is much more pronounced for smaller water bodies in 2021 (20% higher than that of larger water bodies) and the midseason low is 40% lower than that of larger water bodies.

RΔ

RΔ represents the mean lateral change along lake shorelines in meters. RΔ varies from year to year (Figure 6). 2019 has the highest range in values (0m-160m) as well as the largest spread in values from (18m-75m). 2019 also has the highest median RΔ value (35m) and highest mean value (54.11m). The year with the smallest spread in RΔ values is 2023, with values concentrated between (10m-30m). 2023 also has the lowest median value (13m) and mean value (19.89m). 2020 has similar RΔ to 2023, with a slightly higher median value (17m). 2020 did have a higher mean (28.48m) and higher range in values (up to 80m) compared to 2023 (up to 60m). 2021 and 2022 had more moderate values (medians 27m and 21m respectively). Although 2022 has a higher mean value (49.56m compared to 2021 mean of 39.45m).

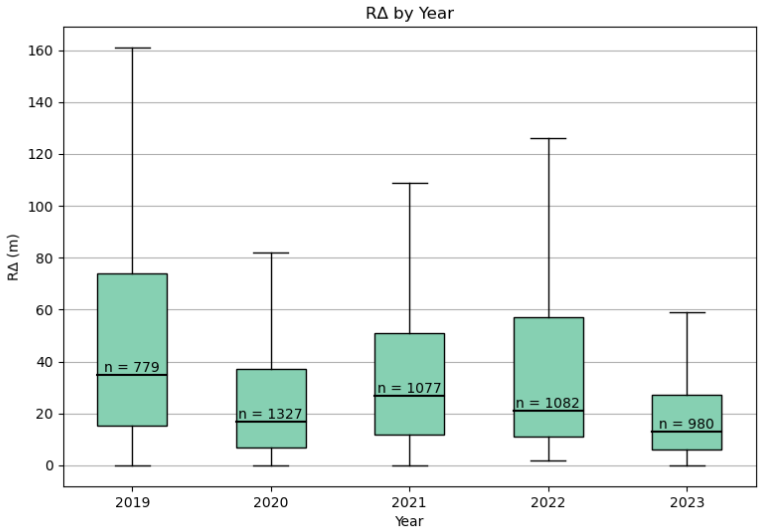

Figure 6. RΔ for each year. The number of lakes included for each year is given above the median. Outliers have been excluded in this graph.

Percent difference from mean area n = 405

Patterns in the n = 405 subset of percent difference from the mean area are similar overall to the seasonal trends seen in all lakes (Figure 7). Differences from the mean are smaller in magnitude for small and larger lakes in 2020, 2021, and 2022, especially for the early season values. Notable differences include that in 2021 there is not the distinct early season split between small and large lakes, but the midseason low of smaller lakes (0.42x the study mean) is much lower than that of larger lakes. The lakes in this subset also had smaller initial values in 2022 than compared to all lakes in that year. Another key difference is that smaller lakes in this subset have much higher early season values in 2023 (over 1.6x the study mean versus 1.2x the study mean for larger lakes).

