

ABSTRACT

Title of Dissertation: ESTABLISHING LINKAGES BETWEEN
THE SATELLITE OBSERVED SURFACE
WATER DYNAMICS AND POTENTIAL
DRIVERS OF CHANGE IN HIGH
NORTHERN LATITUDES OF NORTH
AMERICA

Mark L. Carroll, Doctor of Philosophy, 2018

Dissertation directed by: Dr. Tatiana Loboda Department of
Geographical Sciences

Climate change is affecting aspects of life across the globe. Nowhere is this more prevalent than in the High Northern Latitudes where the Arctic Amplification of climate change has resulted in rates of warming that are twice the global average. Rising air temperatures drive deeper thawing of permafrost which is expressed, among many other ways, through changes in surface water extent. In this dissertation I developed annual maps of surface water extent from a 30 year series of satellite observations from Landsat over a large region of North American tundra. These maps were used in an object based approach to identify water bodies that show a significant trend in surface area over the past 30 years. Over 25% of the 675,000 water bodies in my study region experienced a statistically significant ($p < 0.05$) trend in surface area change between 1985 and 2015. The analysis reveals that water bodies with a net

increasing trend and those with a net decreasing trend are spatially clustered. A distinct pattern of increasing extent in the Northwest and decreasing extent in the Southeast of the study region became clear when change was related to specific watersheds. The watersheds that were dominated by decreasing extent were found to be correlated with presence of bedrock on the surface indicating that shallow soils limit subsurface connectivity and enhance potential for evaporation. There is limited observational data for climate and weather in the region with only four weather stations unevenly distributed in the study region. Therefore, reanalysis data from Modern Era retrospective Reanalysis for Research and Applications (MERRA-2) is used to explore potential climate drivers of surface water change. Surface temperature and ground heating flux in the spring and fall transition periods (shoulder seasons) were found to be good predictors of surface water change. The methodological advances in this dissertation including the object based analysis of water bodies through annual time series and the use of machine learning techniques in high end computing will facilitate future continental scale assessments of surface water extent and the attribution of that change to environmental drivers of change.

ESTABLISHING LINKAGES BETWEEN THE SATELLITE OBSERVED
SURFACE WATER DYNAMICS AND POTENTIAL DRIVERS OF CHANGE IN
HIGH NORTHERN LATITUDES OF NORTH AMERICA

by

Mark L. Carroll

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:
Professor Tatiana Loboda, Chair
Dr. Megan Lang
Dr. Wayne McIntosh
Dr. John Schnase
Dr. Kathleen Stewart

© Copyright by
Mark L. Carroll
2018

Dedication

For my wife Angie and my children Brent and Evelyn who bore with me through this process.

Acknowledgements

First I would like to thank my committee chair for all of her support, advice and determination to guide me through this process. Thank you to my committee members for all of the guidance and feedback that made this work better. I would like to thank John Jones for providing access to the Dynamic Surface Water Extent data that was used in the annual maps. I would like to acknowledge the NASA Center for Climate Simulations for providing access to the Advanced Data Analytics Platform (ADAPT) where most of the data processing was performed. I would also like to specifically thank Jian Li and Roger Gill for their help with programming and analysis of MERRA data and Maggie Wooten who picked up the slack on other work when I was distracted with this dissertation.

Table of Contents

Dedication.....	ii
Acknowledgements.....	iii
Table of Contents.....	iv
List of Tables.....	vi
List of Figures.....	viii
List of Abbreviations.....	xiii
Chapter 1: Introduction.....	1
1.1: Background and Motivation.....	1
1.2: Research Questions.....	9
1.3: Organization of the study.....	10
Chapter 2: Multi-decadal Surface Water Dynamics in North American Tundra.....	13
2.1: Introduction.....	13
2.2: Study area.....	16
2.3: Methods.....	17
2.3.1: Annual product generation.....	19
2.3.2: Annual product accuracy assessment.....	21
2.3.3: Identifying unique water bodies.....	22
2.4: Results.....	24
2.4.1: Accuracy assessment results.....	24
2.4.2: Long term water dynamics.....	24
2.5: Discussion.....	32
2.6: Conclusions.....	36
Chapter 3: The sign, magnitude and potential drivers of change in surface water extent in Canadian tundra.....	38
3.1: Introduction.....	38
3.2: Study area.....	42
3.3: Data and Methods.....	44
3.3.1: Identifying spatial patterns of change.....	45
3.3.2: Random forest analysis.....	46
3.3.3: Relating spatial patterns of change to ecological drivers.....	47
3.4: Results.....	49
3.5: Discussion.....	58
3.6: Conclusions.....	61
Chapter 4: Identifying potential drivers of change in surface water extent in Arctic tundra using reanalysis data.....	63
4.1: Introduction.....	63
4.2: Data and Methods.....	66
4.2.1: Dependent variable.....	68
4.2.2: Independent variables.....	69
4.2.3: Identification of primary predictors.....	71
4.3: Results.....	74
4.4: Discussion.....	78
4.5: Conclusions.....	82

Chapter 5: Conclusion	84
5.1: Major research findings	84
5.2: Study limitations	88
5.3: Big data analytics	90
5.4: Contribution of this research to the broader North American Arctic research	93
5.4.1: Broader Arctic agenda	93
5.4.2: Future directions	95
Appendix.....	98
Chapter 3: Supplemental Material	98
A.1: Land cover data.....	98
A.2: Vegetation Cover and Condition	100
A.3: Other Surface and Subsurface Features	101
A.4: Weather/Climate Data.....	102
Bibliography	106

List of Tables

Chapter 2

Table 2.1 Confusion matrix for accuracy assessment of annual water maps using WorldView-2 multi-spectral data.

Table 2.1 Distribution of water bodies detected in the study region by size. All sizes are in hectares and based on average area over the 31 year study period. (The very low fractional representation of the largest water bodies in the total water body count necessitates the use of precision values to the third decimal point.)

Table 2.3 Overall 31 year trend in size of water bodies.

Table 2.4 Distribution of water bodies that exhibit significant trend in surface water area at $p < 0.05$ by size and trajectory of change.

Chapter 3

Table 3.1 Ranking of variables from the random forest analysis after all runs.

Table 3.2 Results from linear regression of surface type per watershed with respect to surface water change using R^2 to measure correlation between the variable. P-values denoted by symbol: no symbol denotes $p > 0.05$; * denotes $p < 0.05$; *** denotes $p < 0.001$.

Chapter 4

Table 4.1 List of MERRA-2 variable names and plain English explanations for what the variables represent and associated units that were used in this analysis.

Table 4.2 The relationship between the results of MaxEnt runs for increasing and decreasing surface water extent relative to the training data.

Table 4.3 Top five predictors for increasing surface water extent ranked by the permutation importance. Numbers are converted to percent contribution to the model.

Table 4.4 Top ten predictors for decreasing surface water extent ranked by the permutation importance. Numbers are converted to percent contribution to the model.

List of Figures

Chapter 1

Figure 1.1 False color composite (Red, Near-Infrared, Blue assigned to RGB respectively) of Landsat scene from a region near Barrow, Alaska. Water can be seen as the dark, mostly black, features in the landscape. This density and diversity of size of water bodies is typical in the Arctic tundra.

Figure 1.2 The distribution of permafrost within the circumpolar Arctic. This figure has been reproduced from International Permafrost Association (<https://ipa.arcticportal.org>) with data from (Brown et al. 1998).

Chapter 2

Figure 2.1 Study area in North American Arctic region, north central Nunavut territory Canada. The study area is located primarily in the Southern Arctic ecoregion and is characterized by low topography and numerous small to moderate sized water bodies. The Queen Maud Gulf Bird sanctuary is indicated with the red polygon.

Figure 2.2 Algorithm flow for the generation of annual water maps from the individual dates of DSWE (reproduced from [28]). In panel 1 (left most panel) individual scenes are converted from four classes to two (land and water) then summed to get total observations of land and total observations of water for the period. Panel 2 shows the “total water” for an individual path/row. Panel 3 shows the mosaicked non-overlapping path/rows which are then summed. Panel 4 shows the final “total water” for the full region of interest.

Figure 2.3 Distribution of WorldView-2 (WV2) scenes used in the accuracy assessment of the annual water maps. Footprints of WV2 scenes are shown as dark grey rectangles distributed throughout the image.

Figure 2.4 In the images above water is shown in black and land is shown in light grey. The first five images show a water body complex (small unnamed water body in northern Nunavut) in five individual years. The final image shows the master map which is the maximum extent in the whole 31 year record. Through time you see the water body shrinks and splits into several components. The master map allows all of these components to be related to the same water body even when they split off into individual pieces. The years shown here are chosen as representative examples of the 31 year record.

Figure 2.5 Total annual area of surface water between 1985 and 2015. Red circles denote local temporal maxima and the minimum in the record.

Figure 2.6 Difference in extent for several water bodies in the study region. The lighter colors indicate that water was present in early years of the study but not present in later years.

Figure 2.7 Variability of river extent and flow between 1985 and 2015. Lighter colors indicate that water was present in some years but not in all years. This is particularly noticeable in the edges of the river and in the islands in the middle of the channel.

Figure 2.8 Spatial distribution of water bodies with a significant ($p < 0.05$) trend in surface water area (ha/yr) over the 31 year time period of the study. Water bodies that are increasing in size are shown in green while the ones that are decreasing in size are shown in red.

Chapter 3

Figure 3.1 Study region in northern Canada, Nunavut territory where multi-decadal trends in surface water extent have been established (Carroll and Loboda, 2017).

Water bodies in the region show opposite trends of surface water extent. Red indicates a water body that is increasing in extent while green indicates a water body that is decreasing in surface water extent. There are four weather stations within or just outside of the study region indicated by purple circles.

Figure 3.2. The left panel shows percent surface water at 3 km spatial resolution with darker shades of blue corresponding to higher percentage of water. The right panel shows percent water change at 3 km spatial resolution. Green colors indicated a higher percentage of water bodies with increasing extent. Red colors indicate a higher percentage of water bodies with decreasing extent.

Figure 3.3 The study region divided into watersheds using the Canada National Hydro Network. Figure 3.3a on left shows the percent surface water in each of the watersheds. Figure 3.3b on right shows the percent net surface water change per watershed. Both results determined by using the annual surface water extent product (Carroll and Loboda, 2017).

Figure 3.4 Annual temperature trends for four weather stations within or in close proximity to the study domain showing individual trends (a - d) and an average of all four stations (e). The overall trend for the domain is increasing temperature at a rate of 0.66° C per decade.

Figure 3.5 Annual precipitation trend for two weather stations within the study region. Cambridge Bay (a) in the north-west shows a slight increasing trend while Baker Lake (b) in the south-east shows a slight decreasing trend.

Figure 3.6 Soil Landscapes of Canada (black boundaries) with zonal statistics showing the sign of surface water change by soil regions.

Figure 3.7 In the left panel (a) is an overview of the Circumpolar Arctic Vegetation Map with 11 cover types described, notably the "Cryptogram Barren Complex" which is shown in salmon color. The right panel (b) shows the result of calculating zonal statistics with CAVM as the zone and pixels of change as the variable. The areas shown in red have decreasing surface water extent and generally relate to the Cryptogram Barren complex of CAVM. (Some CAVM long names have been shortened using the following abbreviations "d-s" is dwarf-shrub and "l-s" is low-shrub)

Figure 3.8 Scatterplot showing the relationship between the percent barren class in each watershed when compared to the net percent change in surface water extent per watershed.

Chapter 4

Figure 4.1 Overview of study domain in Nunavut, Canada. The regional map shows net change in surface water extent at 60 km spatial resolution, to match MERRA spatial resolution, derived from the significant change in Carroll and Loboda 2018.

Figure 4.2 Workflow for generating inputs for the MaxEnt software runs from MERRA-2 Land Surface Diagnostic variables. The MERRA Analytic Service (Schnase et al. 2017) is used to derive weekly and monthly averages for each of 14 variables. Ordinary Least Squares (OLS) regression is performed at each time step (week and month) to show trend at those time steps.

Figure 4.3 Workflow describing the Monte Carlo simulation to determine the top ten predictors for positive change and negative change, respectively at weekly, monthly and combined weekly/monthly time steps. All processing was completed in the ADAPT system NCCS at Goddard Space Flight Center.

Figure 4.4 Overall results from MaxEnt final runs for decreasing surface water extent and increasing surface water extent with the training data shown at 70% transparent in the background. The separate results have been overlain on top of each other setting probability less than 0.25 to no color. Increasing surface water extent is shown with diagonal lines increasing from left to right while negative change is shown with diagonal lines decreasing from left to right. Heavier line weights indicating a higher probability of occurrence in that grid cell. The areas of overlap can be seen where the crosshatch makes a grid pattern.

List of Abbreviations

ABOVE – Arctic and Boreal Vulnerability Experiment

ADAPT – Advanced Data Analytics Platform

AVHRR – Advanced Very High Resolution Radiometer

AUC – Area Under Curve

CAVM – Circum-Arctic Vegetation Map

CPU – Central Processing Unit

DEM – Digital Elevation Model

DSWE – Dynamic Surface Water Extent

ETM – Enhanced Thematic Mapper

HEC – High End Computing

HNL – High Northern Latitudes

MAS – MERRA Analytic Service

MERRA – Modern Era Retrospective reanalysis for Research and Applications

MODIS – Moderate Resolution Imaging Spectroradiometer

NCCS – NASA Center for Climate Simulation

NDVI – Normalized Difference Vegetation Index

TEOW – Terrestrial Ecoregions of the World

TM – Thematic Mapper

VCF – Vegetation Continuous Fields

VHR – Very High Resolution

VM – Virtual Machine

Chapter 1: Introduction

1.1: Background and Motivation

The High Northern Latitudes (HNL) include parts of the circumpolar Arctic and Boreal ecosystems and are generally described as the region above 60° N. The HNL region plays a key role in moderating the climate of the Earth by dissipating heat generated at mid-latitudes, serves as a net carbon sink, and contains large stores of natural resources (Cole et al. 2007; Kortelainen et al. 2004; Post et al. 1990; Stokstad 2004; Turetsky et al. 2008; Walter et al. 2006). The climate in HNL is driven by a large seasonal variability in the amount of incoming solar radiation with a strong impact from the Arctic Ocean, much of which is perennially frozen. This is a temperature limited system with abundant water resources present on the surface or bound in the subsurface in the form of ice (Hinzman et al. 2005; Hobbie 1984). Although the overall precipitation in the region is low - some areas receiving in total less than 150 mm per year with much of that as frozen precipitation in winter (Linacre 1992) – the short (3-5 month) period of snow- and ice- free conditions strongly limits the potential for evapotranspiration.

The HNL region is changing at a much faster rate than other areas of the world (Barber et al. 2008; Bunn et al. 2007; Hinzman et al. 2005; IPCC 2014; Kaufman et al. 2009; Lawrence et al. 2008; Serreze and Francis 2006). There is strong evidence that temperatures have increased more in the HNL region over the past two decades than in the rest of the world, a phenomena referred to as the Arctic amplification of the greenhouse effect (Francis et al. 2009b; McGuire et al. 2006;

Serreze et al. 2009; Serreze and Francis 2006). Increasing temperatures are likely to result in deeper thaw depth in the Arctic and Boreal ecosystem soils (Grosse et al. 2011; Jorgenson et al. 2010; Kokelj and Jorgenson 2013) resulting in a myriad of effects to the ecosystem (Bates et al. 2008). Coincident with the increase in temperatures has been a decrease in sea ice extent (Serreze and Barry 2011; Stroeve et al. 2011) which is projected to continue to decline as temperatures continue to increase through the 21st century (Collins et al. 2013; IPCC 2014). Changes in winter precipitation amounts and type (Barnett et al. 2005; Francis et al. 2009a; Hinzman et al. 2005), increase in cloud cover (Screen and Simmonds 2010; Serreze and Barry 2011), and lower surface albedo (Chapin et al. 2005; Francis et al. 2009b; Subin et al. 2012) are all anticipated responses to climate change. There has been an increase in the occurrence of rain on top of snow pack which causes a change in the moisture content of the snow that subsequently freezes into a solid rather than a granular texture. The so called “Rain on Snow” events impact travel for both humans and wildlife, availability of forage for wildlife as well as the timing of snowmelt (Barnett et al. 2005; Putkonen et al. 2009). For the local population, these rain on snow events have become a hallmark of climate change that affects their regular activities.

Changing climate is expressed through physical changes to the land surface as well as new weather patterns. Hence, characterizing the land surface of the HNL region is critical to understanding the ecosystem processes and developing capabilities for modeling ecosystem change. The water bodies in the HNL region are a key feature of the landscape with several million water bodies in North America alone (Carroll et al. 2011b; Downing 2010). The majority of these water bodies are

smaller than 1 ha in size (Figure 1.1), and are primarily shallow - less than 3 m deep in most cases. The water bodies of this region are also predominantly closed systems with little lateral movement of water between them due to the limited topographic relief (Hobbie 1984). Water levels, and hence surface water extent, in these water bodies are dependent on recharge from local precipitation events

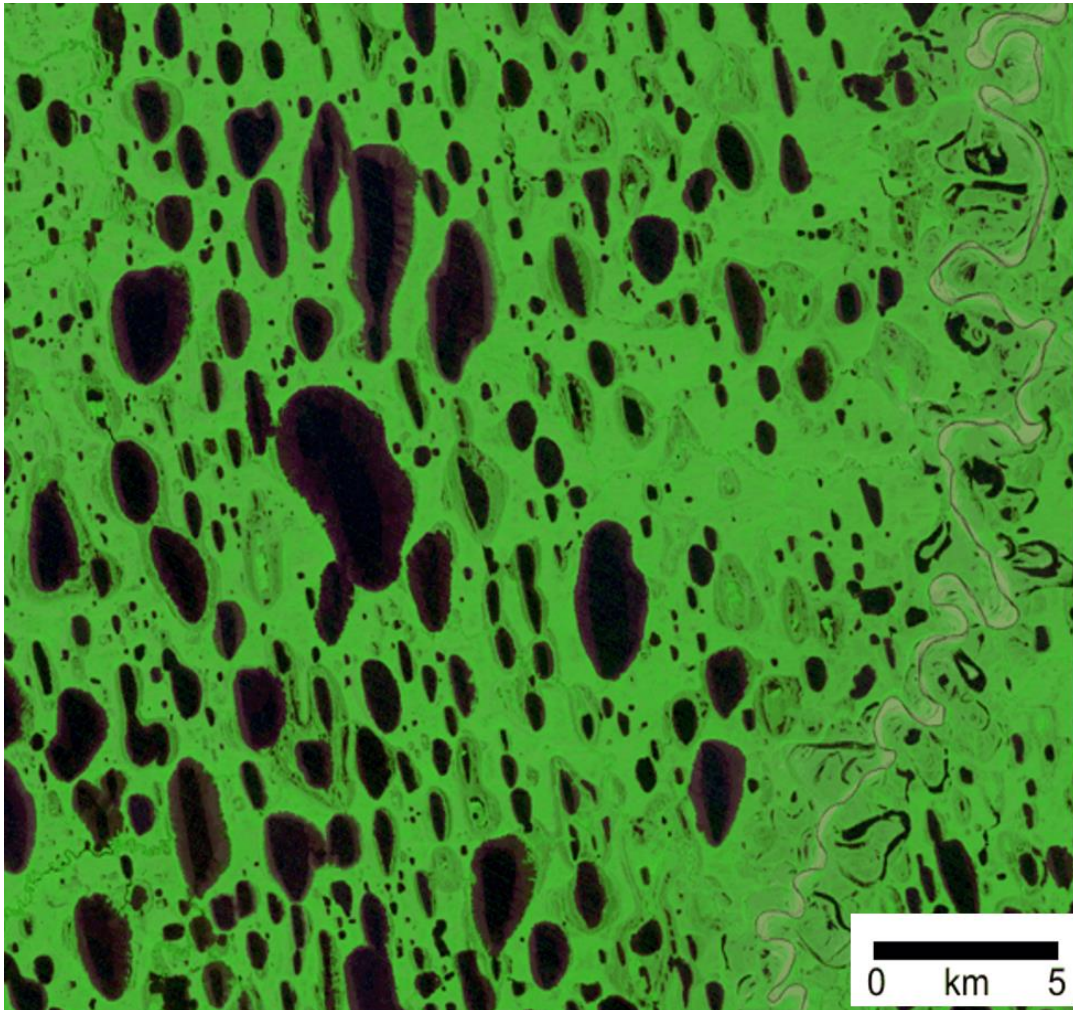


Figure 1.1 False color composite (Red, Near-Infrared, Blue assigned to RGB respectively) of Landsat scene from a region near Barrow, Alaska. Water can be seen as the dark, mostly black, features in the landscape. This density and diversity of size of water bodies is typical in the Arctic tundra.

and on the amount of evaporation during the ice-free period. Ice acts as a cap preventing evaporation of water from a water body. Water, and hence ice, is transient

in the environment and requires frequent measurements to track fluctuations through time. With rising temperatures comes a lengthening of the ice-free period (late May to September) which provides a greater opportunity for evaporation. Increased sediment exposure associated with greater evaporation may lead to the alteration of the biogeochemical cycles resulting in the potential release of soil carbon into the atmosphere which can further influence the Arctic Amplification (Cole et al. 2007). Rising temperatures also result in a longer growing season which provides a greater potential for transpiration from plants, hence a greater potential for changes in surface water extent.

A unique feature of the HNL region is the permanently frozen ground - permafrost. This affects, among other things, the subsurface movement of water both laterally and vertically (Hinkel et al. 2001; Karlsson et al. 2012; Langer et al. 2016; Woo and Guan 2006). The distribution of permafrost is not uniform throughout the Arctic, rather it is stratified by latitude with sporadic and discontinuous permafrost in the southern reaches and becoming increasingly continuous in farther north regions (Figure 1.2) (Brown et al. 1998). Permafrost is particularly sensitive to changes in climate with rising air temperatures causing greater depth of thaw in the upper soil layers that are seasonally thawed, also called active layer (Grosse et al. 2011; Jorgenson et al. 2010). As the permafrost thaws in response to climate change there are three different responses that can occur depending on surface characteristics and

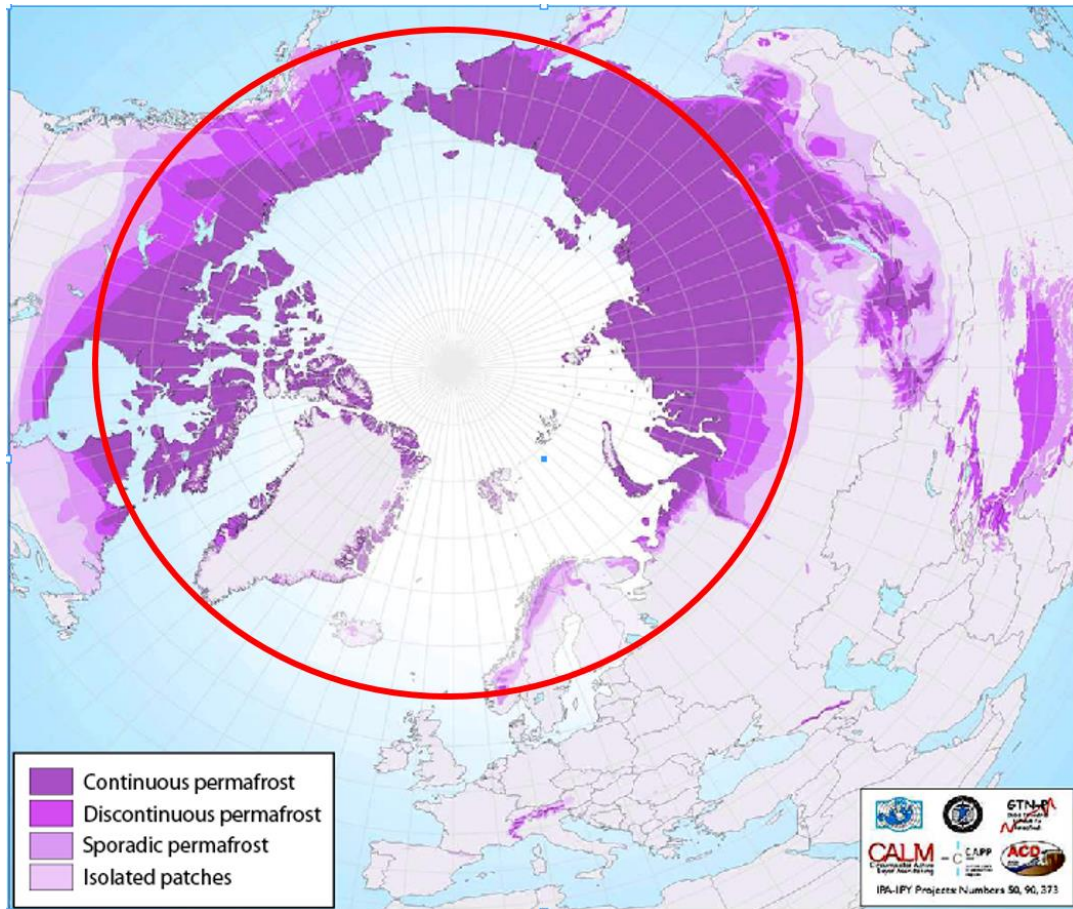


Figure 1.2 The distribution of permafrost within the circumpolar Arctic. This figure has been reproduced from International Permafrost Association (<https://ipa.arcticportal.org>) with data from (Brown et al. 1998).

subsurface composition 1) the soil can collapse forming a depression that fills with water creating a thermokarst lake, 2) the soil around the edges of an existing lake can slump resulting in lake expansion, and 3) deeper thaw depth melts ice wedges or ice in pores of coarse soil enabling vertical and lateral movement of water in the subsurface resulting in lake failure (Grosse et al. 2008; Karlsson et al. 2012; Kokelj and Jorgenson 2013).

The hydrological cycle describes the movement of water through the environment and together with controls to the flow of water both natural (tree fall, beaver dam, etc.) and anthropogenic (dams, levees, etc.) determine the amount of

water that we see on the surface of the earth. Water bodies fluctuate in size at time scales from seasonal (winter snowmelt filling rivers and lakes, monsoonal rains, etc.) to annual and decadal (drought, increased consumption, etc.). In spite of the dynamic nature of water most maps show water as a permanent feature even over long time periods. Generation of a reference map for the location and extent of water is partly dependent on the scale and scope of the research question at hand. Static maps of the location and extent of persistent water bodies have been generated from remotely sensed data at various spatial scales (Carroll et al. 2009; Carroll et al. 2016b; Feng et al. 2015; Lehner and Doll 2004; Pekel et al. 2016; Watts et al. 2012). Only the maps from Carroll et al. (2016), Feng et al. (2015), and Pekel et al. (2016) offer maps of water at spatial resolution finer than 250 m, which is necessary to capture the true range of water body features in the HNL region. Of these only the datasets from Pekel et al. (2016) and Carroll et al. (2016) use a time series of inputs that helps minimize issues related to disparate satellite acquisitions with Landsat data. They also are the only maps that represent multiple epochs with consistently generated maps that facilitate quantification of change in individual water bodies through time.

Maps of surface water extent are used as an ecological determinant for identifying suitable habitat for plants and animals (Adrian et al. 2009), as a component of landcover (Bartholome and Belward 2005; Chen et al. 2015; Defourny et al. 2006; DiMiceli et al. 2011; Friedl et al. 2010) or as an input to a climate or carbon model (Poulter et al. 2017; Subin et al. 2012). Of more environmental interest, work has been done using multiple satellite images or time series to show changes in the extent of water, mostly on a local basis or at moderate to coarse

resolution (Andresen and Lougheed 2015; Carroll et al. 2011b; Karlsson et al. 2012; Pavelsky and Smith 2008; Pietroniro et al. 1999; Rover et al. 2012; Smith et al. 2005). The change in extent of surface water at coarse resolution (25 – 40 km) has been used to inform carbon models that estimate the amount of CH₄ emissions (Poulter et al. 2017). Another study looked at overall depth and general stability of water bodies over time to support the creation of winter ice roads (Jones et al. 2009). Recent work by Pekel et al. (2016) used the Google Earth Engine to produce global annual maps of surface water extent at 30 m spatial resolution. However, instead of using the time series to assess change they divided the time series in half producing two overview maps covering 1985 – 1999 and 2000 – 2015 then simply subtracted the two maps to show where “change” occurred over that time span. All of these studies have been limited in spatial or temporal scope (or both) and in most cases the drivers of change have been broadly attributed to climate change impacting the permafrost and have not been fully explored.

Despite the large body of research that has been performed in the HNL region there remains a significant knowledge gap with regards to the dynamics of surface water. The HNL region is simultaneously data poor and data rich. There is a lack of physical measurements including weather information and subsurface composition. At the same time, a wealth of satellite data is available as well as some thematic maps describing the surface characteristics. In the past, there have been limitations on accessing available data and computing resources to process that data. When the Landsat archive was made freely available, it opened up new avenues for the analysis of research problems (Wulder et al. 2012). However, taking advantage of the dense

time series of Landsat data exposed a new problem with processing and data storage issues due to the large volumes of data. This so called “big data” problem is not new as it has been well-recognized in climate modeling and other areas of environmental studies and requires significant computing resources to overcome. The NASA Center for Climate Simulation (NCCS) provides high end computing resources to the scientific community that are used to run climate models and produce reanalysis data. More recently NCCS has developed the Advanced Data Analytics Platform (ADAPT) to support other science data processing needs including supporting science projects funded under NASA Terrestrial Ecology program’s Arctic and Boreal Vulnerability Experiment (ABOVE). The ADAPT system provides access to Petabyte scale data storage as well as configurable compute resources that enable large scale processing that was previously not possible. High end computing resources are becoming more commonly available at large research institutions and are reducing barriers to large scale analysis of big data including long time series of Landsat data in the Arctic.

As discussed above, much work has already been done to map surface water extent in the Arctic but often the results have been more contradictory than complementary. Surface water dynamics are directly linked to climatic variability, land use and to local conditions including permafrost, drainage networks and soil types. Previous studies have uncovered differences in the rate and sign (increasing or decreasing) of surface water change in local regions but no consensus. The goal of this research is to evaluate the existing data to identify and fill gaps in the knowledge surrounding surface water dynamics at the regional scale including separating inter-annual variability from long term change and identifying potential drivers of change.

The work within this dissertation aims to bring together the cutting edge of big data analytics in the fields of satellite remote sensing and Earth System science to develop the first regional scale analysis of surface water change trends in the tundra regions of North America and provide the context and methodological framework for assessing the impact of various environmental and climatological drivers of the observed trends.

1.2: Research Questions

The overarching research question for my dissertation is: *What is the spatial pattern of variability in the extent of tundra water bodies in High Northern Latitudes (HNL) of North America and what are its environmental determinants?* To answer this question I explore the relationships between surface water, vegetation, climate, and weather in three sub-questions:

1. *What is the long-term dynamic of surface water in HNL tundra?*

There is significant evidence in the literature and anecdotally that small water bodies change in size and shape through time. However there has been no comprehensive work showing whether these are short-term or long-term changes. This question addresses both the frequency of maps required to identify change as well as the composition of maps from each period.

2. *What are the spatial patterns of change in surface water extent and what are the potential environmental drivers of this pattern?*

Through this question I address the spatial relationships, specifically spatial clustering, of change in surface water extent by looking at long term change in

specific water bodies. Furthermore, I examine the potential drivers of the patterns of change by exploring the relationships between change and land cover type.

3. *Through what mechanisms do trends in climate variables in the region explain surface water variability?*

There is a clear trend towards higher temperatures in the circumpolar Arctic. The trend in other climate variables is less well defined. With this question I explore the relationship between these known trends and the spatial patterns found in question 2 using climate reanalysis data from Modern Era Retrospective reanalysis for Research and Applications (MERRA-2) and maximum entropy modeling.

1.3: Organization of the study

The remainder of the dissertation is organized into three chapters, supplementary material and the reference list. Chapter 2 addresses research question 1 with the development of annual maps of surface water extent for the study region in Nunavut, Canada. The maps were generated from the full time series of Landsat data to show the nominal extent of surface water for each year from 1985 – 2015.

“Nominal extent” is used here to describe the extent of water bodies as computed using the probability of water [calculated as (number of water observations/(sum of water and land observations))] with a threshold such that any pixel with a value of greater or equal to 50% probability of water is considered water for that period. These maps were then analyzed to identify overall inter-annual variability in surface water as well as long term trend per water body. The results from these analyses formed the

basis for the work in chapters 3 and 4. This work has been peer reviewed and published (Carroll and Loboda 2017).

Chapter 3 corresponds to research question 2 and highlights the analysis results for spatial clustering of surface water change as well as exploring land surface properties as potential drivers of change. All available maps, satellite based and observational, were interrogated to identify specific features of the surface and subsurface that are related to change in surface water extent. Analysis techniques including spatial averaging, zonal statistics and ordinary least squares regression were used to identify the relationships between descriptive datasets and change results. Important properties of ecosystem function are possible even when observations of features are sparse. Research presented in this chapter has also been peer reviewed and published (Carroll and Loboda 2018).

Chapter 4 corresponds to research question 3 and highlights the results of analysis of MERRA-2 data as a surrogate for sparse weather observations. The large pool of available climate variables from MERRA-2 necessitates the use of high end computing and inferential statistics to identify key variables related to change in surface water extent. Maximum entropy modeling in a machine learning environment is capable of finding relationships between predictor variables and training data even when the training sample is small and some of the predictors are highly correlated. The small sample size is due to the coarse spatial resolution of MERRA-2 which has only 100 grid cells in the study region. Random forest analysis is commonly used in this type of study but requires a large number of training samples which are not available in this case. A Monte Carlo simulation is a series of model runs each with a

random set of inputs performed to identify a convergence in results that is not possible in a single run. The large pool of predictors from MERRA-2 make it difficult to achieve a result from a single run of the maximum entropy model, hence a Monte Carlo simulation is used here to take advantage of the full set of predictors and find a solution.

Chapter 5 presents the major conclusions for the doctoral dissertation with a summary of major scientific and methodological findings from chapters 2 – 4. This chapter also puts the overall dissertation into the broader context of Arctic science and a direction for future work in this area.

Chapter 2: Multi-decadal Surface Water Dynamics in North American Tundra

2.1: Introduction

The North American tundra is a complex landscape where vegetation is interspersed with over one million water bodies, the majority of which are much smaller than 100 hectares (ha) (Carroll et al. 2011b; Downing et al. 2006). Understanding where the water is located and how it is changing over time is a critical component of understanding the role of water in the carbon cycle (Billings et al. 1982; Post et al. 1990), albedo (Chapin et al. 2005; McGuire et al. 2006), energy balance (Francis et al. 2009b; Slater et al. 2007; Subin et al. 2012), and quantifying habitat for migratory wildlife (Hinzman et al. 2005; Stow et al. 2004). In the context of global warming, the Arctic region is warming much faster than lower latitudes (Serreze et al. 2009; Serreze and Barry 2011; Serreze and Francis 2006). The Arctic also has a greater proportion of terrestrial surface occupied by surface water than lower latitudes (Downing et al. 2006; Muster et al. 2013; Prigent et al. 2012). The sediments in Arctic water bodies are known to have high concentrations of carbon that can be released to the atmosphere if the surface water changes and/or if the active layer deepens (Cole et al. 2007; Kortelainen et al. 2004; Stokstad 2004; Turetsky et al. 2008; Walter et al. 2006). Given the higher proportion of surface water in the Arctic, the large amount of carbon in the terrestrial Arctic, and the observations of changes in surface water it is important to understand not only the extent of surface water in the Arctic but also how that surface water is changing through time.

In recent years the extent of surface water has been mapped with several approaches and at regional to global scales. Each of these products has represented an advance in the mapping capabilities but also still has significant limitations including spatial resolution (Carroll et al. 2009; Lehner and Doll 2004; Salomon et al. 2004), temporal resolution (Feng et al. 2015; Verpoorter et al. 2014) or both (Pekel et al. 2016). Most recently, work has been done to generate continental scale maps of surface water extent at the decadal time step using a time series of Landsat data as input (Carroll et al. 2016a; Carroll et al. 2016b). Unlike many other 30 m spatial resolution water maps, the decadal water maps use three years of inputs to generate a single map in order to have enough repeated observations throughout the study region to have confidence that the water that was detected was not a false detection caused by cloud shadows or burn scars, and did not represent a local minimum or maximum caused by short-term weather events (i.e. drought or flood). These decadal water maps represent an advance in the capability to determine surface water change but still do not provide the temporal detail necessary to determine if these decadal maps represent actual long-term trends in water dynamics.

In addition to the mapping efforts there have been attempts to quantify change in surface water extent over time. Most of these change studies (Briggs et al. 2014; Carroll et al. 2011b; Smith et al. 2005) quantify change between two distinct time periods and focus on either a few distinct lakes or the overall regional change. This methodology works reasonably well for mapping forest extent or urban growth that are distinct and comparatively stable, but is less effective with water because water bodies are highly dynamic both inter-annually and seasonally. Quantifying change in

surface water extent over large regional to continental areas requires a different approach.

There are two main challenges with generating accurate maps of surface water change. First is having an accurate map of the nominal (not maximum or minimum) extent of water. Second is having enough representative maps in a time series to ensure that any detected change is an actual long term change and not simply seasonal or inter-annual fluctuations in the size of the water body. This can become more problematic when the water bodies are small and shallow because they are more susceptible to short term weather effects (heavy rains or long dry spells).

Inland water accounts for between 10% and 20% of the continental surface in northern Canada depending on the source of the water map (Carroll et al. 2009; Feng et al. 2015; Lehner and Doll 2004). Previous work focused on the abundance (Carroll et al. 2009; Downing et al. 2006) and importance of small water bodies (Downing 2010) both globally and in the Arctic. Though the overall coverage of surface water is significant the impact is often overlooked in regional/global models that only consider large contiguous water bodies (Bonan 1995; Subin et al. 2012). Even studies of volumetric storage and movement of water do not completely include small water bodies (McGuire et al. 2008; Rawlins et al. 2010; Slater et al. 2007; Tang et al. 2010). Here the focus is on generating a series of maps for the purpose of quantifying change rather than simply to map the location and extent of surface water.

The primary goal of this work is to determine the extent of surface water change over a representative region of the North American tundra in north central Canada. To accomplish this goal the current study aims to 1) create annual maps of nominal

surface water extent (1985-2015); 2) quantify the trajectory of overall change; 3) assess and quantify the trajectory and trend in surface area for individual water bodies. For this work, water bodies are defined as all open surface water including lakes and rivers but excluding oceans and wetlands dominated by emergent vegetation.

2.2: Study area

This research will focus on an area in northern Nunavut territory of Canada (Figure 2.1) that is bounded by a region with upper left 105W, 70N and lower right 95W, 65N. This area is in the Southern Arctic ecoregion (Olson et al. 2001) which is characterized by wide expanses of shrublands, wet sedge meadows, low topography and small polygonal water bodies. The area is underlain by continuous permafrost (Brown et al. 1998), soils are relatively consistent throughout the area (Tarnocai et al. 2002) though the best available dataset is very coarse (1:10,000,000), it is away from most human settlements, and away from most industry (oil, gas, and other mining) (NRC 2015) all of which could be factors in a land cover analysis. These factors, if present, could influence water levels through human extraction of water from water bodies or diversion of water from rivers and streams for consumption or industrial use.

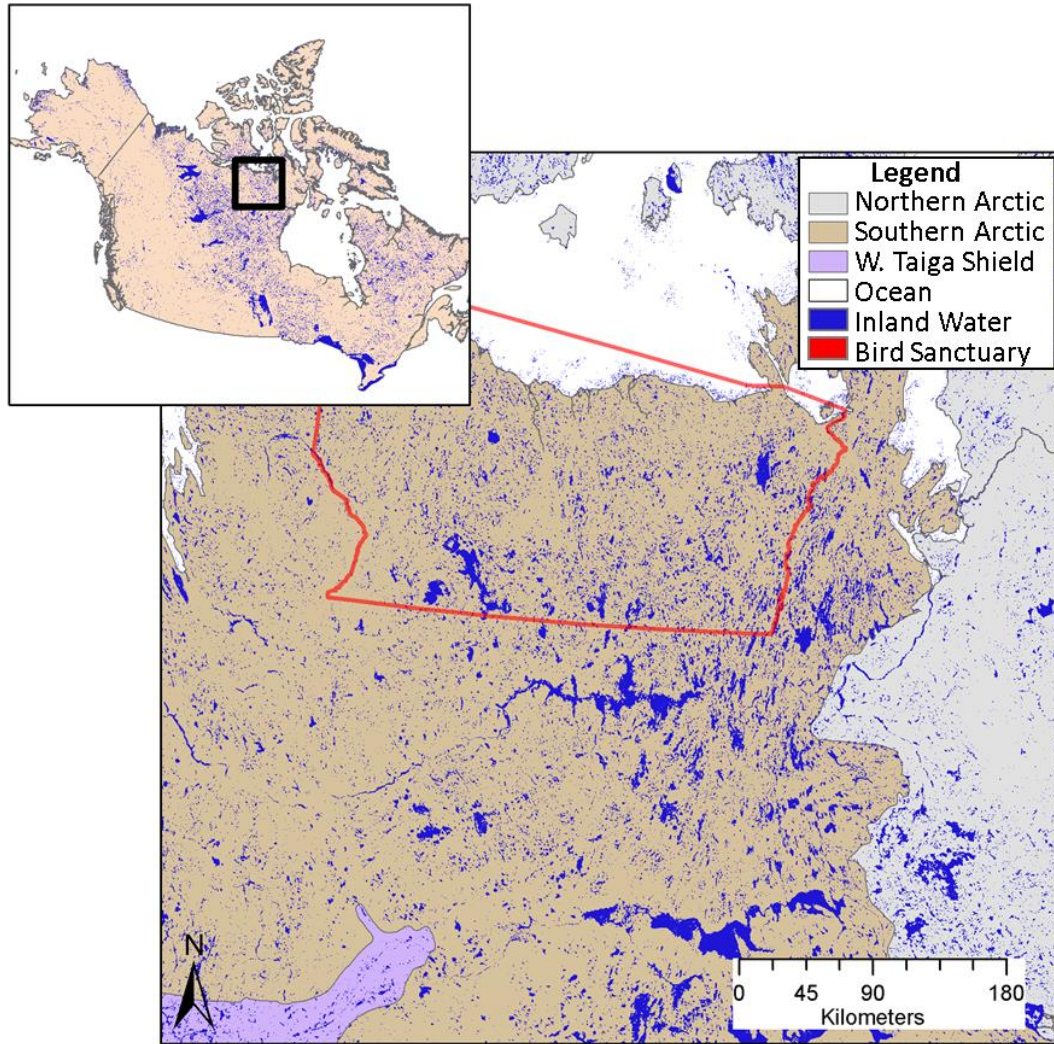


Figure 2.1 Study area in North American Arctic region, north central Nunavut territory Canada. The study area is located primarily in the Southern Arctic ecoregion and is characterized by low topography and numerous small to moderate sized water bodies. The Queen Maud Gulf Bird sanctuary is indicated with the red polygon.

Finally, the study area includes a large protected area, the Queen Maud Gulf Bird sanctuary, which also decreases the likelihood of disruptive human activity affecting the landscape.

2.3: Methods

The Landsat data archive over Canada is dense (both temporally and spatially) and is publicly available through the US Geologic Survey (USGS 2016). The

Dynamic Surface Water Extent (DSWE) algorithm (Jones 2015) uses the Landsat surface reflectance data to produce maps of surface water extent for each Landsat scene that is available in the archive. This results in a standardized mapping of water across the Landsat suite of instruments including Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+). Using the full suite of Landsat data enables the creation of a time series of maps since 1984. Annual maps of water extent can provide a rich time series to evaluate change in surface water extent over the past 31 years of the satellite record.