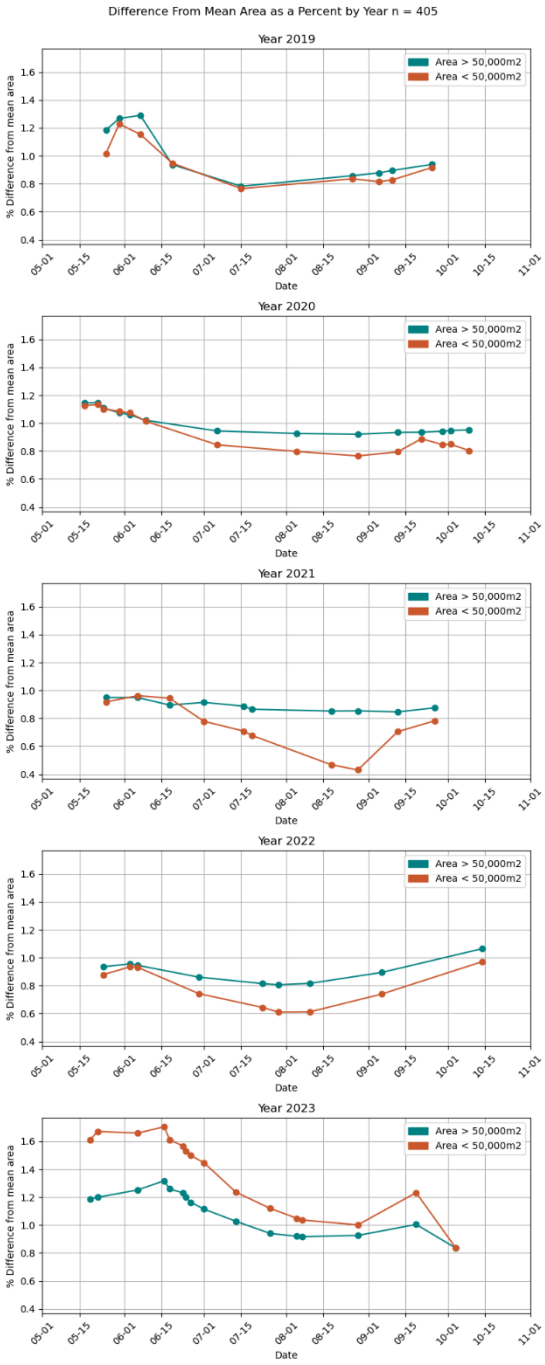

Figure 7. Percent difference from mean area by year for the same lakes over all five years. Here lakes for each year are grouped by area into 2 groups. Teal represents larger lakes and orange represents smaller lakes. Observation dates range from May 15th – October 15th.

RΔ n = 405

Lakes in the n = 405 subset have very similar patterns in RΔ values compared to all lakes for all years (Figure 8). One difference is that the highest values are higher in this subset of data for 2019 (180m versus 161m), 2021 (127m versus 110m), and 2022 (148m versus 125m). Median values also overall appear slightly higher in the n = 405 subset.

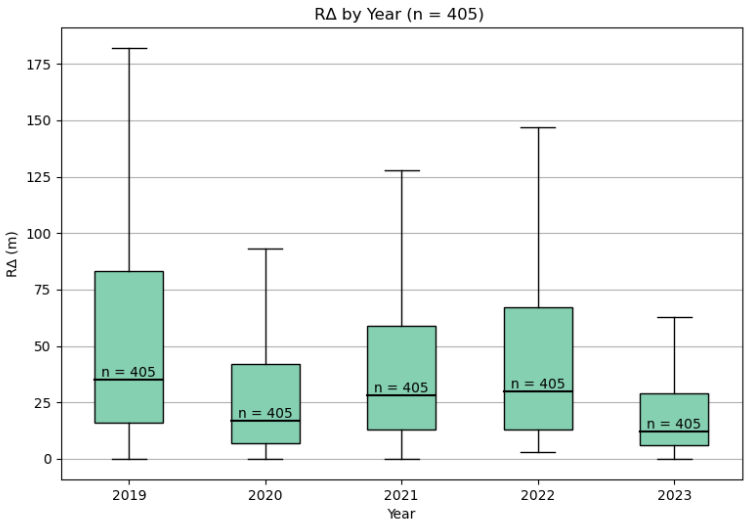

Figure 8. RΔ for each year n = 405. The number of lakes included for each year is given above the median. Outliers have been excluded in this graph.

Figure 9 shows the average RΔ value for the 405 lakes. In this map there are a range of RΔ values across lakes of varying areas. There are also lakes with very different RΔ values near one another. Mean RΔ and mean area are not correlated ($R^2 = 0.1458$).