Mapping water in cold regions using visible, near infra-red, short wave infra-red (such as the Landsat instruments) limits the inputs to the ice-free season, defined here as June through September. Each of the Landsat satellites images every place on Earth every 16 days globally but more frequently at high northern latitudes due to the orbital overlap that increases near the poles (Goward et al. 2006). This results in a maximum possible observations of ~35 daytime observations over the study region per year. Given that cloud cover is frequent in the region (White and Wulder 2014) and the persistence of ice in the centers of water bodies the number of potential cloud free images is reduced to 3 – 8 cloud/ice free observations per year, based on evaluation of data in the current study. To ensure that there are enough observations for the maps to depict nominal (neither maximum nor minimum) extent, a rolling set of cloud free observations over three years of inputs are used to create a single map (e.g. data from 1984, 1985, 1986 used for 1985 map; data from 1985, 1986, 1987 used for 1986 map; etc.). This approach was used effectively with Landsat data to generate the decadal water maps in the high northern latitudes (Carroll et al. 2016a).

For the current study, the DSWE data product (Jones 2015) was selected as input for the annual water maps instead of the custom classification algorithm in the decadal water maps. The DSWE algorithm uses a hierarchical series of spectral tests on the visible and short wave infrared bands of Landsat to determine the presence of water. This algorithm is run on Landsat surface reflectance data that have been ortho-rectified and terrain corrected. The ortho-rectification has a stated error of approximately $\frac{1}{2}$ pixel for over 96% of the pixels (Storey et al. 2014). The Landsat data (from Landsat 5 and 7) used in this study are provided in standard Landsat WRS-2 grid (path 35 – 42 row 11 – 15). In this format there is a significant amount of overlap between adjacent scenes that increases the total number of observations but is not conducive to large area studies. For this study the DSWE data were processed into annual maps, projected and mosaicked into Canada Alber’s Equal Area projection, full description to follow.

2.3.1: Annual product generation

The “Raw” unmasked DSWE product was used as input for the annual map generation. The DSWE product has 4 output values:

- 0 Not water
- 1 High confidence water
- 2 Low confidence water
- 3 Partial water

A combination of high and low confidence water was processed through the same post classification procedure (described below) that was used to generate the decadal water maps (Carroll et al. 2016a) to produce annual maps (Figure 2.2). Though the

partial water class correctly identifies edges of water features it also seems to identify a lot of false positives and hence was treated as “not water” in the annual map creation.

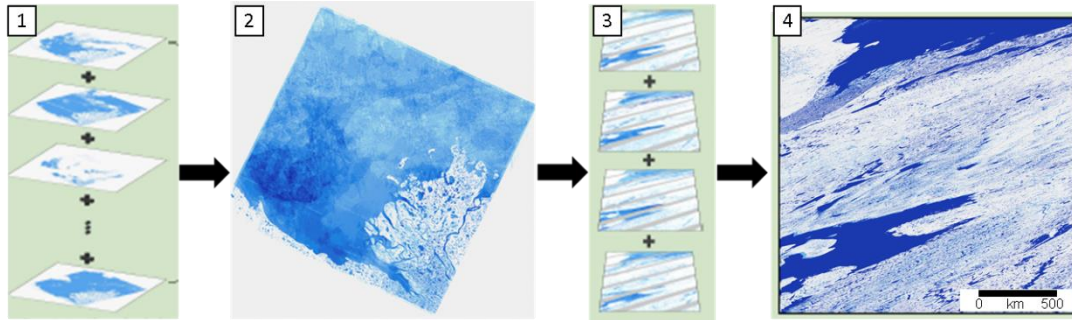


Figure 2.2 Algorithm flow for the generation of annual water maps from the individual dates of DSWE (reproduced from [28]). In panel 1 (left most panel) individual scenes are converted from four classes to two (land and water) then summed to get total observations of land and total observations of water for the period. Panel 2 shows the “total water” for an individual path/row. Panel 3 shows the mosaicked non-overlapping path/rows which are then summed. Panel 4 shows the final “total water” for the full region of interest.

Binary maps of water/ not water were made from DSWE for each time period.

Although each time period included the total of three years, more than one DSWE layer was available for each year. Moreover, the clear surface views within each scene are impacted by clouds and shadows resulting in a varying number of total valid observations for each pixel within the scene. These were summed to get “total water” and “total land” for the three year input period. Probability of water was calculated per pixel as

$$\text{Prob}_w = (W_t / (W_t + L_t)) \quad 1$$

where Prob_w is the probability of a pixel being water, W_t is the total observations of water for the period, and L_t is total observations of land for the period. Water in the annual map is anything that has greater than or equal to 50% probability of being water. Following this methodology annual maps were generated for each year from 1985 to 2015, full description can be found in (28).

2.3.2: Annual product accuracy assessment

Accuracy assessment was performed by comparing our delineations with very high resolution (VHR) data following the current best practices for remotely sensed data (Olofsson et al. 2014). Commercial VHR data are provided to NASA funded scientists through a contract with the National Geospatial Intelligence Agency (Neigh et al. 2013). Multi-spectral data from WorldView-2 was identified for 10 areas within the study region (Figure 2.3). WorldView-2 multi-spectral data has a nominal spatial resolution of 2 m and provide data in up to 8 bands (4 used in this study: Red, Green, Blue, Near-Infrared) (Globe 2016). These data were ortho-rectified using the Ames Stereo Pipeline open source software (Moratto et al. 2010). A random sample of 643 points (see Olofsson et al. 2014 for procedures for sample size determination) within the boundaries of the VHR data in the study area was chosen for evaluation. Each point was inspected by visual inspection and assigned a value of water or land.

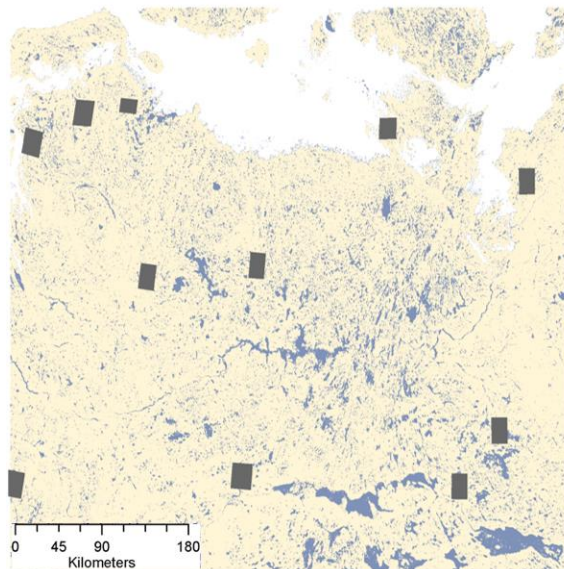


Figure 2.3 Distribution of WorldView-2 (WV2) scenes used in the accuracy assessment of the annual water maps. Footprints of WV2 scenes are shown as dark grey rectangles distributed throughout the image.

These results were compared to the 2010 annual water map because most of the VHR data used in the analysis was acquired in this year.

2.3.3: Identifying unique water bodies

The overall goal of this analysis is to determine if a trend is present for individual water bodies in the study region, thus requiring object-oriented analysis (i.e. changes in discrete water bodies). We define a “water body” as any group of adjacent pixels and water body is inclusive of lakes, rivers and ocean. Each annual raster map was vectorized and individual water body objects were defined using the “queen” adjacency rule where any pixel that touches the candidate pixel is included in the polygon and the area of the polygon was calculated. However, the highly dynamic nature of small and shallow water bodies determines that the water bodies can split, merge, disappear, or appear over time. To ensure that these changes are captured in our analysis, a master map was generated showing the extent of water over the entire 31 year study period.

For this project, annual maps were created that show land and water as 0 and 1 respectively. Calculating the sum of these 31 maps yields a single map showing the frequency of occurrence of water per pixel giving values from 0 – 31. Using the adjacency rules described earlier we identify all contiguous water bodies with a frequency of occurrence of 1 – 31. The resulting master map is the maximum possible extent of a water body over the 31-year study period. In an individual year it is possible for there to be more than one distinguishable water body but we want to consider them as part of the same water body through time to make sense of the area statistics (Figure 2.4). The master map was used to specify water body object

identifiers which were propagated through the full temporal extent of the data record. Subsequently, area of individual water body objects was calculated at the annual time step and the total change in water extent is quantified on the per object basis.

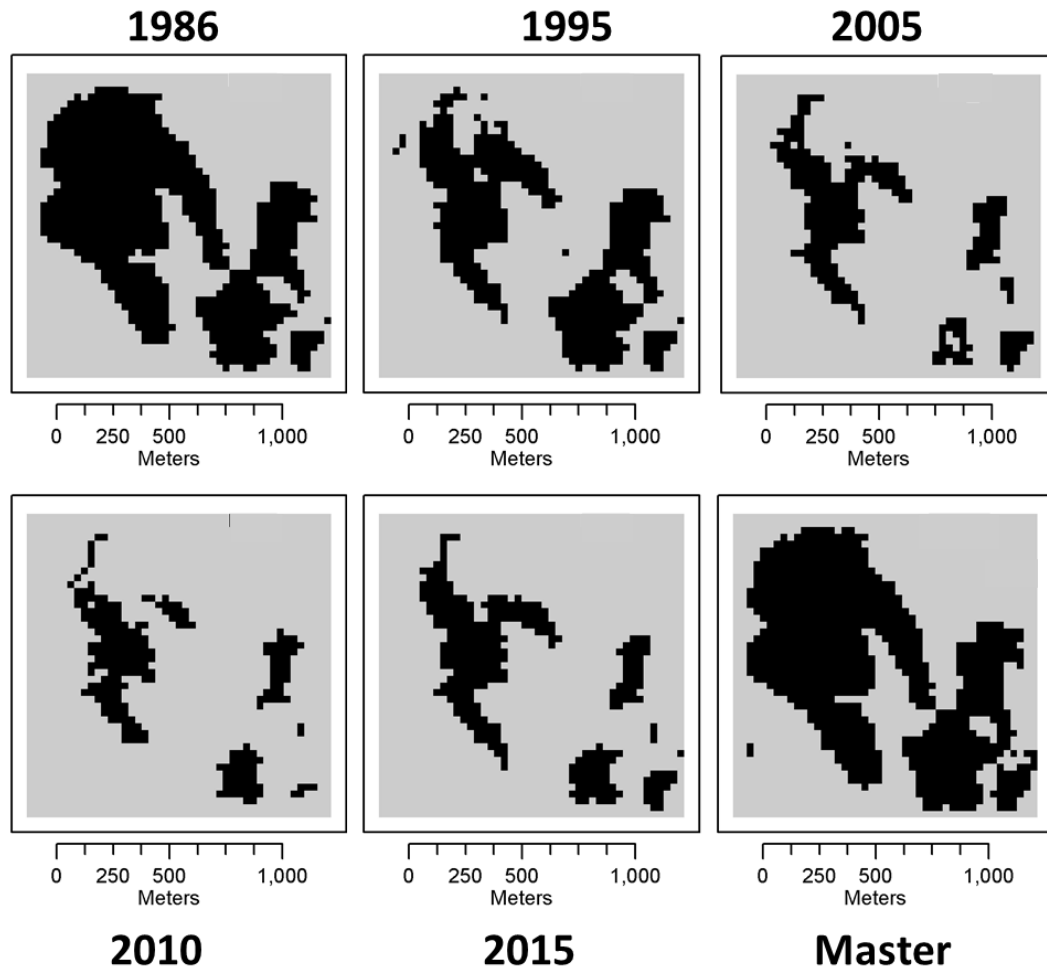


Figure 2.4 In the images above water is shown in black and land is shown in light grey. The first five images show a water body complex (small unnamed water body in northern Nunavut) in five individual years. The final image shows the master map which is the maximum extent in the whole 31 year record. Through time you see the water body shrinks and splits into several components. The master map allows all of these components to be related to the same water body even when they split off into individual pieces. The years shown here are chosen as representative examples of the 31 year record.

The combined record of annual area estimates for each of the water body objects was processed in R statistical analysis software to generate a linear regression

(ordinary least squares regression) through time. The slope, intercept, R^2 , RMSE, and p-values were calculated and recorded.

2.4: Results

2.4.1: Accuracy assessment results

The results of the accuracy assessment show that the annual water maps for year 2010 compared to WorldView-2 data represent water distribution of the tundra landscape of Northern Canada with the overall accuracy of 95% (Table 2.1). Although the producer's accuracy for water bodies was lower (87%) visual inspection of the erroneous pixels shows that the misclassified pixels occurred at the edge of a water body or in a pond smaller than one Landsat pixel.

Table 2.1 Confusion matrix for accuracy assessment of annual water maps using WorldView-2 multi-spectral data.

		Reference (from VHR)			
		Land	Water	Total	User's accuracy
Predicted (annual map 2010)	Land	486	19	505	96%
	Water	13	125	138	91%
Total		499	144	643	
Producer's accuracy		98%	87%		
				Overall Accuracy	95%

2.4.2: Long term water dynamics

The study region encompasses over 30,000,000 ha divided into two categories: land 25,820,000 ha (84.4%) and inland water 4,784,100 ha (15.6%), all values based

on average area of our maps over the 31 year study period. Oceans were eliminated from evaluation as were ocean coastlines because changes to oceans and coastlines are outside the scope of the study question. Over 675,000 individual water bodies were identified during the 31 year study period, using the master map of contiguous water bodies described earlier. The inland water bodies range in size from 0.09 ha (one Landsat pixel) to 350,000 ha based on average size over the 31 year study period (Table 2.2). The majority of detected water bodies (67%) are small (<1 ha) and the total number of water bodies decreases as size class increases.

Table 2.2 Distribution of water bodies detected in the study region by size. All sizes are in hectares and based on average area over the 31 year study period. (The very low fractional representation of the largest water bodies in the total water body count necessitates the use of precision values to the third decimal point.)

Size in ha	<0.1	0.1 to 1	1 to 10	10 to 100	100 to 1,000	1,000 to 10,000	10,000 to 100,000	>100,000
Count	251,884	202,412	167,450	48,495	4,836	257	29	9
Percent of total water bodies	37.296%	29.970%	24.794%	7.180%	0.716%	0.038%	0.004%	0.001%

Comparing maps from individual years exposes differences in extent of water bodies between those years. The total area of inland water varies annually from a low in 1999 of 4,435,692 ha to a high of 4,764,440 ha in 1989 (Figure 2.5) with a significant amount of inter-

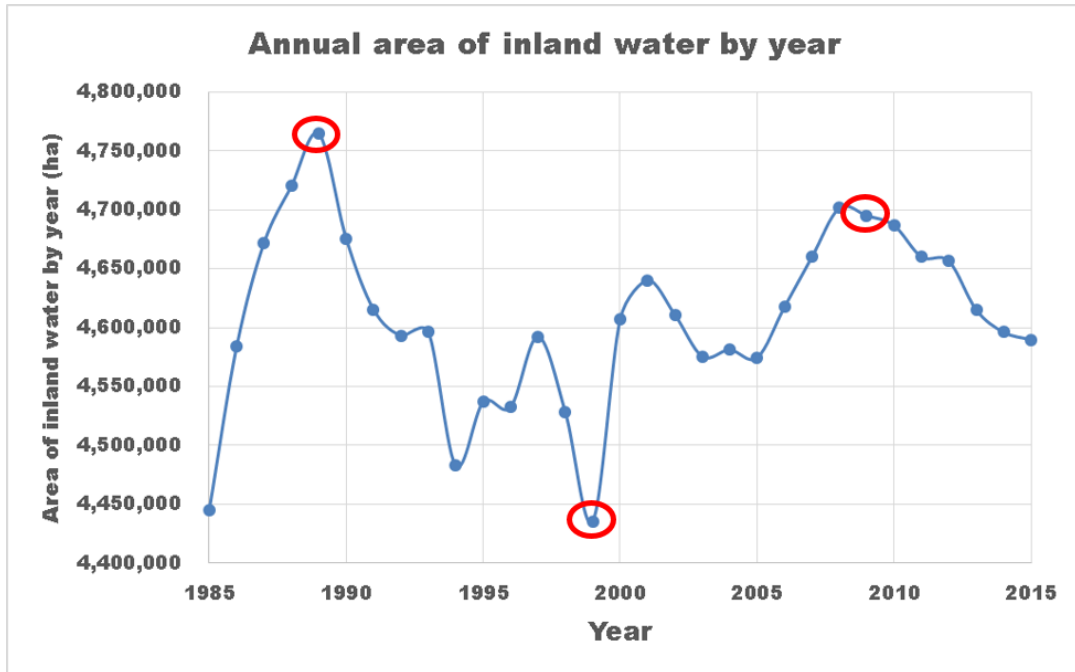


Figure 2.5 Total annual area of surface water between 1985 and 2015. Red circles denote local temporal maxima and the minimum in the record.

annual variability. The three red circles are the area values at 10-year intervals beginning with 1989 suggesting a local maximum, minimum and another maximum respectively. Water bodies that are less than 1 ha routinely fluctuate in size through the data record. By overlaying multiple years together it is possible to see the progression of change in specific water bodies (Figure 2.6). Typical fluvial processes can be seen as rivers meander and midstream islands change shape over time (Figure 2.7).

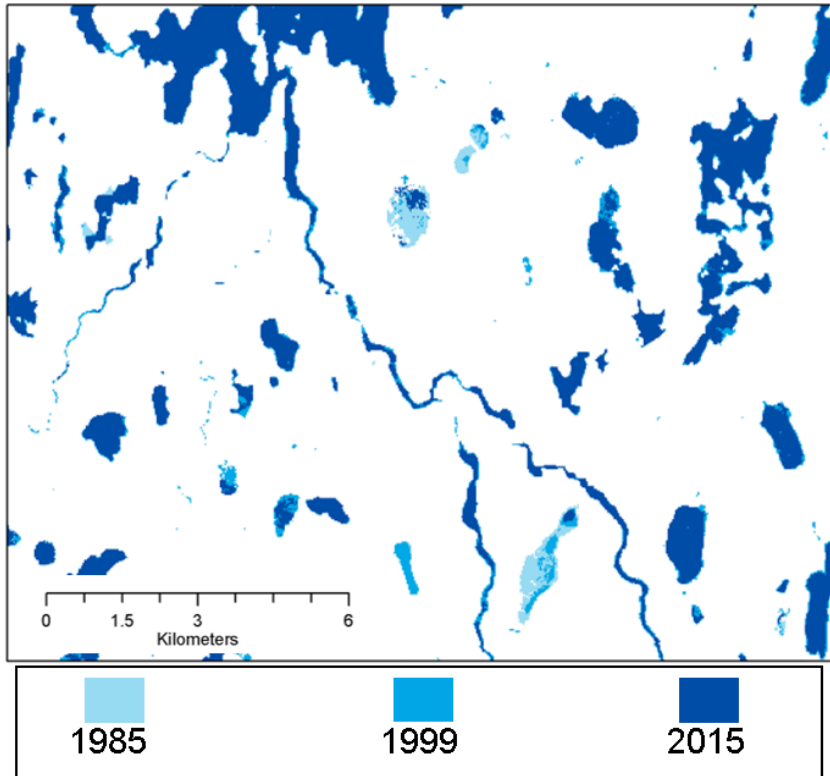


Figure 2.6 Difference in extent for several water bodies in the study region. The lighter colors indicate that water was present in early years of the study but not present in later years.

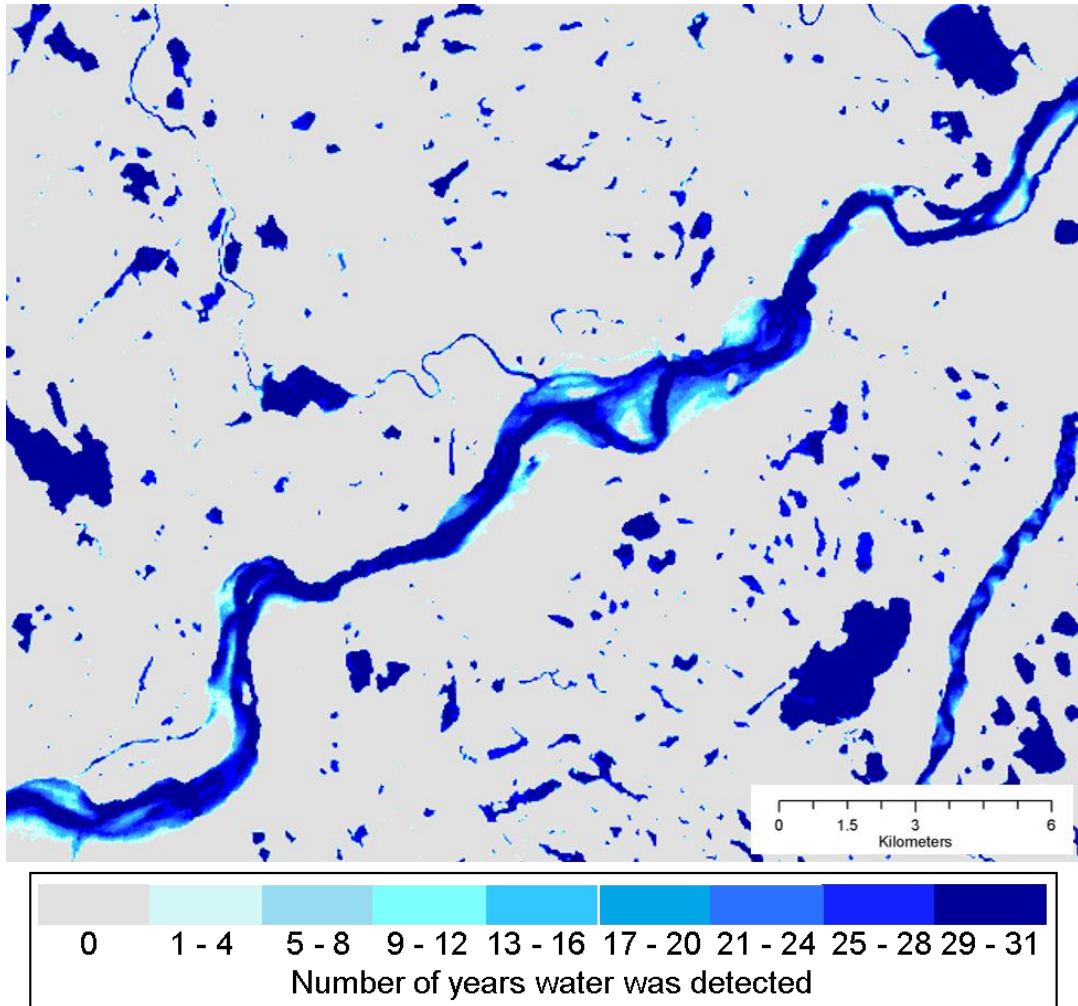


Figure 2.7 Variability of river extent and flow between 1985 and 2015. Lighter colors indicate that water was present in some years but not in all years. This is particularly noticeable in the edges of the river and in the islands in the middle of the channel.

The amount and direction of change was quantified by performing a linear regression analysis (ordinary least squares) on area of water bodies over 31 years and is reported in Table 2.3. Nearly 87% of all individual water bodies in the study region showed a trend in area with 45% showing positive trend (increasing size) and 42% showing negative trend (decreasing size). Over 168,000 (25%) water bodies that show a trend had p-value < 0.05 and nearly 73,000 (11%) had p-value < 0.01 .

Table 2.3 Overall 31 year trend in size of water bodies.

Study area analysis	Decreasing size	Increasing size	No change	Total number of water bodies
Count of water bodies	282,904	304,204	88,264	675,372
Count of water bodies with trend in surface water area with $p < 0.05$	75,988	92,059		168,047
Count of water bodies with trend in surface water area with $p < 0.01$	30,194	42,528		72,722

Small (0.1 – 1 ha) water bodies constitute the majority of water objects with significant (at $p < 0.05$) trend in surface water area (Table 2.4). This distribution is similar to the overall size distribution of water bodies (Table 2.2). The results show that for the majority of water bodies, specifically those between 0.1 and 1,000 ha in size, the fraction of bodies that exhibited a significant trend in surface water area at $p < 0.05$ is approximately 27%. This fraction is smaller for very small (< 0.1 ha) bodies at 21% and large water bodies at 24%, 14%, and 22% for 1,000 – 10,000 ha, 10,000 – 100,000 ha, and $> 100,000$ ha, respectively. Breaking this down further into water bodies that are increasing compared to those that are decreasing more differences in the trends become apparent.

Table 2.4 Distribution of water bodies that exhibit significant trend in surface water area at $p < 0.05$ by size and trajectory of change.

size in ha	<0.1	0.1 to 1	1 to 10	10 to 100	100 to 1,000	1,000 to 10,000	10,000 to 100,000	>100,000
Count of water bodies with trend in surface water area at $p < 0.05$	52,475	55,081	45,724	13,330	1,369	62	4	2
Fraction of total water bodies by size	21%	27%	27%	27%	28%	24%	14%	22%
Count of water bodies decreasing in size	31,810	21,438	17,216	4,960	527	32	3	2
Count of water bodies increasing in size	20,665	33,643	28,508	8,370	842	30	1	0
Fraction of water bodies decreasing	61%	39%	38%	37%	38%	52%	75%	100%
Fraction of water bodies increasing	39%	61%	62%	63%	62%	48%	25%	0%

The smallest size class shows a distinct trend toward decreasing in size: 61% of all water bodies in size <0.1 ha that exhibit a significant trend decreased in extent over time. Since a single Landsat pixel is ~0.09 ha any water body that is <0.1 ha that decreased in size either disappeared or became too small to be detected by Landsat anymore. In comparison, the next three size classes show a distinct trend toward increasing in size at approximately the same proportion (61 – 63% of water bodies that exhibit a significant trend in each category). Approximately half of the medium (52% of the 1,000 – 10,000 ha) and most large (75% and 100% of 10,000 – 100,000 ha and >100,000 ha, respectively) water bodies with significant trend have decreased in size over time. The analysis here is possible because the use of the

master map prevents the counts from inflating artificially when several water bodies coalesce into a single water body or conversely a single water body dries and splits into two or more water bodies. The master map keeps those water bodies together throughout the analysis.

The water bodies that show a significant change can be found throughout the study area (Figure 2.8). Water bodies that exhibit a trend toward growth are colored

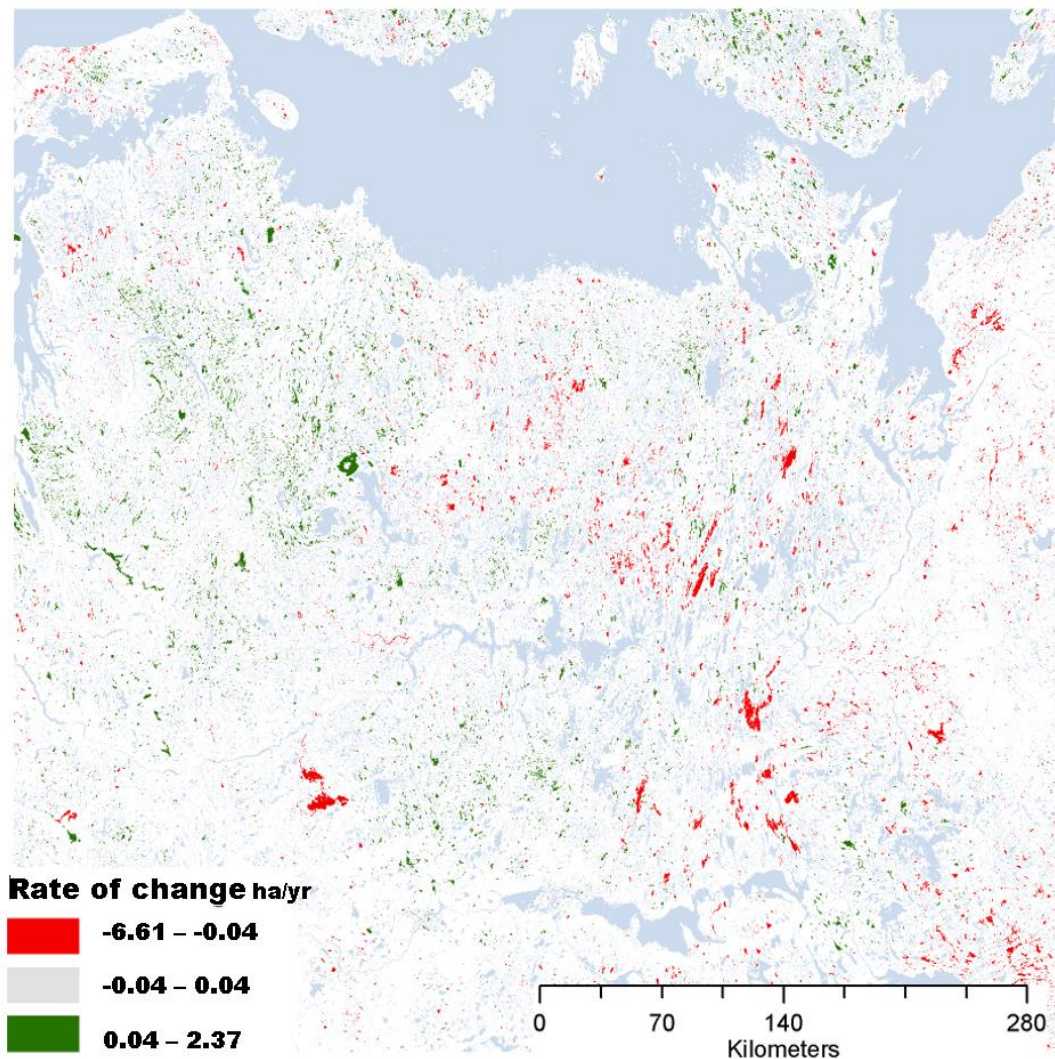


Figure 2.8 Spatial distribution of water bodies with a significant ($p < 0.05$) trend in surface water area (ha/yr) over the 31 year time period of the study. Water bodies that are increasing in size are shown in green while the ones that are decreasing in size are shown in red.

green, those that show a trend towards shrinking are colored red and those water bodies that have no significant trend are light blue-grey. There are water bodies that are increasing in size that are spatially adjacent to water bodies that are decreasing in size. It is also apparent that the water bodies are distributed throughout the study region regardless of size, i.e. small water bodies and large water bodies can be found anywhere in the study region.

2.5: Discussion

In this study, the annual maps are generated from a rolling set of three years of inputs which results in a map of nominal extent of surface water over the period of the inputs. By doing this the likelihood that an individual map is showing maximum or minimum extent is reduced thereby increasing the validity of inter-comparison with other maps produced in the same way, i.e. the time series created here. Accuracy assessment was performed on the map for the year 2010 to closely match the VHR data that were used for the assessment. It is assumed that the accuracy is similar for maps from other years because the input product (DSWE) and the methodology remain the same.

Figure 2.4 suggests that the overall area of surface water changes from year to year. This clearly shows that studies that use any two years as inputs and employ simple differencing to identify change will inevitably produce erroneous results. The majority of water bodies that show significant trends are growing in size at an average rate of 0.03 ha per year. The net trend in area of surface water per water body is 0.009 ha calculated as average trend in growth minus average trend in decline. This finding is consistent with predictions that as the Arctic warms the active layer of

permafrost deepens (Jorgenson et al. 2010; Lawrence et al. 2008). As the active layer deepens, this can result in the erosion of the margins of a water body often resulting in expansion of the water body. Conversely, deepening of the active layer could result in the collapse of the water body bottom which would likely cause the draining of the water body (Hinzman et al. 2005). The changes to a water body by deepening of the active layer are not necessarily uni-directional, in some cases a water body could initially expand until a critical threshold is exceeded at which point an edge fails or the bottom falls out and the water body drains (i.e. thermokarst).

The novelty in this analysis comes in the object based analysis that tracks water body area through time by spatial location. Through this analysis any water body that is shown to disappear has not been absorbed by the growth of a larger immediately adjacent water body. Table 2.4 shows that size of water body plays a role in the direction of change (increasing or decreasing) where the smallest water bodies (<0.01 ha) are decreasing in size over the study period and the remaining small water bodies (< 1 ha) are increasing in size. This could imply that water from the smallest water bodies is moving laterally into larger adjacent water bodies or it is exiting the system in a different way (possibly evaporation or drainage into ground water). However, determining exactly where the water is going is beyond the scope of this study.

While there are specific instances of complete drainage of a water body or the creation of new water bodies these instances are rare. The more interesting result is the quantification of a significant trend showing expansion of surface water area over the long time series. While the changes are distributed throughout the study region (Figure 2.8) there appear to be regions of increasing and decreasing water bodies.

The water bodies that are increasing appear to be concentrated in the northwest portion of the image and the water bodies that are decreasing appear to be concentrated on the eastern half of the image. Further investigation with ancillary data such as ecoregions, soils, elevation and vegetation cover may provide additional insight into the potential patterns of change. The implication of wide spread growth in surface water lies in the extensive stores of carbon in the soils and sediments in the region. Greater surface water area can lead to (or be caused by) deepening of the active layer beneath and around the water bodies. One potential implication of changing surface area of water is the possibility of mobilizing the significant amount of soil based carbon to the atmosphere through methanogenesis, denitrification, aerobic decomposition/respiration and other biogeochemical processes.

The methodology used to create the annual water maps from a series of water classifications is robust and has been employed for both Landsat and MODIS instruments (Carroll et al. 2016a; Carroll et al. 2009) and is similar to the method adopted for the Global Landsat Water maps (Pekel et al. 2016). The development of these new time series products of surface water make it possible to apply our object based analysis methodology in other regions. The challenge arises in large regions with a rapid increase in the number of objects. The current process with over 600,000 objects pushed the limits of what a standard software package (ArcGIS in this study) would handle. Significant improvements in data handling in this software or the development of custom routines that are outside of proprietary software are needed to facilitate processing larger numbers of objects. Ideally, these new processes would be developed to take advantage of “Big Data” processing environments and parallel

processing to both speed up the processing time and enable larger datasets to be considered.

The DSWE water detection algorithm was originally designed and tuned to detect water in the contiguous United States. This is the first application of the algorithm in the tundra ecotone and the algorithm has performed well as demonstrated by the validation presented in this text. Future enhancements to the DSWE algorithm including improvements to the “partial water” class will have important implications for advancing the methodology presented here to support studies that quantify carbon exchange in the tundra. In addition, the identification of a snow/ice coupled with improved cloud detection would increase the overall number of observations that could be used. These improvements have been suggested to the developers of the DSWE and are currently under consideration.

Previous analyses of water body change in the Arctic have focused on only a few observations and there has been more work done in the Alaskan North Slope than the Canadian high Arctic (Bouchard et al. 2013; Brown and Young 2006; Chen et al. 2013; Hinkel et al. 2007). Coarse resolution analysis using MODIS data (Carroll et al. 2011b) was limited to differencing over a decade and at 250 m spatial resolution does not fully capture the dynamics because so many of the water bodies are well below detection levels with MODIS (1 MODIS pixel at 250 m resolution ~6.25 hectares) see table 2.4 showing over half of the water bodies with a significant trend are 1 hectare or less. Coarser resolution analysis (25 km) using microwave data (Watts et al. 2012) has the advantages of fractional coverage and the ability to collect data under cloudy conditions but is still too coarse to capture the dynamics of small

water bodies. Most recently global annual water maps have been generated from Landsat (Pekel et al. 2016), however the change mapping provided combines all maps up to 1999, all maps after 1999 (yielding two maps) and then does a simple difference. Based on the information in figure 2.5, this approach is likely to lead to erroneous results at least for the current study region. The ability to use the full time series of Landsat data to generate a time series of object based (individual water bodies) measurements provides a new perspective on the true dynamics of water bodies in the region. The density and diversity of remote sensing observations has increased with the launch of Sentinel 1A/B (radar observations) and Sentinel 2 A/B (10 – 20 m visible and near infrared observations). Combining the historical time series with the current and potential future observations from the Sentinel constellation will provide an extended record into the future.

2.6: Conclusions

This is one of the first analyses to produce annual water maps that show the nominal water extent in each year over 31 years for a region. A large inter-annual variability in the results demonstrates that it is necessary to do time series analysis to fully understand the dynamics of surface water in the region. Size analysis of water bodies confirms previous studies that show that small water bodies dominate the landscape. Our results show that over the past 31 years the smallest water bodies have been shrinking or disappearing entirely whereas slightly larger to moderate water bodies have been increasing in size. The area based linear regression analysis shows an overall net growth of $0.009 \text{ ha year}^{-1}$ per water body. The occurrence of significant trend in surface water area in the study area, which is located far from

most anthropogenic activity, implies that environmental or climate factors are driving the observed change. Further investigation is required to determine if the potential periodicity seen in the graph of total area of surface water over 31 years (i.e. figure 2.5) is related to a climate driver or other environmental factors and an analysis of spatial patterns and clustering will reveal potential areas for further investigation.

Chapter 3: The sign, magnitude and potential drivers of change in surface water extent in Canadian tundra

3.1: Introduction

The North American High Northern Latitudes are currently the subject of intensive research focus in both Canada and the United States. Over the past 30 years air temperature in this region has increased faster than in the rest of the world with an average of 2 – 4° C (Miller et al. 2010; Serreze and Barry 2011) increase over that time span. The observed warming has been identified as the primary driver of a broad spectrum of environmental changes from enhanced vegetation response (i.e. “greening”) (Goetz et al. 2010; Ju and Masek 2016; McManus et al. 2012), to the deepening of the active layer (seasonally melted layer above permafrost) and talik formation (unfrozen layer of soil beneath lakes and above the permafrost) in the Arctic (Arp et al. 2016; Brown et al. 1998; Camill 2005; Jorgenson et al. 2010), and to an overall trend toward lower sea ice cover (Stroeve et al. 2011). Accounting for more than 20% of the land surface across the Arctic tundra, water bodies are a critical component of tundra ecosystems (Downing 2010; Downing et al. 2006). These water bodies are essential to the hydrological cycle, carbon cycle and overall energy balance in the region (Bring et al. 2016; Cole et al. 2007; Francis et al. 2009b; Stokstad 2004; Walter et al. 2007; White et al. 2007). Not surprisingly, numerous recent studies have focused on extensive changes across various components of the hydrological cycle including longer inland water ice-free season (Chen et al. 2013; Smejkalova et al. 2016), increased inland water temperatures (O'Reilly et al. 2015; Schneider and Hook 2010; Sharma et al. 2015), and a change in the overall extent of

surface water across High Northern Latitudes (Hinzman et al. 2005; Jepsen et al. 2013; Jorgenson et al. 2006; Roach et al. 2013; Rover et al. 2012; Smith et al. 2005; Smol and Douglas 2007b).

Many factors including changes in temperature, precipitation, evaporation, vegetation cover, and subsurface soil structure affect the surface water extent of inland water bodies (Bring et al. 2016; Gibson and Edwards 2002; Yoshikawa and Hinzman 2003). Incoming precipitation is balanced by runoff, surface storage, evaporation, and subsurface drainage. In High Northern Latitudes (HNL), however, permafrost prevents surface water from percolating into ground water and thus largely controls subsurface drainage (Jorgenson and Grosse 2016; Jorgenson et al. 2010; Kokelj and Jorgenson 2013). In regions of permafrost this, in general, results in water ponding in depressions where there is not enough slope or an outlet to permit overland flow of water. Surface water change in discontinuous permafrost has been shown to be primarily drying due to connection between surface water and subsurface water. Whereas surface water change in continuous permafrost has been more complex to characterize with the hypothesis of initial wetting followed by drying due to talik formation under the water bodies (Smith et al. 2005).

Many studies involving the use of satellite imagery have been conducted to ascertain the direction and magnitude of surface water extent change in the HNL over the past 30 years. Most studies of surface water extent change are performed using a thresholding method to classify an image pair (or a few dates) to get the difference in water extent between those dates (Hinkel et al. 2007; Smith et al. 2005; Yoshikawa and Hinzman 2003). However, those methods that pick a difference between two

arbitrarily selected points in time are likely plagued with considerable uncertainty due to a naturally large inter-annual variability of water extent as demonstrated in Carroll and Loboda (2017). Better approaches focus on multi-temporal assessment of long-term trends in surface water change. A time series of Landsat data (one image per year) was used to explore long term change in Yukon Flats, Alaska (Rover et al. 2012) which was subsequently used to investigate natural variability in extent of lakes (Chen et al. 2013). The most robust currently available methods utilize the full archive of Landsat imagery to account for seasonal variability of surface water extent in addition to the inter-annual variability and obtain the long-term trends in nominal surface water extent as compared to potentially abnormally seasonally high or low gained from a single image per year (Carroll and Loboda 2017; Pekel et al. 2016).

The approaches to analyzing the resultant change in surface water extent vary. The global analysis from Pekel et. al. (2016) compares the difference in a 15-year mean extent of water bodies before and after the year 2000: annual maps from 1984 – 1999 and 2000 – 2015 were combined to produce two period maps, respectively, which were then differenced to produce change. Another example represents a regional study with four sites (two in Alaska and two in Siberia), that used Landsat time series with a machine learning approach to perform a per pixel classification into four classes (stable water, stable land, water to land, land to water) based on the trend in spectral indices over time (Nitze et al. 2017). Lastly a regional study in northern Nunavut (Canada) used the full time series of Landsat to generate annual maps of the nominal surface water extent. These maps were converted to polygons and used to generate a trend in areal extent for each water body (Carroll and Loboda 2017).

Unlike the methods from Carroll and Loboda 2017 and Pekel et al. 2016, the method from Nitze et al. 2017 produces a single output map that shows the classification with change as a class rather than annual maps that are used in a secondary analysis for change.

Although mapping and quantifying surface water extent change over time at the regional scale is sufficiently challenging, developing an understanding of the drivers of the observed changes at this scale is particularly difficult since datasets representing subsurface parameters are largely unavailable or are grossly oversimplified. The studies that have examined the mechanisms driving surface water extent are limited to local areas with local maps or field measurements (Rover et al. 2012) or local airborne data (Jepsen et al. 2013; Minsley et al. 2012). These studies both concluded that subsurface composition, specifically coarse gravel, was a better determinant of water body change than the overall depth of talik. In a study on Ellesmere Island it was found that water bodies on exposed bedrock disappeared completely in the summer of 2006 (Smol and Douglas 2007a). Maps of permafrost (Brown et al. 1998) and soils (SLCWG 2010; Tarnocai et al. 2002) are coarse in spatial resolution, hence provide little ability to discriminate drivers of change in specific water bodies. Considering the remoteness and difficulty of access for most of circumpolar tundra, large scale studies based on observed (rather than modeled) land surface conditions rely on satellite-derived data products. One approach to developing an understanding of the regional scale drivers is to use other directly observable land surface parameters as a proxy for more general environmental conditions and subsurface characteristics. Specifically, land cover and vegetation

fraction – both routinely mapped from satellite observations – can to some degree reflect the type of soils and the depth of active layer in otherwise uniform climatic conditions. Datasets describing land cover (Didiuk and Ferguson 2005; Friedl et al. 2010; Olthof and Fraser 2014) and vegetation (Carroll et al. 2011a; Hansen et al. 2003; Sexton et al. 2013; Townshend et al. 2012) are available both globally and regionally.

In this paper we use an existing surface water change map (Carroll and Loboda 2017) to explore potential drivers of that change at the regional scale in Nunavut, Canada. Maps of land cover at coarse (DiMiceli et al. 2011; Friedl et al. 2010; Walker et al. 2005) and moderate (Didiuk and Ferguson 2005; Olthof et al. 2014; Wulder et al. 2008) spatial resolution are used to derive proxies for ecological factors from land cover type. The specific objectives of this study are to 1) quantify the spatial relationships and sign (increasing or decreasing) of change shown in Carroll and Loboda 2017, and 2) to relate that change to ecological factors observable at the regional scale including weather station data and various land cover types.

3.2: Study area

The study area is located in Nunavut territory primarily in north central continental Canada and extends into the Queen Maud Gulf including parts of several islands (figure 3.1) bounded by the geographic coordinates 69.2 N, 108.6 W, and 63.5 N, 93.6 W. This area falls within the Arctic Tundra biome (vegetation limited to mosses, grasses, sedges and shrubs) underlain by continuous permafrost (Brown et al. 1998). The study area is characterized by low topographic relief and the best available DEM (Canadian Digital Elevation Model used in this study) is at 1:50,000

spatial resolution (~30 m). The average annual temperature and precipitation range from -11° C to -14° C and 273 mm to 148 mm from south to north respectively (CCN 2017).