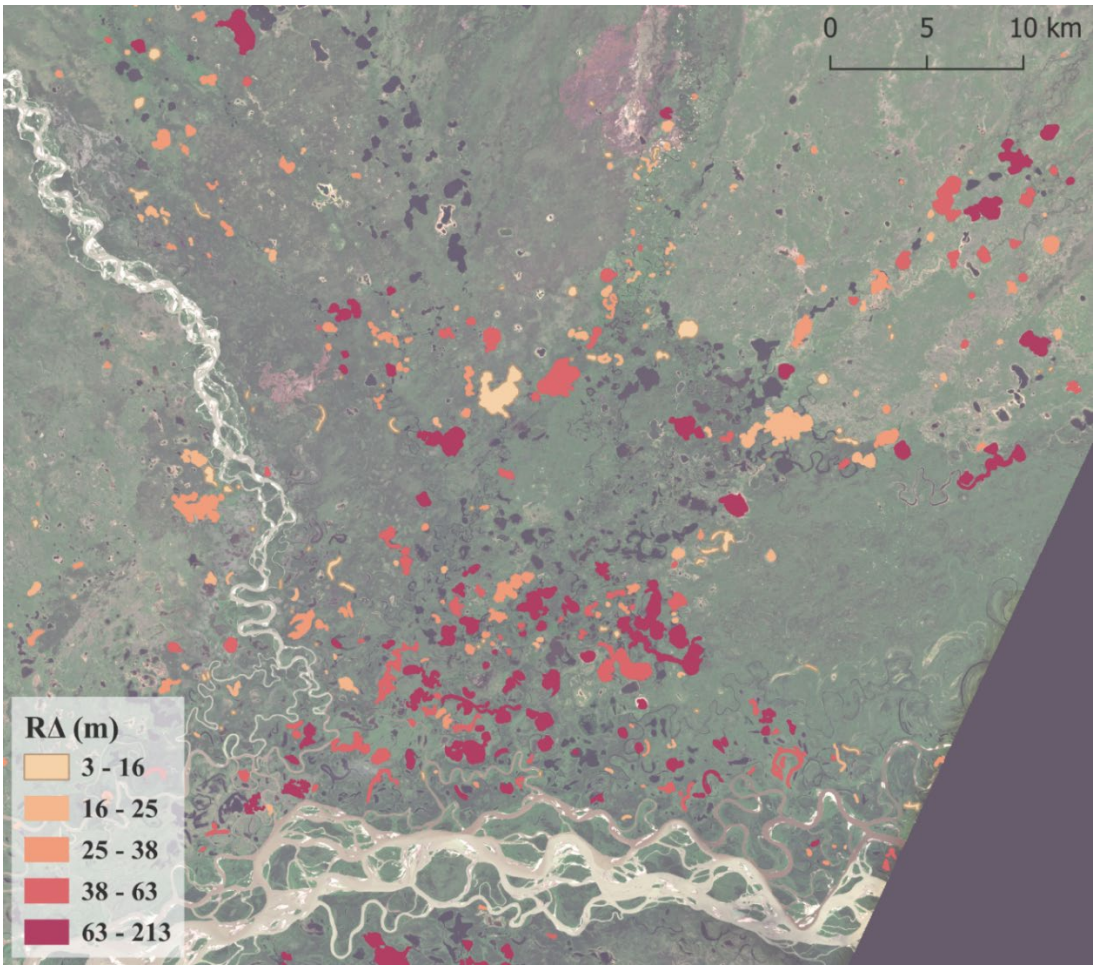

Figure 9. Map of RΔ values for the n = 405 subset of lakes. Red lakes have higher RΔ values. These classes were created using equal interval (quantile) breaks. The Yukon River runs across the south of the map, the Chandalar River runs vertically down the left side.

Quality assessment results

Overall, it appears that the classification works well. Figure 10 shows three example quality assessment results. The patterns in these three lakes are representative of the 40 lakes I sampled and visually assessed. The lake in Figure 10A shows that the classification is generally good for lakes with landforms within a lake not counted as water, although it slightly underestimates water. The lake in Figure 10B shows that the classification works well on lakes that appear different colors in optical imagery. Figure 10C shows a small waterbody that is a lake in the input shapefile but is a river channel in the image from 2020, reflecting the high interannual variability in such water bodies.