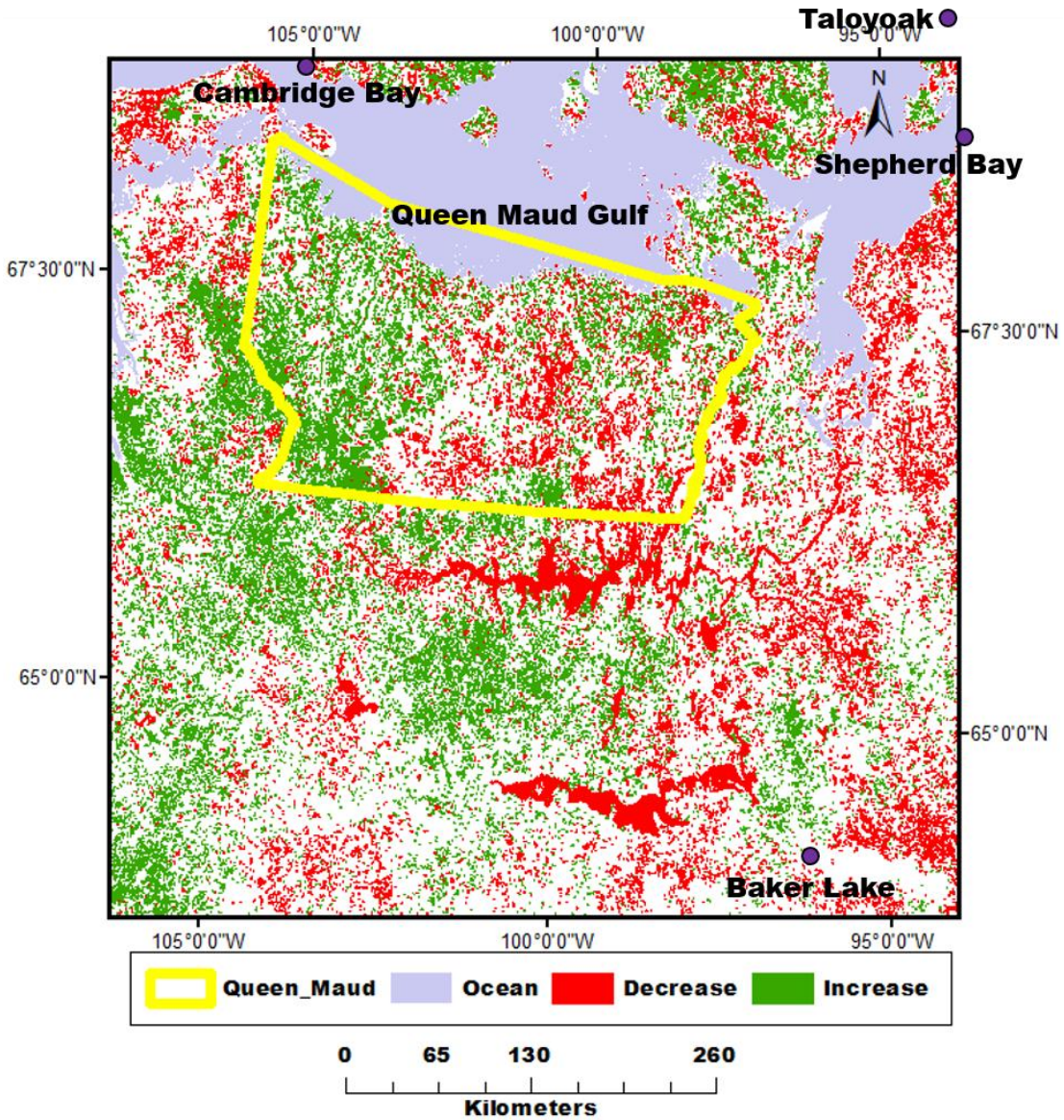


Figure 3.1 Study region in northern Canada, Nunavut territory where multi-decadal trends in surface water extent have been established (Carroll and Loboda, 2017). Water bodies in the region show opposite trends of surface water extent. Red indicates a water body that is increasing in extent while green indicates a water body that is decreasing in surface water extent. There are four weather stations within or just outside of the study region indicated by purple circles.

There is limited infrastructure with no major cities or roads in the study domain and the most prominent feature is the Queen Maud Gulf Bird Sanctuary on the northern coast. This sanctuary is the largest federally owned protected area in Canada covering over 61,000 km² and is listed in the Ramsar Convention on Wetlands of International Importance. The study area boundaries were selected such that the sanctuary covers approximately half of the continental land area in the study region. This is significant because the protected area designation puts strict limits on the amount of human activity that can be in the protected area, so any changes that are seen here are less likely to be caused by direct human intervention such as dam building or other infrastructure construction.

3.3: Data and Methods

In our previous work in this study area of North American tundra (Carroll and Loboda 2017) we mapped over 675,000 water bodies, which constitute over 20% of the land cover (excluding ocean), with over 25% of water bodies showing a significant increasing or decreasing trend in surface extent over time ($p < 0.05$). The goal of the current work is to put the previously identified trend in surface water extent into context with landscape and ecology of the region as well as other previously reported changes including the recent findings regarding vegetation greening and browning trends (Goetz et al. 2010; Ju and Masek 2016; McManus et al. 2012). We conducted a review of available regional-scale land cover and land cover change data to identify key data sets that we used in this analysis (see supplemental section for dataset review).

From our previous work identifying trend in surface water area per water body, we identified the actual areas of observed change (individual per pixel observations) within the water bodies that showed significant increasing or decreasing trends. These grid cells were dichotomized into decreasing and increasing classes by assigning values of -1 for water bodies showing negative trend and 1 for water bodies showing positive trend irrespective of the actual magnitude of the trend. Ocean (Queen Maud Gulf) was masked from all analysis using its maximum extent between 1984 – 2015 (Carroll and Loboda 2017) to eliminate the coastal shoreline change which is governed by different ecological processes and does not represent the focus of this study. Through the process described above, we identified over 4.3 million grid cells representing ~3870 km² of area with statistically significant trend in long-term surface water extent change within our study region.

3.3.1: Identifying spatial patterns of change

With surface water change identified and quantified down to the grid cell level the next challenge is to identify the driving ecological factors. First order analysis involved identifying spatial patterns of observed dichotomized change. A straightforward way to assess spatial grouping of data is to aggregate from finer resolution to coarser resolution to reveal spatial clustering of areas of change (Lam and Quattrochi 1992). In this case, we took the positive and negative change at 30 m resolution and averaged to 3 km resolution and recorded the percent of grid cells showing change within the resulting 3 km grid cells. Because each 30 m grid cell has a sign (positive or negative) when the averaging occurs, the end result is a “net” change within the grid cell where equal number of positive and negative values will

result in a “0” value within the 3 km grid cell. Spatial aggregation to coarser resolution provides a systematic way to assess the spatial association of nearby change irrespective of other local-scale underlying ecological processes.

Another way to identify spatial patterns is to group the change in terms of other meaningful spatial subsets. We used watershed boundaries to provide a meaningful ecological stratification of the landscape for regional-scale analysis. Surface water area and surface water change were aggregated per watershed, defined by the Canadian National Hydro Network (Natural Resources Canada 2007), using zonal statistics. This produces percent surface water area and percent surface water change per watershed. As in the 3 km percent change, the change per watershed is net change because the sign of change is used in the calculation of area of change.

3.3.2: Random forest analysis

With several datasets available describing surface features it was necessary to determine which datasets are statistically related to surface water change. Random forest analysis (RFA) (Breiman 2001) is a machine learning technique that facilitates assessing the combined influences of multiple inputs on the dependent variable. Unlike many other multivariate statistical models, RFA is not sensitive to the number of independent variables and readily allows for inclusion of both discrete and continuous variables in the analysis. A byproduct of the RFA is a ranking of the independent variables in the context of importance to the RFA. We used this feature of RFA to identify key variables for determining where surface water change will occur.

The 4.3 million individual pixels showing change were used as the dependent variable in the RFA. Values for Terrestrial Ecoregions of the World (TEOW), Circum Arctic Vegetation Map (CAVM), Soil Landscapes of Canada (SLCWG), National Hydrography Network-watershed (watersheds), National Hydrography Network-connectivity (connectivity), Canadian DEM-elevation (elevation), Canadian DEM-slope (slope), and MODIS percent bare (bare) were extracted under the raster pixels of change into a table and used as the independent variables in the RFA. A random sample of 10% of the pixels (~430,000) from the table was used in successive runs of the random forest with increasing numbers of trees starting with 5 and increasing to 1000 in 30 runs.

3.3.3: Relating spatial patterns of change to ecological drivers

Weather- specifically temperature and precipitation - is a critical factor controlling surface water extent with particularly complex conditions in cold regions where ice cover plays a role. Temperature and precipitation data from four weather stations, figure 3.1, were examined in this study (<http://climate.weather.gc.ca/>). To establish the overall trend per weather station the daily temperature data was averaged to a single annual average value per year consistent with previous studies (Easterling et al. 1997; Jones et al. 1999). Most years had between 0 and 7 missing days of data, any year that was missing 10% or more of the observations was set to no data and was excluded from further analysis. Of the four weather stations used only two (Cambridge Bay and Baker Lake) had a consistent record of precipitation while the other two had significant gaps in coverage and were excluded from analysis. The precipitation data (total daily precipitation including frozen precipitation) was

summed for each year to produce total annual precipitation for each year from 1985 – 2015.

There are a number of datasets that describe the land surface of the study region. Many of these land cover characteristics can be used as proxies for ecological parameters. We selected datasets that describe cover type (land cover maps), vegetation condition (Normalized Difference Vegetation Index (NDVI)), elevation and slope, and surface water extent/change across the full study area. A full description of datasets that were investigated is included in the supplemental section. All variables that were used in this analysis provide a continuous representation (i.e. percent, trend, or fraction) except the three classes from the land cover of Canada that describes areas that are either exposed or lichen covered bedrock (henceforth “barren”). Grid cells that were identified as barren were dichotomized into a separate dataset with two classes – barren and vegetated. All datasets were aggregated at the watershed level using zonal statistics retaining the average (for continuous) or percent (for discrete) variable per watershed. Univariate statistical relationships between individual surface characteristics and surface water change as the dependent variable were subsequently tested using R^2 values for each variable to determine the goodness of fit.

3.4: Results

Initial evaluation of the map of surface water change at 30 m resolution (figure 3.1) suggests that there are groups or areas within the map where water bodies with positive trend or negative trend are clustered. The clustering becomes more apparent with aggregation of the original map to 3 km grid cells shown in figure 3.2. The clusters of significant surface water change in the right panel of figure 3.2 do not simply mimic locations where there are water bodies in the left panel, but rather they show a distinct and different spatial pattern. Visually there does not appear to be a

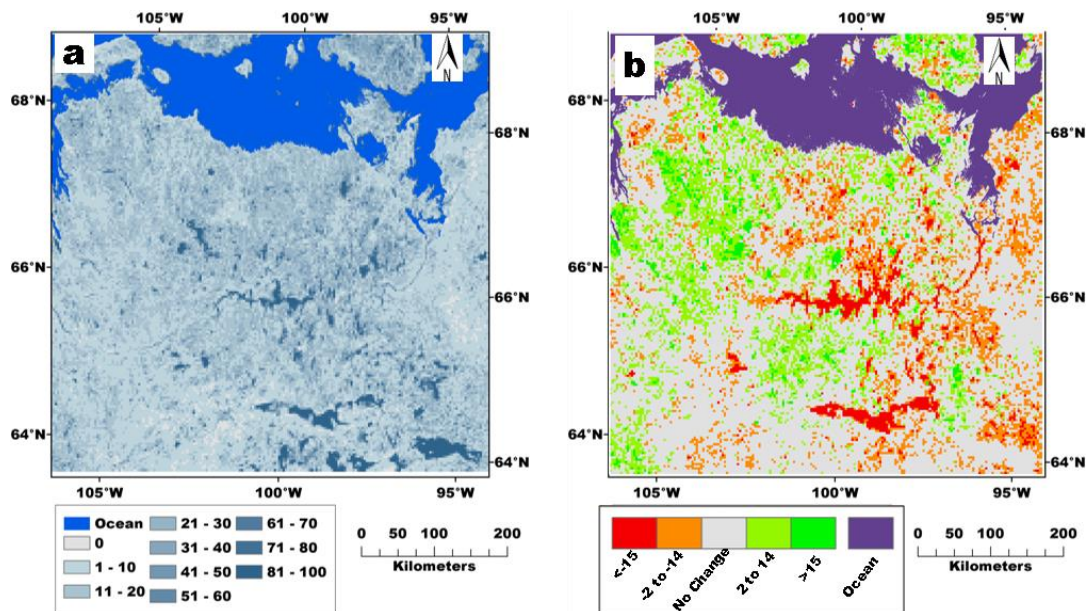


Figure 3.2. The left panel shows percent surface water at 3 km spatial resolution with darker shades of blue corresponding to higher percentage of water. The right panel shows percent water change at 3 km spatial resolution. Green colors indicated a higher percentage of water bodies with increasing extent. Red colors indicate a higher percentage of water bodies with decreasing extent.

relationship between the amount of surface water area per watershed and surface water change per watershed. The quantitative assessment also shows a poor correlation ($R^2 = 0.01$) between surface water abundance in the watershed and

change. Overall figures 3.2 and 3.3 highlight a pattern of change in the study region that goes from north-west to south-east.

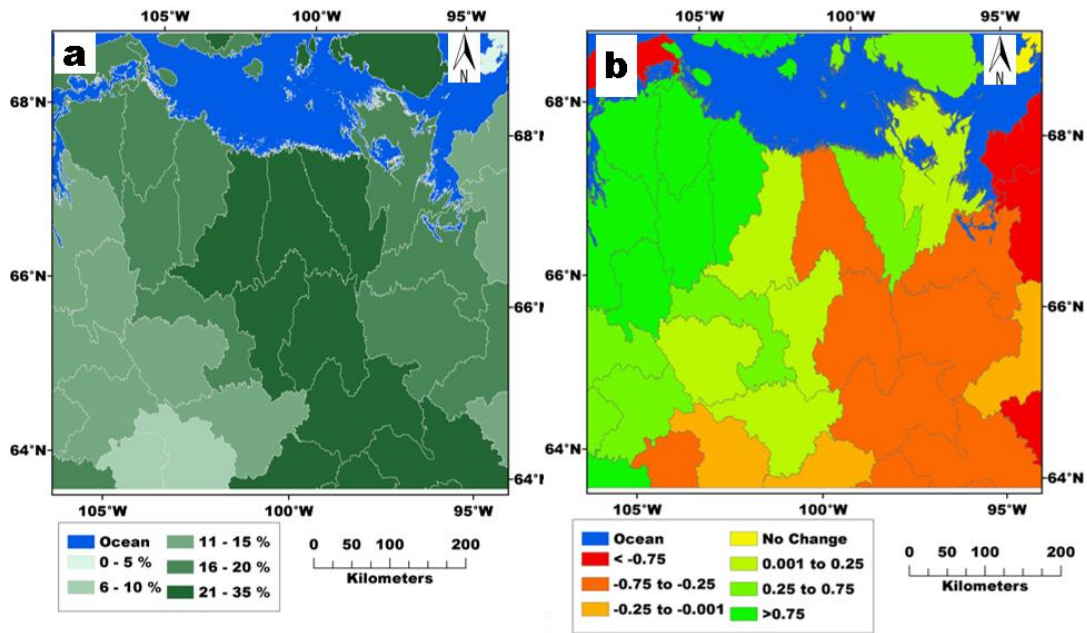


Figure 3.3 The study region divided into watersheds using the Canada National Hydro Network. Figure 3a on left shows the percent surface water in each of the watersheds. Figure 3b on right shows the percent net surface water change per watershed. Both results determined by using the annual surface water extent product (Carroll and Loboda, 2017).

Weather is a fundamental driver of surface water change where year to year differences in amount of precipitation can substantially affect the surface water extent for climatologically short periods of time. The purpose of the current study is to identify long term trends in surface water extent and their drivers so we approached analysis of weather data with a similar long term approach. We evaluated temperature and precipitation records from the four official weather stations that are within, or just outside of, the study region with records of at least 30 years (figure 3.1). Taloyoak, Shepherd Bay and Cambridge Bay stations are located more than 500 km (~6° of latitude) to the north of Baker Lake station and thus show substantially

lower mean annual temperatures. The annual temperature trend at three of the four stations is positive over the available data record with an average increase of $\sim 2^{\circ}$ F across the study region (figure 3.4). This rate of increase is consistent with reports of increasing temperatures throughout the Arctic (Miller et al. 2010; Serreze and Barry 2011). The station at Shepherd Bay shows a slight decline in temperatures over the study period. The negative trend at the Shepherd Bay station is strongly influenced by the comparatively high mean annual temperature in the late 1980s. However, a trend beginning from the early 1990s would show a similar direction and magnitude to the trend recorded at all other stations in the study area. Averaged over the four stations, the regional trend in mean annual temperature remains positive (figure 3.4e).

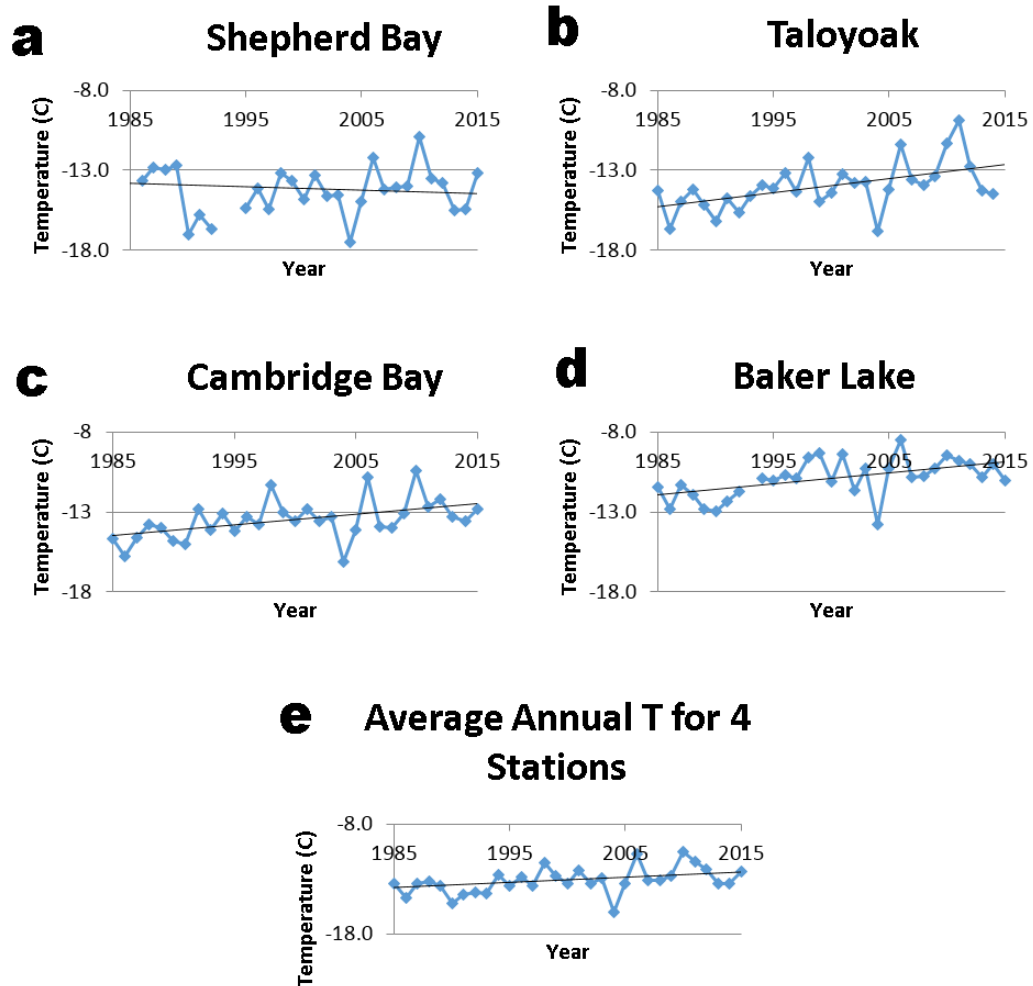


Figure 3.4 Annual temperature trends for four weather stations within or in close proximity to the study domain showing individual trends (a - d) and an average of all four stations (e). The overall trend for the domain is increasing temperature at a rate of 0.66° C per decade.

The annual precipitation, available from only two stations, shows divergent trends for Cambridge Bay (slightly increasing) and Baker Lake (slightly decreasing) (figure 3.5). Neither trend was significant; however it should be noted that the inter-annual variability was very high at both stations which may indicate an inconsistency in the record, which is not uncommon in cold region precipitation records due to under-catch in winter precipitation (Adam and Lettenmaier 2003).

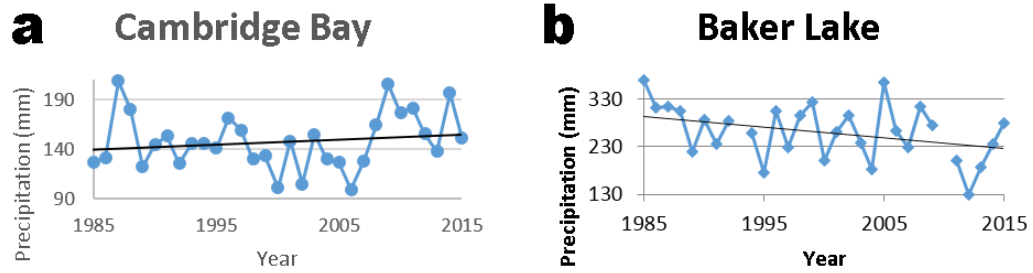


Figure 3.5 Annual precipitation trend for two weather stations within the study region. Cambridge Bay (a) in the north-west shows a slight increasing trend while Baker Lake (b) in the south-east shows a slight decreasing trend.

The result of the multivariate analysis of surface parameters as determinants of surface water change from all random forest runs (table 3.1) shows that the watersheds dataset was the highest ranked predictor of an increasing or decreasing trend in surface water extent followed by elevation, soil types (SLCWG) and arctic vegetation (CAVM) with an Out of Bag (OOB) error (Breiman 2001) of 48.67%. The bare fraction, terrestrial ecoregions (TEOW), connectivity and slope were the least important in all runs.

Table 3.1 Ranking of variables from the random forest analysis after all runs.

Variable	Mean Decrease Accuracy	Mean Decrease Gini
Watersheds	-0.01320	49,525
Elevation	0.06494	27,477
SLC	0.04625	18,873
CAVM	0.01664	10,612
Bare	0.01013	6,830
TEOW	0.01790	4,254
Connectivity	0.00427	2,518
Slope	0.00139	2,112

Further zonal statistics analysis, performed on the top four variables, was aimed at understanding how these variables are related to the observed pattern of surface water change. The relationship between the percent water change and watersheds can be seen in figure 3.3b with primarily increasing water bodies in the northwest and primarily decreasing water bodies in the southeast. The elevation data provides a per pixel height above sea level that the random forest analysis identified as important on a per pixel level. However, elevation is a continuous variable that, by itself, does not provide a convenient way to identify groups or spatial patterns without creating arbitrary thresholds of elevation. The soil types (SLCWG) show only a few broad regions that do not appear to be related to the distinct spatial pattern of surface water change (figure 3.6). The final predictor from the top four of the random forest analysis is the arctic vegetation map (CAVM) where a single vegetation class - “Cryptogram barren complex (bedrock)” – is linked to a decreasing trend in surface water (figure 3.7). The spatial resolution (1 km) of CAVM is coarse compared to the features of the region so finer resolution information is needed to confirm and to fully understand the relationship between direction of change and barren or bedrock surface condition. Small outcroppings of bedrock may not be observable in the 1 km base data used to create CAVM and hence the bedrock class may underestimate the true coverage of bedrock in the region.

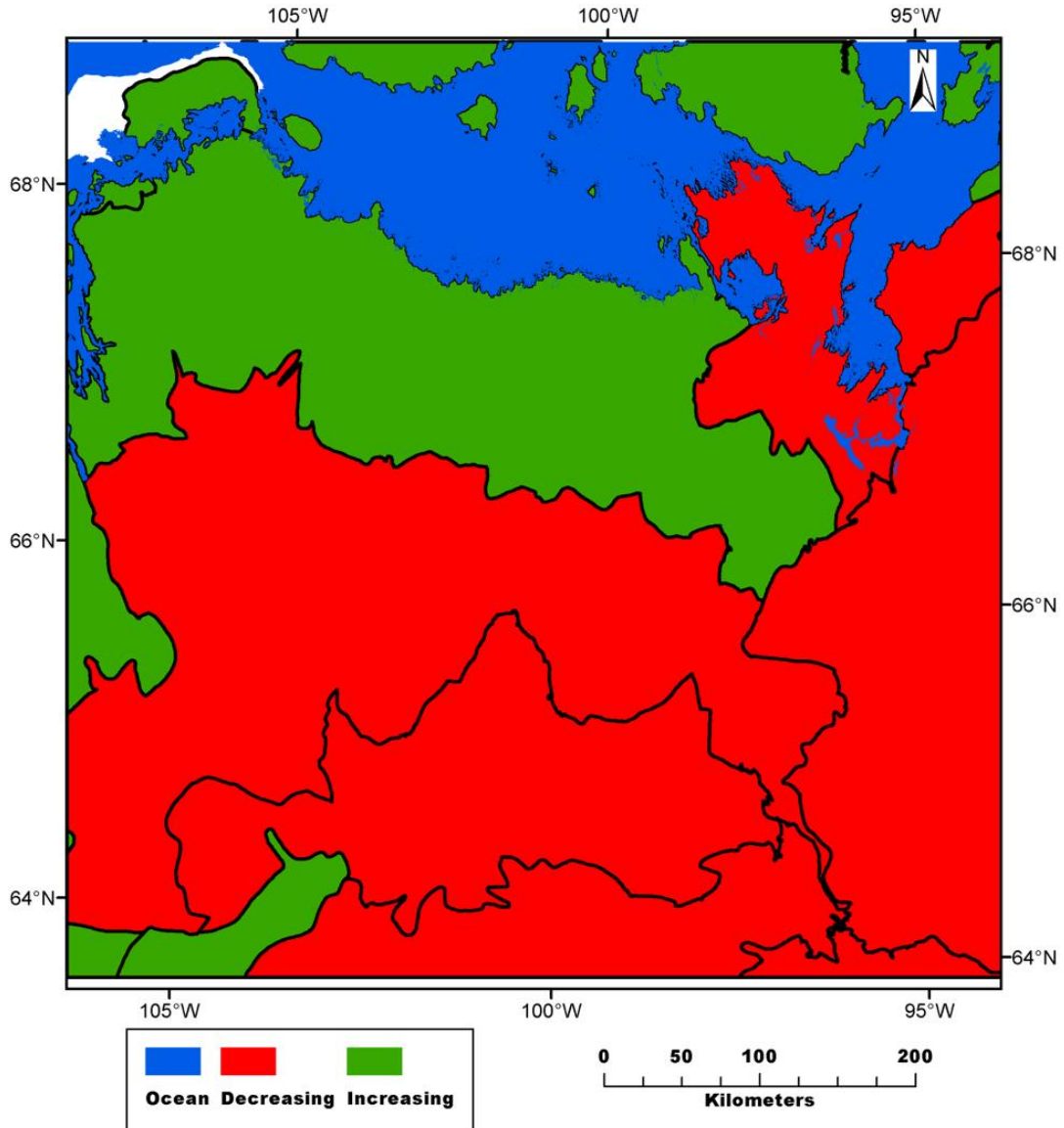


Figure 3.6 Soil Landscapes of Canada (black boundaries) with zonal statistics showing the sign of surface water change by soil regions.

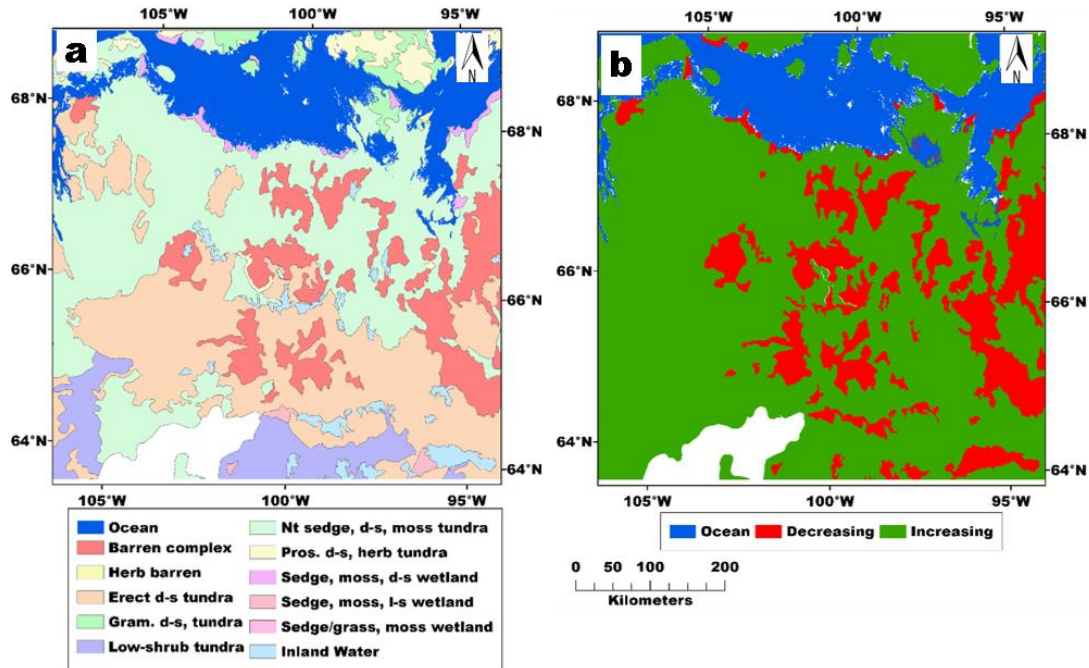


Figure 3.7 In the left panel (a) is an overview of the Circumpolar Arctic Vegetation Map with 11 cover types described, notably the "Cryptogram Barren Complex" which is shown in salmon color. The right panel (b) shows the result of calculating zonal statistics with CAVM as the zone and pixels of change as the variable. The areas shown in red have decreasing surface water extent and generally relate to the Cryptogram Barren complex of CAVM. (Some CAVM long names have been shortened using the following abbreviations "d-s" is dwarf-shrub and "l-s" is low-shrub)

The Canadian National Hydro Network Watersheds data (Natural Resources Canada 2007), which identify 38 distinct hydrologic units within our study area, are the most ecologically meaningful and the most statistically important (as specified by the random forest analysis) spatial zoning parameter for assessing the relationships between trends in surface water and other satellite-derived surface parameters related to vegetative cover and barren surfaces. Table 3.2 shows the results of the linear regression between the median of each satellite-derived descriptor and the percent surface water change per watershed. Median value is used to avoid anomalous values (caused by clouds, shadows, or other poor-quality data) that can contaminate an average or maximum calculation.

Table 3.2 Results from linear regression of surface type per watershed with respect to surface water change using R^2 to measure correlation between the variable. P-values denoted by symbol: no symbol denotes $p > 0.05$; * denotes $p < 0.05$; *** denotes $p < 0.001$.

Dataset	R^2
NCETM-LC:Barren	0.46***
MODIS: Median ET	0.15*
Landsat NDVI: Vegetation Fraction	0.11*
Elevation	0.04
Landsat NDVI: Trend over 30 years	0.03
MODIS: Percent Bare	0.03
Surface Water Area – percent	0.01

The percent surface water area per watershed has a low R^2 and $p > 0.05$ which implies that the amount and direction of change is not correlated with how much surface water is present in a given watershed. Though the relationship is only moderately strong, the best predictor is the percent bare defined by the barren class of the Landsat ETM+ Land Cover of Northern Canada (Olthof et al. 2014) with an R^2 of 0.46, figure 3.8. Though the p value is significant for the analysis of the evapotranspiration (ET) data, the R^2 of 0.15 showing that ET was a weak predictor of surface water change. This is likely due to the coarse spatial resolution of the

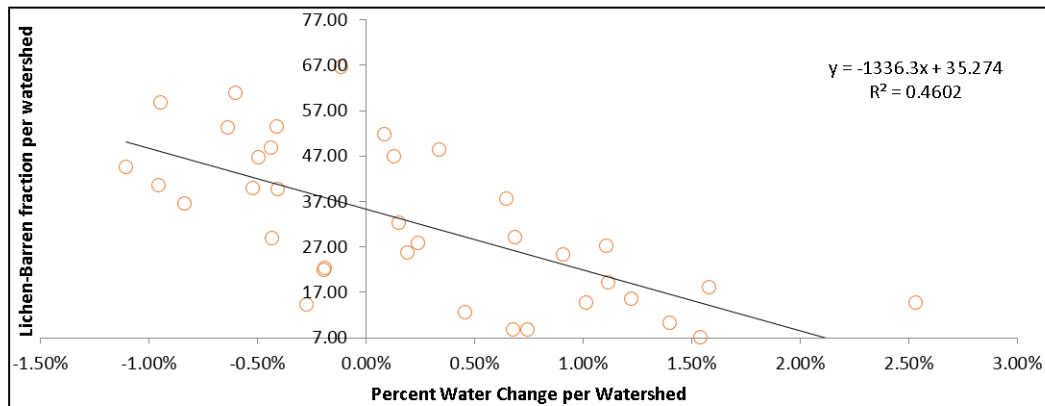


Figure 3.8 Scatterplot showing the relationship between the percent barren class in each watershed when compared to the net percent change in surface water extent per watershed.

product, 500 m, and it was only available for the second half of the study period (2000 – 2015) because the Terra satellite was not launched until December 1999 and data from the MODIS instrument was not available until February 2000. As a consequence of the spatial resolution it is likely that mixed land and water pixels confound the results. Similarly the Landsat NDVI Vegetation Fraction has a significant p value but a poor R^2 of 0.11 showing it is a weak predictor. In this case it is likely that the NDVI is confounded by soil color or wetness leading to poor correlation.

3.5: Discussion

The areal extent of water in a water body at any given time is determined by the hydrological cycle and specifically three primary factors: incoming water (precipitation, inflow from streams/rivers), outgoing water (runoff, evaporation), and topographical depressions that are capable of holding the water. With remotely sensed data we can measure surface features influencing the amount of water stored on that surface. Measurements of subsurface conditions are significantly harder to get because they often require field data collection which can be difficult, expensive, and time consuming. Numerous datasets that describe surface features in various ways are publicly available, however only coarse spatial resolution datasets of subsurface conditions, namely the Circum Arctic Permafrost map (Brown et al. 1998) and the Soil Landscapes of Canada (SLCWG 2010), are available at continental scales.

Building upon our previous work (Carroll and Loboda 2017) we have demonstrated that though there is no distinct pattern in fraction of surface water

“cover” there is a pattern from Northwest to Southeast for surface water “change” in our study region. Direct observations of weather are limited in the study region with only four official weather stations in or near the area that have more than sporadic measurements. Mean annual temperature across the entire region has been increasing at an average rate of $\sim 0.78^{\circ}$ C per decade over the past 32 years with no distinct spatial variability that could explain the observed spatial patterns of surface water change. While it cannot be ignored that the precipitation data from the two weather stations with complete precipitation data show that precipitation is increasing in the north west (Cambridge Bay) and decreasing in the south east (Baker Lake) - the same pattern that is shown in the surface water change - having only two stations does not provide sufficient evidence to draw definitive conclusions because there could be significant variation locally.

Watersheds are defined by topographical features that cause water to flow to a specific outlet. The spatial relationship of change determined by location within a watershed shows that there are localized controls on the flow of water. The analysis of surface water change presented here shows that a larger fraction of barren surface and bedrock are associated with the watersheds that decrease in surface water extent. This is significant because the presence of bedrock at or near the surface suggests that the soil is shallow in these locations. Evaluation of surface water change with NDVI (surrogate for vegetation fraction), NDVI trend (change in NDVI over time), and trend in evapotranspiration all yielded weak correlations with surface water change. One possible explanation for this is that further distinction of vegetation type is needed to provide an adequate relationship between surface water change and

vegetation (i.e. NDVI is a general measure of vegetation vigor without giving an indication of vegetation type). Generation of a new land cover map to investigate this potential is a large undertaking and well beyond the scope of the work described here.

The observed changes in surface water extent are found throughout the study region. Both air and water temperature (Schneider and Hook 2010; Sharma et al. 2015) are increasing in this region. With similar weather and climate conditions for all water bodies (i.e. they do not vary across the study region) the relationship that stands out is the amount of barren or bedrock that is present. In areas where barren or bedrock is less prevalent, hence the vegetation component is higher, the water bodies are expanding. This supports the hypothesis that as the near surface permafrost thaws the water bodies are expanding due to a deepening of the talik beneath the water body. The deeper talik can create a connection to ground water which can have a positive (expansion of water body) or negative (contraction of water body) effect depending on the hydraulic gradient. Talik formation is one possible explanation for the surface water changes that are reported here.

Microtopography is another potential control on the flow of water both on the surface and below through its control on the hydraulic gradient. A DEM finer than the CDED used in this study would be necessary to investigate this and may be possible with future releases of the ArcticDEM (PGC 2017) but at the time of this writing the ArcticDEM is incomplete in the study region which prevents a detailed analysis of the impact of microtopography on surface water. Depth of thaw as well as the composition of the bedrock (fractured or permeable vs solid or impermeable) provide another control on the flow of water which impacts wetting and drying.

The water bodies are drying where barren or bedrock is more prevalent and the vegetation component is low. We hypothesize that the presence of near surface bedrock implies shallow soils in the surrounding area that prevents subsurface lateral water flow and, barring the presence of cracks and fissures in the underlying bedrock, connection to the ground water. There is evidence of similar occurrence in water bodies situated over bedrock on Ellesmere Island, Nunavut (Smol and Douglas 2007a) where changes in evaporation to precipitation ratios have resulted in water deficit and hence reduction (or complete loss) of water body extent. In contrast, in areas underlain by deeper soil layers, permafrost thawing associated with rising temperatures creates new pathways for water flow due to melting of ice wedges (Liljedahl et al. 2016), deepening talik (Arp et al. 2016) or increased surface connectivity between water bodies (Liljedahl et al. 2016; Woo and Guan 2006). Change in surface area of water bodies is not uniform and clearly depends on a suite of environmental conditions including subsurface composition, depth of water, as well as evaporation and precipitation ratios.

3.6: Conclusions

Over 25% of the 675,000 water bodies in the study region experienced a significant multi-decadal trend in change. Though the entire region experienced the same amount of warming the surface water extent responded differently in different parts of the study domain. The results provided here offer evidence of a direct relationship between surface water change and barren or bedrock substrate at the regional scale. This relationship was identified using two independent datasets (CAVM and Landsat ETM+ Land Cover of Northern Canada) and at different spatial

resolutions (1 km and 30 m) respectively. Relating surface water extent change to bedrock detected on the surface provides an extensible method that does not rely on local maps of soils/subsurface or airborne observations with ground penetrating radar. Improved land cover datasets with specific focus on barren (including lichen covered bedrock) land cover types would facilitate expansion of this method to larger region or continental scale. Further improvement would come from a more complete understanding of the subsurface makeup including the amount of fracturing in the bedrock which can also facilitate subsurface flow.

Chapter 4: Identifying potential drivers of change in surface water extent in Arctic tundra using reanalysis data

4.1: Introduction

There is an abundance of evidence that rising air temperatures are having a significant impact on the aquatic and terrestrial environment in the High Northern Latitudes (HNL) (Hinzman et al. 2005; Miller et al. 2010; Serreze and Barry 2011; Serreze et al. 2000; Stroeve et al. 2011). Among the many observed changes, a change in inland surface water extent has been widely reported in both local and regional studies (Bring et al. 2016; Carroll and Loboda 2017; Carroll et al. 2011b; Jepsen et al. 2016; Rover et al. 2012; Smith et al. 2005). In most cases increasing thaw depth of permafrost resulting in the connection of the water body to other surface water bodies or to ground water is identified as the primary mechanism of change (Liljedahl et al. 2016; Smith et al. 2005; Woo and Guan 2006). Rising temperatures in the region are driving permafrost thaw (Grosse et al. 2011; Jorgenson et al. 2010) but other components of the surface energy balance (evaporation, transpiration, latent and sensible heat flux, etc.) are also important factors in changing surface water extent. In a study of water bodies in coastal Alaska Bowling and Lettenmaier (2010) modeled surface energy balance using inputs from weather stations to understand why surface water extent changed. They found that they were able to reproduce temperature and ice cover of the water bodies as well as explain variance in snow water equivalent using the model outputs. A different study compared modeled surface water fluxes based on precipitation records from weather

stations to reanalysis data across the Arctic and found good agreement between the reanalysis data and a modeled result using direct observations (Su et al. 2006). While this study was focused on representing discharge, it demonstrates the ability of reanalysis data to represent surface water fluxes in the Arctic.

The prior studies demonstrate a relationship between surface water change and surface energy balance, however the sparse distribution of weather stations in HNL (Chapin et al. 2000), and the limited diversity of measurements at the existing stations (usually only air temperature and precipitation) mean that researchers need to derive the surface energy components independently (Bowling and Lettenmaier 2010). Calculating surface energy components from weather station data requires implementation of a model and the computational resources to run that model. These requirements can deter researchers from exploring the impact of surface energy balance on their studies. This creates an information gap that has limited, among other things, attempts to relate surface water change directly to potential drivers of change represented by climate variables (Carroll and Loboda 2018). As shown in Su et al. (2006), there is potential for climate reanalysis data to fill this gap by providing data on surface energy balance that can inform analyses of surface water change.

Reanalysis data are generated by assimilation of observational data in a model framework to produce a consistent set of outputs based on the physics that are described in a climate model (Decker et al. 2012). Several different reanalysis datasets exist that are generated from, primarily, the same observations but with different underlying models. For example the Modern Era Retrospective Analysis for Research and Analysis version 2 (MERRA-2) is generated at NASA Goddard Space

Flight Center using the Goddard Earth Observing System (GEOS-5) model (Gelaro et al. 2017; Rienecker et al. 2011) whereas the European Centre for Medium-Range Weather Forecast (ECMWF) model is used to generate the ECMWF Re-Analysis (ERA-Interim) (Dee et al. 2011; Simmons et al. 2006) and the National Oceanographic and Atmospheric Administration (NOAA) National Centers for Environmental Prediction (NCEP) reanalysis (Kalnay et al. 1996). Reanalysis data provide projections that describe hundreds of atmospheric and land surface variables, at sub-daily time steps, including temperature, precipitation and energy fluxes resulting in a comprehensive suite of metrics exceeding 750 Terabytes in size (Gelaro et al. 2017).

There are key differences between reanalysis data, observations and interpolated results. Observations are direct physical measurements of phenomena, such as weather station or radiosonde measurements (e.g. temperature, precipitation, barometric pressure, etc.). These observations are relevant to the physical location where they were collected, but can be representative of a local region and are often taken in concert with similar measurements in other locations. It is possible to estimate the values in between point source observations by performing mathematical interpolation functions (often distance weighted functions) to estimate values at those locations that do not have actual measurements. This data interpolation can work well over short distances but suffers from increased uncertainty as the distance between adjacent observations increases. Reanalyses take observations as input and produce a modeled output of what the value of a variable might be based on the physics constraining the model. Positive aspects of reanalysis data include consistent

global representation, high temporal resolution, and a wide diversity of variables. Challenges with using reanalysis data include limited spatial resolution (0.5° or coarser) and the values that are represented are simulated rather than measured observations. The simulated values can present a challenge when interpreting results because they demonstrate one possible scenario rather than a definitive measured value. The inherent advantage of this approach is that there is a comparable level of confidence between adjacent pixels that is not possible with observations because the measurements are not acquired at equal intervals around the world.

In this study we will use locations of surface water change (Carroll and Loboda 2017; Carroll and Loboda 2018) as a dependent variable in a maximum entropy model with MERRA-2 reanalysis land surface diagnostic variables as input to identify key variables for identifying locations of change. The primary research questions addressed in this study are: 1) *What components of the surface energy balance from the MERRA-2 land diagnostics suite are the best predictors of surface water change for positive change and negative change?* 2) *What do those climate variables tell us about the environmental conditions driving the change in surface water extent?* This analysis is based on a Monte Carlo simulation with MaxEnt software and random combinations of predictors derived from MERRA-2 variables to identify the top ten predictors and the analysis of the resulting map and the set of ranked predictors.