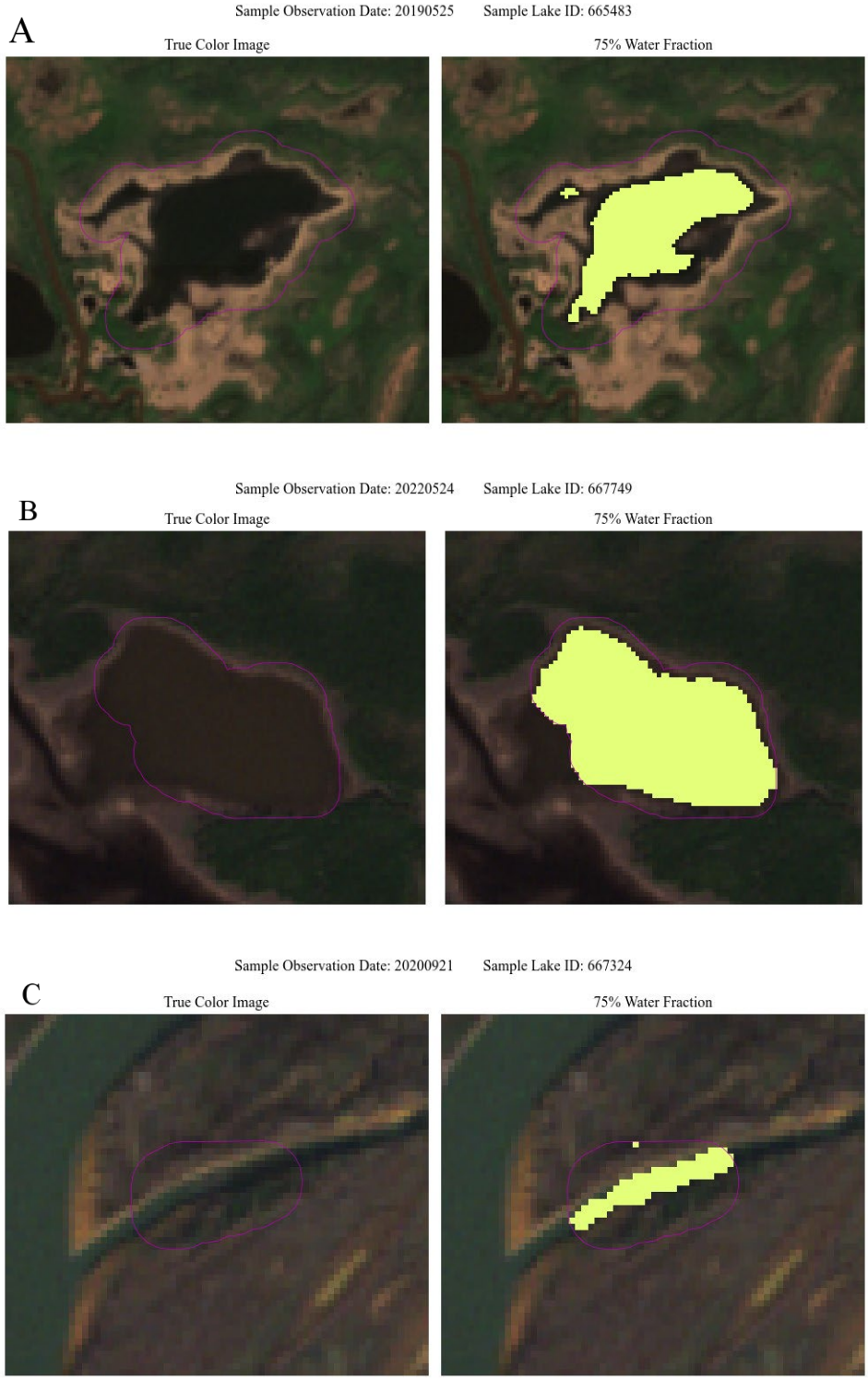

Figure 10. Quality Assessment Examples. Lake object showing the water classification in yellow and the buffered lake polygon in pink. A. a lake that shows underestimation. B. a lake with different coloration. C. a lake that is more of a river channel in 2020.

Chapter 6: Discussion

Out of the 2176 lakes in the original shapefile of the AOI, only 405 lakes had 90% cloud free observations over the entire study period. On average these lakes are slightly larger the mean lake size. The mean (median) area of the 2176 lakes and ponds in the original shapefile is 69,981m^2 (27,215 m^2) compared to 122,772 m^2 (11,508 m^2) in the n = 405 subset. This is likely due to the transient nature of these small water bodies that do occur over the entire five-year period. It also reflects the high dynamism in this region.