4.2: Data and Methods

The study area is in northern Nunavut territory in Canada, figure 4.1. This region is located within the North American tundra as defined by Terrestrial

Ecoregions of the World (Olson et al. 2001), is underlain by continuous permafrost (Brown et al. 1998), and is characterized by low topographic relief with between

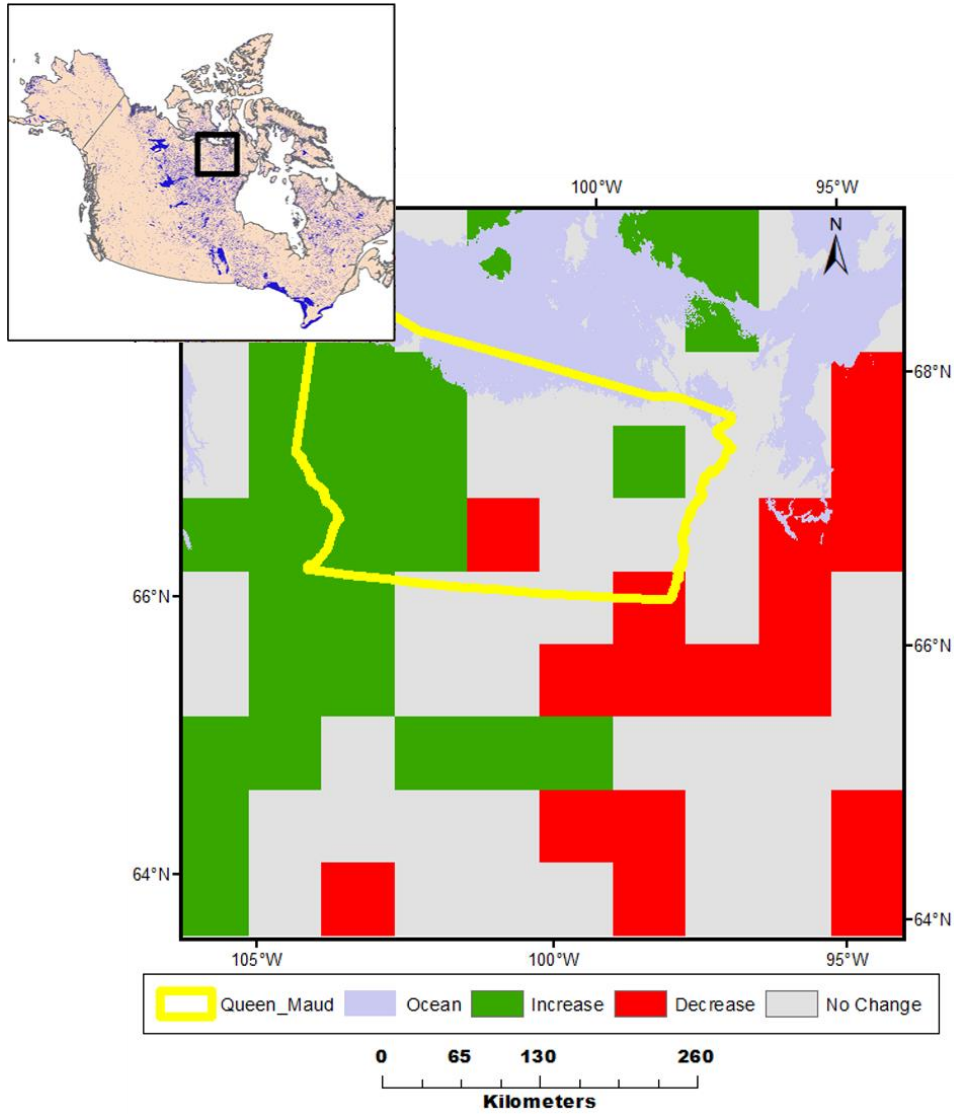


Figure 4.1 Overview of study domain in Nunavut, Canada. The regional map shows net change in surface water extent at 60 km spatial resolution to match MERRA spatial resolution derived from the water bodies with significant change in Carroll and Loboda 2018.

20% and 40% surface water. This area contains the Queen Maud Gulf Bird sanctuary, a global protected area, and has limited anthropogenic activity.

The overall approach combines the ability of machine learning software to rank the predictive power of independent variables with a Monte Carlo convergence

method to identify key determinants of surface water change. The workflow entails generation of a dependent variable from observations of surface water change and independent variables from a suite of climate variables. Surface water change was identified in previously reported work (Carroll and Loboda 2017; Carroll and Loboda 2018) and used here as the dependent variable. The independent variables were 30 year (1985 – 2015) trends derived from 14 MERRA-2 land surface diagnostics variables (Gelaro et al. 2017).

4.2.1: Dependent variable

The surface water change in the study area was quantified using annual maps of surface water created from a time series of Landsat data at 30 m spatial resolution in (Carroll and Loboda 2017; Carroll and Loboda 2018). Over 25% of the 675,000 water bodies in the study region showed a statistically significant amount of change between 1985 and 2015 (Carroll and Loboda 2017). Specific grid cells (locations) of change were identified within water bodies with significant change and dichotomized into a positive (1) and negative (-1) state. The resulting dataset with significant locations of change was aggregated from the native 30 m spatial resolution to 60 km spatial resolution ($\sim 0.5^\circ$) by obtaining the mean value from the dichotomized dataset (figure 4.1) providing net change per 60 km grid cell. The resulting net change in surface water extent was projected from its native Canada Albers Equal Area projection to Geographic (WGS-84) projection to match the MERRA-2 data and converted to three values: 1 for net increase in surface water extent (positive change), -1 for net decrease in surface water extent (negative change), and 0 for no change.

Positive and negative change were addressed in separate models to analyze potential climate drivers independent of the sign of the change.

4.2.2: Independent variables

The inherent complexity and richness of model outputs combined with the sheer volume of the MERRA-2 dataset (>750 TB) necessitates the use of high end computing (HEC) and inferential analytics to extract meaningful information for specific science questions. The MERRA Analytic Service (MAS) provides a straightforward way to access one or multiple variables of MERRA-2 data for user defined temporal and spatial subsets (Schnase et al. 2017). Through MAS, canonical operations (max, min, average, etc.) were calculated for the user defined temporal range (day, week, month, season, etc.), thus reducing the overall size of the dataset, which facilitates incorporation into subsequent analyses. MAS facilitates the use of MERRA-2 data by eliminating the need for the end user to download the large volumes of sub-daily data and reduces the programming/processing burden for the user. The reduced dimensionality of the outputs from MAS still results in hundreds of output files that require inferential statistics to relate specific variables to ecological function.

A total of 50 variables are available in the MERRA-2 land surface diagnostic variables (Bosilovich et al. 2016). From this pool of available variables MAS was used to acquire MERRA-2 data for 14 variables of interest representing temperature, moisture and energy balance, in weekly and monthly averages (table 4.1). Figure 4.2 shows the workflow for the generation of a pool of “predictors” used in the MaxEnt model runs from the MERRA-2 variables. The variables were converted from

representing a fixed point in time (e.g. average temperature for week 7 in 2006) to trend for that time period (e.g. week) over the 30 year record. After completing the trend generation there were 728 (14 variables * 52 weeks/year) weekly “predictors”, and 168 (14 variables * 12 months/year) monthly “predictors”. Filtering was

Table 4.1 List of MERRA-2 variable names and plain English explanations for what the variables represent and associated units that were used in this analysis (Bosilovich et al. 2016).

Variable Name	Variable description	Units
ECHANGE	rate of change of total land energy	W/m ²
EVLAND	evaporation from land	kg/m ² /s ²
EVPSOIL	evaporation from soil	W/m ²
EVPTRNS	evapotranspiration	W/m ²
GHLAND	ground heating of land	W/m ²
GWETTOP	soil moisture surface	1
LHLAND	latent heat	W/m ²
PRECTOTLAND	precipitation	kg/m ² /s ²
SHLAND	sensible heat	W/m ²
SPLAND	spurious land energy source	W/m ²
TELAND	total energy storage land	W/m ²
TSURF	surface temperature	K
TWLAND	available water storage land	kg/m ²
WCHANGE	rate of change of total land water	kg/m ² /s ²

performed to eliminate predictors with only “null data” at locations with training points, which occurred for several of the energy balance variables during the winter. After filtering out the predictors with only null values the final pool contained 490 weekly and 113 monthly predictors. These predictors were used in the model because inputs tied to a fixed point in time can confound machine learning algorithms that are creating a single result over a long time period (Carroll et al. 2011a; Hansen et al.

2002). The final pool of predictors was used to select random subsets of 10 predictors for use in each MaxEnt run.

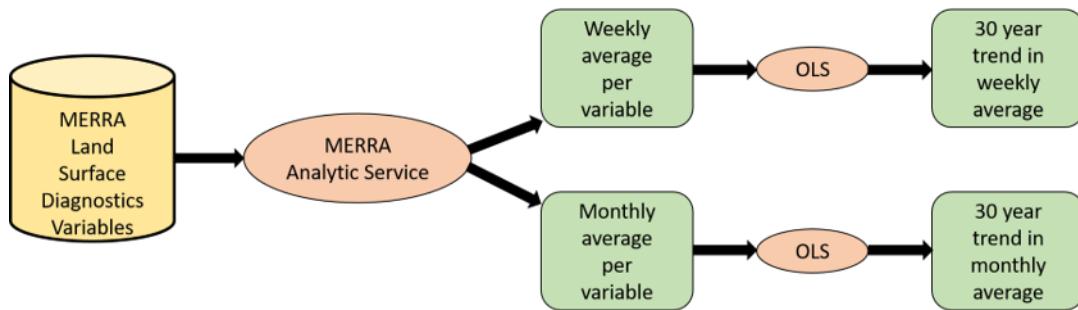


Figure 4.2 Workflow for generating inputs for the MaxEnt software runs from MERRA-2 Land Surface Diagnostic variables. The MERRA Analytic Service (Schnase et al. 2017) is used to derive weekly and monthly averages for each of 14 variables. Ordinary Least Squares (OLS) regression is performed at each time step (week and month) to show trend at those time steps.

4.2.3: Identification of primary predictors

Maximum entropy modeling is a type of statistical inference based on estimated probability distribution that enables the identification of dominant variables without a-priori conditions (Jaynes 1957). This enables us to extract meaningful information based on partial knowledge, when a dataset is too large to extract information directly, and generate descriptive statistics. Two key characteristics of maximum entropy modeling make it an ideal choice in this study: 1) the model is insensitive to correlation between predictor variables, which is necessary because many of the climate variables are auto-correlated; 2) the model functions well even in cases where the training sample size is small which is also true in this case. The classic implementation of maximum entropy modeling is in a machine learning algorithm that takes a dependent variable and models a probability distribution based on a set of independent variables (predictors). The dependent variable is typically a

set of observations of a phenomena (e.g. locations where a species is observed, measurements of concentration of particulates in the atmosphere, etc.) while the independent variables are measurements or other representations of the physical environment that can be used to predict the dependent variable (e.g. landcover, elevation, temperature, precipitation, etc.). MaxEnt software is a popular example of maximum entropy modeling as a machine learning implementation (Merow et al. 2013).

The overall processing scheme (figure 4.3) employs MaxEnt software within a Monte Carlo simulation to identify the best predictors of change using weekly and

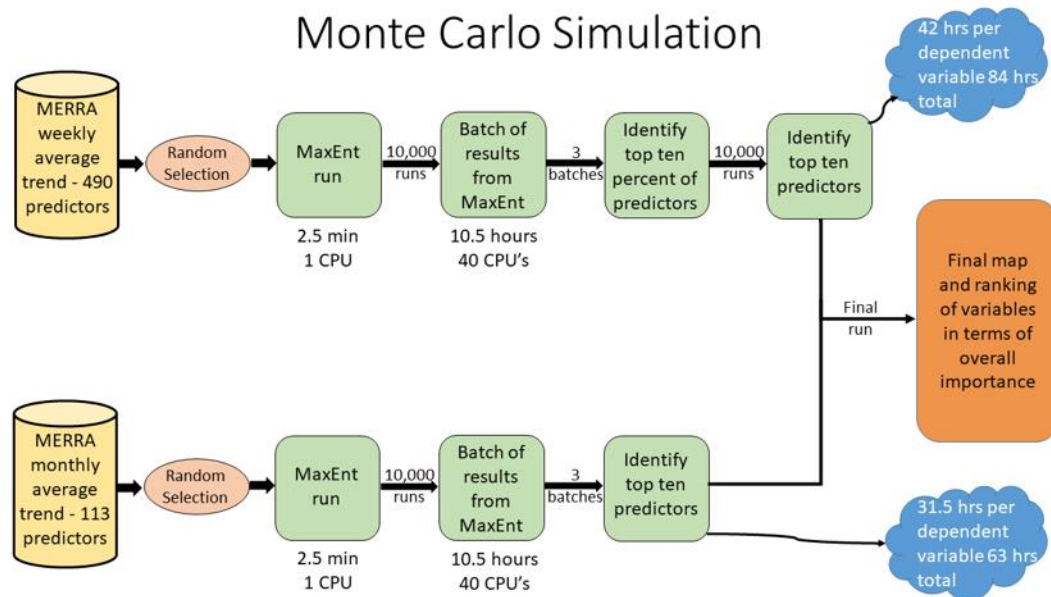


Figure 4.3 Workflow describing the Monte Carlo simulation to determine the top ten predictors for positive change and negative change, respectively at weekly, monthly and combined weekly/monthly time steps. All processing was completed in the ADAPT system NCCS at Goddard Space Flight Center.

monthly predictors. A Monte Carlo simulation is an iterative set of runs of a given model with a random selection of inputs performed sequentially to determine if a model converges on a common result. The large pool of predictors from MERRA-2 makes it difficult to achieve a result from a single run of the maximum entropy model

because the complexity of relationships between variables goes up exponentially with increasing number of predictors. While it is technically possible to perform a single run with all predictors it would require larger computing capacity than was available for this study. Hence, the Monte Carlo simulation is used to facilitate use of the full set of predictors and still enable completion of processing in the available computing environment. In this case we performed the simulation to determine which predictors perform best in the MaxEnt model. The random selection of inputs was performed using the random seed generator in python. Testing was performed to ensure that the random seed was not the same with successive runs.

A “run” was defined as a single execution of MaxEnt software using a random selection of ten predictors. A run produces a ranking of the predictors based on percent contribution to the resulting model and a map based on the model. A batch was defined as a set of 10,000 runs each with a random selection of predictors. After three batches there was a clear convergence in the top ten percent (~50) of weekly predictors. These predictors were selected for a final batch of 10,000 runs to determine the top ten weekly predictors. Monthly predictors were processed in the same way as the weekly with the omission of the intermediate step to identify the top ten percent because the original pool of monthly predictors only had 113 members (less than $\frac{1}{4}$ of the weekly predictors). As with the weekly simulation, there was clear convergence on the top ten monthly predictors after three batches. Finally, the top ten weekly predictors and top ten monthly predictors (20 total predictors) were used as independent variables in a final run of MaxEnt to determine an overall ranking of predictors. In total the simulation entailed 140,000 runs of MaxEnt

software on ten virtual machines with five CPU's each (one CPU on each VM was left idle for system tasks). Overall time to complete the runs was ~150 hours of wall time and a similar amount of actual CPU time because the software maximizes the CPU usage.

4.3: Results

The MaxEnt software produces a map, displaying the result graphically, and two measures of the performance of the model. First, the area under the receiver-operator curve (Miller et al.) provides a measure of the predictive accuracy of the model. As values of AUC approach 1 the predictive power of the model is assumed to be higher. For positive change the AUC of the final MaxEnt model was 0.997, and for negative change the AUC of the final MaxEnt model was 0.996. Both values are approaching "1" which suggests that the model is able to reproduce the training data. Second, separate rankings of variables by percent contribution and percent permutation importance, respectively, gives the user a sense of which variables the model used to make its predictions.

The results (figure 4.4) show the MaxEnt predictions for positive change superimposed on negative change giving a sense of how the results relate to each other. The output shows the probability that a change will occur in a given grid cell on a scale from 0 to 1. This result shows that we can replicate the observed pattern of change only modeled land surface energy predictors derived from a suite of MERRA-2 variables in a maximum entropy modeling system. A quantitative comparison of the training data as it relates to the results from models of increasing surface water

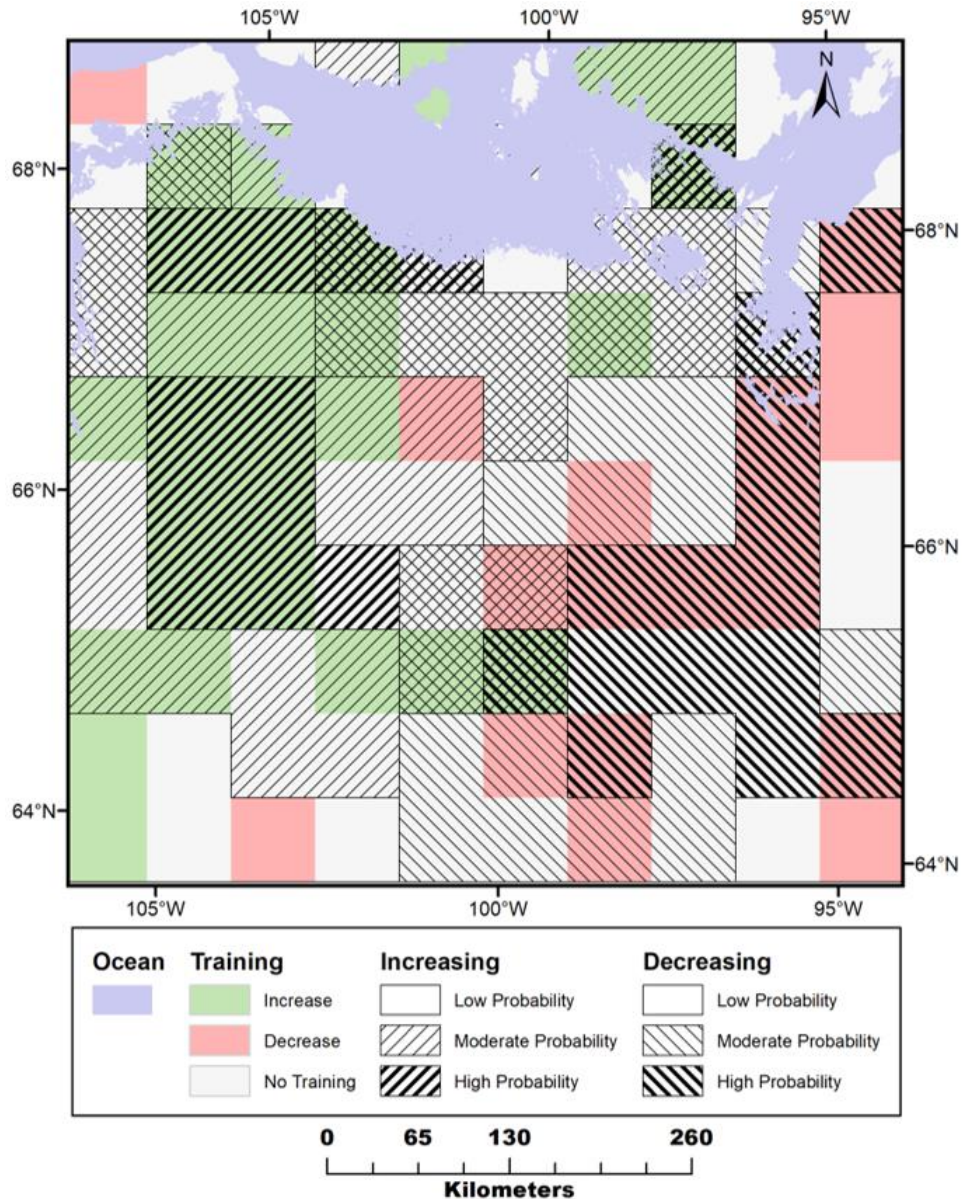


Figure 4.4 Overall results from MaxEnt final runs for decreasing surface water extent and increasing surface water extent with the training data shown at 70% transparent in the background. The separate results have been overlain on top of each other setting probability less than 0.25 to no color. Increasing surface water extent is shown with diagonal lines increasing from left to right while negative change is shown with diagonal lines decreasing from left to right. Heavier line weights indicating a higher probability of occurrence in that grid cell. The areas of overlap can be seen where the crosshatch makes a grid pattern.

extent and decreasing surface water extent was performed and the results are shown in table 4.2. This evaluation shows that MaxEnt labeled 11 out of 18 (61%) training pixels for decreasing, and 18 out of 28 (64%) of training pixels for increasing as

either moderate or high probability of change. In addition 24 out of 36 (67%) "no net change" pixels are labeled as equal probability of increasing or decreasing.

Table 4.2 The relationship between the results of MaxEnt runs for increasing and decreasing surface water extent relative to the training data.

Class	Grid cells	Training - Increasing	Training - Decreasing
High probability increasing	9	8	0
Moderate probability increasing	19	10	1
Low probability of net change	22	2	5
Moderate probability decreasing	13	0	3
High probability decreasing	12	0	8
High probability decreasing & Moderate probability increasing	2	2	0
High probability increasing & Moderate probability decreasing	3	2	0
Moderate probability decreasing & Moderate probability increasing	14	4	1
Majority Ocean	6		
Total	100	28	18

The top five predictors for increasing surface water extent and decreasing surface water extent based on permutation importance are shown in tables 4.3 and 4.4 respectively. The permutation importance measures the effect of removing a particular variable from the training on the value of Area Under Curve (Miller et al.). In cases where individual predictors may be correlated with other predictors it is more instructive to use the permutation importance than the percent contribution because it

does not depend on the structure or paths within the model (Phillips 2005) hence we use permutation importance to rank the predictors for this analysis. The top five predictors explain 84% and 96% contribution to the final model for increasing surface water extent and decreasing surface water extent respectively.

Table 4.3 Top five predictors for increasing surface water extent ranked by the permutation importance. Numbers are converted to percent contribution to the model.