Few studies have examined how sub-seasonal variability varies from year to year. This analysis provides a comparison of five years of sub-seasonal variability data. Lake area changes generally follow the subseasonal progression seen in previous research (Cooley et al. 2019). Lake area starts high early in the season then decreases until July/August before increasing slightly through the end of September. The specific timing of these patterns varies from year to year, especially when the midseason low occurs. The early season peak has the highest variability values ranging from 20-40% above the mean area values. Whereas the midseason low is between 80-90% of the mean area values (excluding for smaller waterbodies in 2021). This seasonal progression is somewhat apparent in the n = 405 subset of lakes. The larger lakes in this subset have low variability overall. The high early season peak of smaller lakes in 2023 and the low midseason low of smaller lakes in 2021 are both large in magnitude in this subset. But the early season peak is virtually nonexistent in both large and small lakes in 2020 – 2022. The late season drawdown is also not as pronounced and is most clear in 2023. This subset is not representative of the lakes overall and has much larger areas on average. This likely contributes to the progression being less apparent in this subset.

Smaller water bodies have more pronounced seasonal progressions with higher early season peaks and lower midseason lows. This distinct pattern in 2021 aligns with Mullen et al. 2023 who find that ponds (< 10,000 m^2) decrease by 40% in total surface area in 2021 (compared to ~10% for larger water bodies).

Patterns in lake seasonality are important in conjunction with recent findings that suggest that lake areas in the Arctic are decreasing faster than models have predicted (Webb et al. 2022, Nitze et al., 2018). Teasing out subseasonal variations in area is crucial in understanding how these large-scale declines in Arctic lake area are occurring.

Interestingly, these patterns with seasonal variability do not necessarily correlate with RΔ. While 2019 has the highest RΔ values, reflecting greater dynamism, 2023 has the lowest RΔ values. RΔ is uncorrelated with lake size. The map of average RΔ values shows that this metric of dynamism may be correlated with connectivity to the Yukon in the southern portion of the study area. The proximity of lakes with ranging RΔ values suggests an area of further research.

Limitations

One limitation with this study is that the shapefile used is not representative of all lakes across the five-year period. Due to the high dynamism in this area, lakes that one year appear as lakes are river channels in the following year, making it difficult to get observations for the same lakes for all five years. Another limitation with this study is the frequency of observation dates. 2020 and 2023 have more observations 14 and 16 respectively. But 2019, 2021, and 2022 all have 9-10 observations. Gaps between these observations could change the appearance of the subseasonal progression in these years.

Another limitation of this study is the sensitivity of the reflectance values of the green band used in the NDWI calculations. In preliminary analyses the green band values for some

lakes appeared speckled. These lakes also had uncharacteristically low green band values. Many of these images coincided with observations that did not have clear bimodal NDWI histograms and were removed from analysis.

Future directions

The patterns of RΔ across the study area warrant further analysis to determine what factors are driving different lakes near each other to have very different amounts of dynamism. Potential sources of data of to connect to include river discharge gauge data from gauges on the Yukon River downstream of the study area and climate data from nearby weather stations.

A second direction I would pursue in the future is to group lakes differently in the comparison of difference from mean as a percent. In this analysis I chose to divide the lakes in half based on area. I would develop this further by comparing different area grouping—using three or more groups or exploring different thresholds for groups. Additionally, I would like to use metrics other than lake area (ex. spatial relationships or geology) to determine patterns.

Chapter 7: Conclusion

This study of lake area dynamics in the Yukon Flats serves as an example of analysis of subseasonal variability in surface water. This is important for understanding how lakes and ponds in the Arctic are connected to thawing permafrost and how they contribute to the global carbon cycle. Seasonal and interannual changes to these lakes and ponds also have implications for the people who rely on these lands. The community of Fort Yukon at the southern edge of the study area is Alaska's largest Athabascan village. The lakes and ponds in this study area are a key part of the hunting areas for local communities. The perspectives of those who have experienced these seasonal changes over generations would be a meaningful addition to future studies in this region.

Surface water extent is quite variable across the study area. Lakes of different sizes show drastically different subseasonal dynamics. Lakes close to each other also exhibit different subseasonal dynamics. From year-to-year temporal and spatial patterns in water area is relatively consistent, although more variable in 2019 and 2023.

Bibliography

Alsdorf, D. E., Rodríguez, E., & Lettenmaier, D. P. (2007). Measuring surface water from space. Reviews of Geophysics, 45(2). https://doi.org/10.1029/2006RG000197

Bastviken, D., Cole, J., Pace, M., & Tranvik, L. (2004). Methane emissions from lakes: Dependence of lake characteristics, two regional assessments, and a global estimate. Global Biogeochemical Cycles, 18(4). https://doi.org/10.1029/2004GB002238

Bring, A., Fedorova, I., Dibike, Y., Hinzman, L., Mård, J., Mernild, S. H., Prowse, T., Semenova, O., Stuefer, S. L., & Woo, M.-K. (2016). Arctic terrestrial hydrology: A synthesis of processes, regional effects, and research challenges. Journal of Geophysical Research: Biogeosciences, 121(3), 621–649. https://doi.org/10.1002/2015JG003131

Carroll, M. L., & Loboda, T. V. (2017). Multi-Decadal Surface Water Dynamics in North American Tundra. Remote Sensing, 9(5). https://doi.org/10.3390/rs9050497

Chen, M., Rowland, J. C., Wilson, C. J., Altmann, G. L., & Brumby, S. P. (2014). Temporal and spatial pattern of thermokarst lake area changes at Yukon Flats, Alaska. Hydrological Processes, 28(3), 837–852. https://doi.org/10.1002/hyp.9642

Cooley, S. W., Smith, L. C., Ryan, J. C., Pitcher, L. H., & Pavelsky, T. M. (2019). Arctic-Boreal Lake Dynamics Revealed Using CubeSat Imagery. Geophysical Research Letters, 46(4), 2111–2120. https://doi.org/10.1029/2018GL081584

Cooley, S. W., Smith, L. C., Stepan, L., & Mascaro, J. (2017). Tracking Dynamic Northern Surface Water Changes with High-Frequency Planet CubeSat Imagery. Remote Sensing, 9(12). https://doi.org/10.3390/rs9121306

Downing, J. A., & Duarte, C. M. (2009). Abundance and Size Distribution of Lakes, Ponds and Impoundments. In G. E. Likens (Ed.), Encyclopedia of Inland Waters (pp. 469–478). Academic Press. https://doi.org/10.1016/B978-012370626-3.00025-9

Downing, J. (2008). Emerging global role of small lakes and ponds: Little things mean a lot. Limnetica, 29, 9–24. https://doi.org/10.23818/limn.29.02

Elder, C. D., Thompson, D. R., Thorpe, A. K., Hanke, P., Walter Anthony, K. M., & Miller, C. E. (2020). Airborne mapping reveals emergent power law of Arctic methane emissions. Geophysical Research Letters, 47, e2019GL085707. https://doi.org/10.1029/2019GL085707

Finger-Higgens, R. (2022). Diminishing Arctic lakes. Nature Climate Change, 12(9), 782–783. https://doi.org/10.1038/s41558-022-01466-7

Huang, C., Chen, Y., Zhang, S., & Wu, J. (2018). Detecting, Extracting, and Monitoring Surface Water From Space Using Optical Sensors: A Review. Reviews of Geophysics, 56(2), 333–360. https://doi.org/10.1029/2018RG000598

Jepsen, S. M., Voss, C. I., Walvoord, M. A., Rose, J. R., Minsley, B. J., & Smith, B. D. (2013). Sensitivity analysis of lake mass balance in discontinuous permafrost: The example of disappearing Twelvemile Lake, Yukon Flats, Alaska (USA). Hydrogeology Journal, 21(1), 185–200. https://doi.org/10.1007/s10040-012-0896-5

Jones, B. M., Grosse, G., Arp, C. D., Jones, M. C., Walter Anthony, K. M., & Romanovsky, V. E. (2011). Modern thermokarst lake dynamics in the continuous permafrost zone, northern Seward Peninsula, Alaska. Journal of Geophysical Research: Biogeosciences, 116(G2). https://doi.org/10.1029/2011JG001666

Lehner, B., & Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. Journal of Hydrology, 296(1), 1–22. https://doi.org/10.1016/j.jhydrol.2004.03.028

Levenson, E. (2024). Alaska Lake and Pond Occurrence Dataset (ALPOD)

McCullough, I. M., Loftin, C. S., & Sader, S. A. (2012). High-frequency remote monitoring of large lakes with MODIS 500m imagery. Remote Sensing of Environment, 124, 234–241. https://doi.org/10.1016/j.rse.2012.05.018

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. https://doi.org/10.1080/01431169608948714

Mullen, A. L., Watts, J. D., Rogers, B. M., Carroll, M. L., Elder, C. D., Noomah, J., Williams, Z., Caraballo-Vega, J. A., Bredder, A., Rickenbaugh, E., Levenson, E., Cooley, S. W., Hung, J. K. Y., Fiske, G., Potter, S., Yang, Y., Miller, C. E., Natali, S. M., Douglas, T. A., & Kyzivat, E. D. (2023). Using High-Resolution Satellite Imagery and Deep Learning to Track Dynamic Seasonality in Small Water Bodies. Geophysical Research Letters, 50(7), e2022GL102327. https://doi.org/10.1029/2022GL102327

Nitze, I., Grosse, G., Jones, B.M. et al. Remote sensing quantifies widespread abundance of permafrost region disturbances across the Arctic and Subarctic. Nat Commun 9, 5423 (2018). https://doi.org/10.1038/s41467-018-07663-3

Olthof, I., Fraser, R. H., & Schmitt, C. (2015). Landsat-based mapping of thermokarst lake dynamics on the Tuktoyaktuk Coastal Plain, Northwest Territories, Canada since 1985. Remote Sensing of Environment, 168, 194–204. https://doi.org/10.1016/j.rse.2015.07.001

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. https://doi.org/10.1038/nature20584

Smith, L. C., Sheng, Y., & MacDonald, G. M. (2007). A first pan-Arctic assessment of the influence of glaciation, permafrost, topography and peatlands on northern hemisphere lake distribution. Permafrost and Periglacial Processes, 18(2), 201–208. https://doi.org/10.1002/ppp.581

Stackpoole, S. M., Butman, D. E., Clow, D. W., Verdin, K. L., Gaglioti, B. V., Genet, H., & Striegl, R. G. (2017). Inland waters and their role in the carbon cycle of Alaska. Ecological Applications, 27(5), 1403–1420. https://doi.org/10.1002/eap.1552

Tranvik, L. J., Downing, J. A., Cotner, J. B., Loiselle, S. A., Striegl, R. G., Ballatore, T. J., Dillon, P., Finlay, K., Fortino, K., Knoll, L. B., Kortelainen, P. L., Kutser, T., Larsen, Soren., Laurion, I., Leech, D. M., McCallister, S. L., McKnight, D. M., Melack, J. M., Overholt, E., … Weyhenmeyer, G. A. (2009). Lakes and reservoirs as regulators of carbon cycling and climate. Limnology and Oceanography, 54(6part2), 2298–2314. https://doi.org/10.4319/lo.2009.54.6_part_2.2298

Turetsky, M. R., Abbott, B. W., Jones, M. C., Anthony, K. W., Olefeldt, D., Schuur, E. A. G., Grosse, G., Kuhry, P., Hugelius, G., Koven, C., Lawrence, D. M., Gibson, C., Sannel, A. B. K., & McGuire, A. D. (2020). Carbon release through abrupt permafrost thaw. Nature Geoscience, 13(2), 138–143. https://doi.org/10.1038/s41561-019-0526-0

Walvoord, M. A., & Kurylyk, B. L. (2016). Hydrologic Impacts of Thawing Permafrost—A Review. Vadose Zone Journal, 15(6), vzj2016.01.0010. https://doi.org/10.2136/vzj2016.01.0010

Webb, E.E., Liljedahl, A.K., Cordeiro, J.A. et al. Permafrost thaw drives surface water decline across lake-rich regions of the Arctic. Nat. Clim. Chang. 12, 841–846 (2022). https://doi.org/10.1038/s41558-022-01455-w

van Everdingen, R. (Ed.) (1998), Multi-language glossary of permafrost and related ground-ice terms, Natl. Snow and Ice Data Cent., Boulder, Colo

Verpoorter, C., T. Kutser, D. A. Seekell, and L. J. Tranvik (2014), A global inventory of lakes based on high-resolution satellite imagery, Geophys. Res. Lett., 41, 6396–6402, doi:10.1002/2014GL060641.