Predictor	Time Frame	Permutation importance
Downward ground heating	Week 36 (mid-September)	32%
Surface temperature	Month 5 (May)	29%
Surface temperature	Week 17 (late April)	9%
Rate of change of total land energy	Week 21 (late May)	7%
Surface temperature	Week 18 (early May)	7%

Table 4.4 Top ten predictors for decreasing surface water extent ranked by the permutation importance. Numbers are converted to percent contribution to the model.

Predictor	Time Frame	Permutation importance
Surface temperature	Week 17 (late April)	71%
Surface temperature	Week 26 (early July)	11%
Downward ground heating	Week 36 (mid-September)	7%
Latent heat	Week 38 (late September)	5%
Surface temperature	Month 7 (July)	2%

For increasing surface water extent, table 4.3 shows the final ranking of the top five predictors from the combined run using top ten weekly plus top ten monthly predictors, 20 total predictors in the final run. The results show that based on the

permutation importance weekly predictors make up all but one of the top five predictors and the top two predictors explain 61% of the contribution to the final model. The trend in downward ground heating flux in week 36 is the top predictor with 32% contribution and the trend in surface temperature for month 5, May, is second with 29% contribution.

The top predictor for decreasing surface water extent (table 4.4) is trend in surface temperature from week 17 with 72% contribution to the final model. Two out of the top five predictors were the same between the increasing compared to decreasing surface water extent. Four of the top five predictors are weekly and three of the top five predictors are based on surface temperature.

4.4: Discussion

The successful replication of observed patterns in surface water change using MERRA-2 surface energy balance variables in the MaxEnt software demonstrates the methodological viability of using MERRA-2 climate variables as a surrogate for weather observations. The small sample size (100 MERRA pixels) precludes the use of many other machine learning algorithms (e.g. random forest) which require a much larger number of samples to deliver robust results. MaxEnt has previously been shown to be successful in generating robust models even in cases where sample size is small (Hernandez et al. 2006; Pearson et al. 2006). The challenge of linking observed phenomena to modeled data (such as MERRA-2 in this study) is considerable in the HNL regions where few observational records are available to fit the physical models. It is, therefore, important to match the strong longer-term predictive capabilities of the reanalysis data with a proper analytical framework to

avoid misattribution of drivers of change to potentially unrelated spurious correlations between observed and modeled parameters within the data record. Some examples of such analytical framework include assessing multi-year trends of a daily parameter (e.g. fire weather) to compare trends in fire occurrence at a decadal scale (French et al 2015). In this work, however, the long-term predictive capacity of the reanalysis data was extracted through assessing 30-year trends in weekly and monthly parameters rather than including the specific values for individual time periods (e.g. week 17 in year 2006). The results show that there are specific variables within the MERRA Land Surface Diagnostics product suite that can reproduce greater than 60% of the training for change in surface water extent in our study region. The top predictors, surface temperature and ground heating flux, are both surface energy balance components that are well characterized in MERRA-2. Incorporation of these variables enables the analysis of atmospheric drivers of change that was not previously possible due to the lack of observations. Future development of continental and circumpolar datasets of long-term trends in surface water extent will facilitate the expansion of this type of investigation to a larger geographic domain.

In addressing the primary science question of this study, the results indicate that the long-term changes in surface energy within shoulder seasons drive the change in surface water extent regardless of the sign of the change (increasing or decreasing). The shoulder seasons are known to be important because they represent the transition of frozen to thawed surface in the spring and vice-versa in the fall (ACIA 2005). This study's findings are unique because they show that weekly trends serve as stronger predictors of long-term trends in surface water extent compared to monthly trends.

Given the brevity of the actual transition during the shoulder season it is clear that working at the weekly time step can provide a clearer signal for the magnitude of the transition in the shoulder season lost in the monthly data assimilation. However, considering that weekly predictors are subject to greater inter-annual variability, care must be taken to ensure that extreme conditions are not driving the statistical inference. In this study we mitigated that concern by generating trends in weekly averages of climate variables as opposed to using inputs from individual weeks or months.

According to our results, increase in surface water extent is driven almost equally by ground heating flux in the fall (e.g. week 36 – mid-September) and the surface temperature in the spring (e.g. month 5 – May). This suggests that the driver of increasing surface water extent is the overall lengthening of the period where temperatures are above freezing which allows for greater depth of thaw in the permafrost. With deeper thaw in the permafrost there is greater opportunity to connect surface water with ground water and also to facilitate slumping around the edges of the water bodies which will expand the surface area.

Decrease in surface water extent is primarily driven by early season temperature (e.g. week 17 – late April). The trend in temperature for this period is declining which is counter-intuitive but may suggest that there are changes in snow cover which allow for changes in surface temperature. A decrease in temperature may indicate a reduction in overall snow cover allowing cold air temperatures to penetrate into the land surface. Snow depth observations, as with most observations in the region, are sparse due to limited number of settlements in the region. A

remotely sensed dataset derived from active microwave sensors is available but only covers the period from 1998 to present (Brown and Brasnett 2010) which is insufficient to test this idea.

Overall, precipitation is a key control for hydrologic processes but the dynamics are poorly understood. MERRA-2 provides information that can be used to understand, at a process level, what may be happening with precipitation. Trends in precipitation were included as predictors in the Monte Carlo simulation but only one appeared in the top ten selected predictors in the final run for increasing surface water extent. Even the one predictor (Precipitation from month 3, March) was ranked low in permutation importance (6.5%). Intuitively, we expected the precipitation variables to have greater importance in the final model. Reanalysis data have had difficulty reconstructing precipitation (Reichle et al. 2017) and while it has been improved in MERRA-2 there are still known issues with the hydrology variables especially at high latitudes where observations are sparse (Gelaro et al. 2017).

The long term trend in temperature in the study region shows a 2° C increase in average annual temperature over the past 30 years (Carroll & Loboda, 2018). This warming results in a small amount of difference in physical comfort difference for humans but can ultimately manifest as a longer frost free season, as shown in our results for increasing surface water extent. Climate projections show further warming in the coming years with as much as 4° C increase in temperature by the year 2100 (IPCC, 2014). Rising air temperatures result in greater annual thaw depth in permafrost and greater evaporation potential from open water. Tipping points exist in any system (climate, ecosystem, etc.) and we often cannot predict them until they

actually occur. The shallow water bodies in the Arctic tundra are highly responsive to changes in climate and may provide an early warning that tipping points are approaching in the way that they are responding to the rising temperatures. At the minimum, a positive outcome of this work would be realized if the density of observations of weather variables could be expanded so that we can understand the variability of temperature at local and small regional scales.

4.5: Conclusions

The accelerated rates of warming in the Arctic are impacting ecosystems' state and functioning in various ways. The satellite-observed change in surface water extent over the past 30 years reveals persistent multi-year trends that cluster regionally into zones of growing and shrinking water bodies. While it is anticipated that climatic changes are driving these trends, the scarcity of in situ meteorological observations in this region substantially limits the ability to explore climatic drivers of changes in surface water extent. Although reanalysis data present a suite of modeled rather than observed values describing the climatic state of the system, they are the only viable source of information to support such studies. The results of this work, focused on identifying MERRA-2 surface energy balance metrics that reflect the observed changes in surface water extent, show that surface temperature and downward ground heat flux in the shoulder seasons can reproduce the pattern of independently observed zones of growing and shrinking water bodies in Canadian tundra.

The short snow and ice free season in the Arctic defines the window when surface energy balance drives the processes within the hydrological cycle that govern

changes in surface water extent. Thus even comparatively small lengthening or shortening of this window can have a substantial impact on water bodies. The results show that temporally narrow (e.g. weekly) changes in shoulder seasons hold the key to understanding the patterns of surface water change. In contrast monthly variables, particularly those at the peak of the growing season, hold little explanatory power for the observed growing and shrinking water bodies. Decrease in surface water extent is linked to decreasing trend in surface temperature in late April. Increase in surface water extent is linked to an overall lengthening of the non-frozen season.

The relatively coarse modeling grid of available climate reanalysis datasets limits the studies aimed at identifying climatic drivers of environmental change to those that can be observed at coarse regional scales. One potential avenue for future work should focus on downscaling the climate variables to finer resolution to investigate localized conditions. At the same time, another avenue for future research using the native resolution of the reanalysis datasets should pursue studies at continental and circumpolar scales to test and verify the linkages between surface energy balance and surface water change across the entire Arctic domain.

Chapter 5: Conclusion

5.1: Major research findings

The goal of this dissertation is to answer the overarching research question: “What is the spatial pattern of variability in the extent of tundra water bodies in High Northern Latitudes (HNL) of North America and what are its environmental determinants?” To answer this question I conducted three integrated studies that took advantage of high end computing (HEC) and long time series (30 plus years) of satellite observations of land surface and reanalysis data.

In question 1 I explored the type of maps required to understand the dynamics of surface water extent in the Arctic tundra so that inter-annual variability could be distinguished from actual long term change. When I began my research there were several continental to global maps that showed surface water extent at 30 m spatial resolution (Carroll et al. 2016a; Feng et al. 2015; Zheng 2014). Of these only the one that I generated in my previous work (Carroll et al. 2016a) had used multiple inputs to create three decadal surveys, i.e. three maps 10 years apart. An important methodological sub-question within the scope of research question 1 was to assess whether the decadal surveys are representative of a long-term trend in surface water extent in the tundra. Initial evaluation quickly revealed that having only three time periods to represent the extent of water was insufficient to identify the actual change. In most cases individual water bodies had different surface area at each of the three points in time with no clear trend. To overcome this limitation I generated annual maps of surface water extent using the full time series of Landsat data to quantify the

location and extent of surface water change. Analysis of the annual maps confirmed that inter-annual variability is extensive and that decadal scale maps are inadequate for analysis of change in surface water extent.

Previous studies considered overall change in the region rather than assigning change to specific water bodies (Roach et al. 2013; Smith et al. 2005), used a sporadic time series of inputs (Jepsen et al. 2013; Rover et al. 2012), or conducted time series analysis on individual pixels (Nitze et al. 2017). In my doctoral research I adopted an object based approach where each water body was identified as a unique entity and change was determined for the object and test for significance provided an additional attribute. By doing this I was able to show not only the difference between inter-annual variability and long term change of the total combined surface water extent but also the spatial relationship between water bodies that have a significant amount of change. Change in surface water extent of individual water bodies is of considerable ecological importance, reflects ecosystem state, and in part determines ecosystem functioning (Adrian et al. 2009). Thus my methodological approach to a regional assessment of change of individual water bodies over a long time frame is a highly novel and significant contribution to the methodological toolbox for Earth System research that can be scaled from local to global studies. The methodology and scientific findings were published and can be followed by other researchers in the field (Carroll and Loboda 2017).

Through question 2 I explored the distribution of changes within my study region to determine if there is a discernible pattern to the location and sign of the change and to expose any potential environmental drivers of that pattern. A clear

transition of wetting to drying from Northwest to Southeast was evident when the surface water change extent was related to watersheds within the study region. Watersheds are defined by the topography of the area such that water generally flows through that area to a common outlet. Relating the surface water change results to the watersheds provided an ecologically meaningful way to group the water bodies to look for patterns. Despite a dearth of information regarding environmental characteristics and specifically subsurface composition of the region there is evidence that exposed bedrock is associated with a net decrease in surface water extent. Given the overall lack of information available in the region I focused on interrogating all available satellite data for explanatory variables. A detailed review of available datasets that were evaluated is included in the appendix.

The entire study region is characterized by low topographic relief which means that the hydraulic gradient between water bodies is small. Analysis of the Canadian Digital Elevation Data (CDED), at 30 m spatial resolution, was unable to reveal meaningful connectivity characteristics between the water bodies. The new ArcticDEM version 2, at 5 m spatial resolution, will likely provide new insights into the microtopography of the region that affects the connectivity between water bodies. The relationship between a net decrease in surface water extent and exposed bedrock makes a compelling case for acquisition of data from ground penetrating radar as has been done in the Yukon Flats in Alaska (Minsley et al. 2012). Finally, the lack of weather observations, particularly reliable precipitation measurements, limits the ability to relate surface water change to obvious potential drivers. These results have been published through peer review (Carroll and Loboda 2018).

With question 3 I expanded that search for environmental drivers of change in surface water extent by exploring the use of climate variables. The lack of consistent and well distributed meteorological stations in the Arctic make it difficult to generate a reliable gridded data product for temperature and precipitation. The Daymet dataset - a gridded interpolation of weather station data that is available for this region (Thornton et al. 2017) - exhibits extensive artifacts in my study region which reflect the extremely sparse network of measurements and introduce a very high degree of uncertainty in the analysis. Reanalysis data are generated by assimilation of available information into climate models to produce projections for a suite of climate variables based on the physics of the model.

In this work I used a subset of Modern Era Retrospective analysis for Research and Applications version 2 (MERRA-2) land surface diagnostic variables describing temperature and moisture conditions for the study region in a machine learning algorithm (MaxEnt) to identify any variables that are good predictors of surface water change. This part of the research used the HEC environment and analysis tools to distill large volumes of MERRA-2 data into meaningful predictors at weekly and monthly time steps. These 490 weekly and 113 monthly predictors were the pool of environmental variables used in a Monte Carlo simulation with MaxEnt software. The resulting map clearly reproduced the pattern of change in the training data (i.e. the surface water change) and the top predictors came from the shoulder (early Spring and late Fall) seasons. This research used maximum entropy modeling to relate surface water change to change in climate variables provided by reanalysis data. To do this I used high end computing to reduce dimensionality and volume of

the climate variables and to convert from fixed point in time to trend over time. The conversion from fixed point to trend helps the machine learning software from getting confused by extreme values associated with a single point in time. These results have been compiled in chapter 4 and will be submitted for peer review.

5.2: Study limitations

The methods adopted within the scope of this dissertation demonstrate research capabilities within the constraints of existing software and data availability. However, the study domain of this research project represents only a comparatively small region within the Arctic tundra. While the majority of methods used within this work are directly applicable to other regions, there are a number of limitations to their deployment in other regions. First, the Canadian Arctic, including the study area of this doctoral research, represents the HNL region with the densest time series of Landsat data. Historically, all Landsat 4 - 7 observations were acquired and archived through a Canadian satellite receiving station (Goward et al. 2006). However, even within this study the overall availability of clear surface views was limited by the total available observations from the snow and ice free season of ~3 months duration. Further constraints are caused by persistent cloud cover and the infrequent (every 16 days) repeat coverage of Landsat observations. Expanding this methodology to other non-Canadian HNL regions, and specifically Alaska, would be challenging for the period prior to 1996 because Landsat data record is not sufficiently dense to generate high quality annual maps. This problem becomes more acute when looking at the Arctic in Europe and Asia where the acquisition strategy for Landsat was to save one or two cloud free observations per year. This policy was in place into the 2000's

making it nearly impossible to build high quality annual maps with multiple observations to ensure that the nominal water extent is captured. Many of the limitations of Landsat data availability are problematic only for the historical analysis. Current satellite coverage with Landsat 8, Sentinel 2 A/B and radar data from Sentinel 1 A/B, not to mention the commercial high resolution instruments, provides enough data to make generation of current maps of surface water extent a more tractable problem. However, the historical record is necessary to establish baseline conditions and, therefore, comprehensive use of the available data archive is key for examining environmental change at climate-relevant temporal scales.

The second major limitation to the expansion of the object-based water bodies analysis stems from the existing software constraints. Identification of individual water bodies is computationally intensive and is difficult to do if the data are in multiple images rather than one large mosaic because the seams where tiles are split will inevitably divide water bodies artificially. The GIS software that is typically used for this task is often inefficient when working with image files that are very large. A solution for this limitation is possible and will involve a substantial customization of GIS routines to enable the analysis over a larger spatial domain.

The third factor is the lack of observational data for weather parameters, specifically temperature and precipitation. As previously mentioned in sections 3.5 and 4.4 there are only 4 weather stations in or near my study region and only two of these have consistent precipitation records. Observational data is needed to help improve the reanalysis data and/or to validate the results of the modeled data. A long

term record would be ideal but even a short term spatially dense observation network could improve the characterization of the modeled reanalysis data.

The fourth limitation is the small number of input observations available at the coarse spatial resolution of MERRA. The study region has only 100 pixels (grid cells) at MERRA resolution which precludes the use of many analysis tools including many common machine learning tools including random forest analysis. With the current constraints it was necessary to use the MaxEnt software, which handles small numbers of input pixels well, and precluding saving out a sample of training pixels for testing and cross validation. A larger spatial domain for the change in surface water extent or a finer resolution version of MERRA would facilitate a much broader analysis. Doubling MERRA spatial resolution to 0.25 from 0.5 would yield 4 times as many pixels.

The final and most constraining factor is the lack of spatially explicit data on subsurface composition at the regional scale. The lack of datasets that characterize subsurface properties in tundra ecosystems prevents definitive attribution of change to potential causes. Specifically measurements of thaw depth, depth of soil (prior to bedrock), soil texture (fine soil as compared to gravel), permafrost ice content, hydraulic gradient, and bedrock type and state (fractured or solid) all have impacts on subsurface water flow.

5.3: Big data analytics

Environmental science in HNL is simultaneously very data poor (particularly for in situ observations) and extremely data rich (for satellite and modeled data

suites). To take advantage of the rich satellite and modeling data archive, big data processing and analytics as well as machine learning algorithms are a must for environmental analysis at regional to continental scales here. My dissertation research specifically delved into 2 discrete activities – 1) regional scale long-term mapping and analysis of trajectories of change for discrete water bodies, and 2) analysis of MERRA-2 as a surrogate for observations of climate variables.

The process to generate annual maps entailed image classification of ~30,000 dates of Landsat data from 55 path/rows. These data represent all available data for the ice-free season (nominally May – September) regardless of cloud cover. In place of the custom classification algorithm used in Carroll et al. 2016 I used a product generated by the United States Geological Survey called Dynamic Surface Water Extent or DSWE (Jones 2015) for the image classification of individual dates. Analysis of the DSWE product demonstrated a comparable result to the decadal water maps (Carroll et al. 2016a) and use of this product over a custom algorithm enables future researchers to test and reproduce this result. The data volume of raw DSWE inputs is over 1 TB including data from Landsat 5 and 7. Using 20 virtual machines (VMs) in the Advanced Data Analytics Platform (ADAPT) at NASA Center for Climate Studies (NCCS) I was able to reduce ~30,000 inputs totaling 1 TB into 30 maps totaling ~7 GB in approximately two weeks of compute time. By way of comparison, a single server processing stream would have required ~ 9 months of run time.

The generation of MERRA-2 data is only possible due to the availability of high end computing because it is generated using the Goddard Earth Observing

System version 5 (GEOS-5) climate model which is computationally intensive. MERRA-2 land surface diagnostic variables are produced at hourly time step which is not ideal as input to the MaxEnt software. The MERRA Analytic Service was used to distill several terabytes of global MERRA-2 data into regional subsets at weekly and monthly time steps. These weekly and monthly data were then used to generate trend per variable for 52 weeks and 12 months. To perform these calculations outside of the HEC environment would have taken several months. Finally, the Monte Carlo simulation which comprised 140,000 runs of MaxEnt to achieve the final map and ranking of predictors was completed on 10 VMs using 4 CPUs on each VM and finished in 147 hours. For comparison this would have taken 40 times as long (5,880 hours or 245 days) on a single server with one CPU.

A significant outcome of my research is the combination of long time series of environmental data and high end computing to produce new and innovative results. This is the definition of “Big Data” analytics and has emerged as an essential component to current environmental research. The proliferation of HEC at NASA institutions as well as universities suggests that this wave will continue. In the private sector there are several cloud computing environments including Amazon Web Services and Google Earth Engine. The latter of these has been used by researchers in the generation of map products from Landsat and other satellite data (Hansen et al. 2010; Pekel et al. 2016). The ease of access to these resources will encourage the next generation of product creation by reducing the constraints of input data volumes and long processing times that make it difficult for researchers to manage big data.

The time recovered from the processing end can be reallocated to the validation and analysis parts of the project as I have done in my research.

5.4: Contribution of this research to the broader North

American Arctic research

5.4.1: Broader Arctic agenda

The Arctic and Boreal regions are the subject of intensive research efforts with initiatives and assessments provided by organizations including the Arctic Council, Interagency Arctic Research Policy Committee (IARPC) as well as the governments of the countries with significant land area in the Arctic. The Arctic Council is an international organization comprised of six working groups that provides a forum for promoting cooperation and coordination between council members (Council 2018). The Arctic Monitoring and Assessment Programme (AMAP) is one of six working groups and is tasked with monitoring and assessing the status of the pollutants and climate change in the Arctic. The most recent report from AMAP concludes that no unified trends have been discovered regarding thermokarst lake development and the further research is needed to understand if this means that there is no trend or that the current body of work has been insufficient to discover the trends (AMAP 2017). The methodological advances in chapters 2 and 3 have direct applicability to solve the need to understand the dynamics of surface water in HNL. Furthermore, the report encourages future analysis to be performed at the catchment or watershed scale (AMAP 2017). IARPC was formed in 2010 and consists of principals from 16 US federal agencies to enhance scientific monitoring and research

in the Arctic (IARPC 2018). One goal of IARPC is to develop a more complete understanding of the effects of thawing permafrost on infrastructure and climate (IARPC 2018). The research presented in this dissertation supports the goals of both the Arctic Council and IARPC by providing a specific methodology that not only quantifies the magnitude and direction of change in surface water extent but also a way to attribute the change to both physical drivers and climate drivers related to surface energy balance.

The North American Arctic is completely within two countries: the United States and Canada. In the US, NASA is funding an effort through the terrestrial ecology program called the Arctic and Boreal Vulnerability Experiment (ABOVE) with the multi-year goal of synthesizing observations and models to generate a fuller understanding of the processes at work in the region. The federal government of Canada is funding several programs including the Global Water Futures (GWF) with the goal of improved disaster warning, predicting water futures, and adapting and managing risk. The major findings in this dissertation provide the groundwork for a continental application of methods that will inform the hydrology related aspects of these programs.

In addition to the direct effects of changing hydrology on infrastructure and resources, the change in surface water extent has an impact on carbon stocks through exposure of carbon rich soils. Quantification of carbon stocks as well as identifying sinks and sources is a global concern. It is well established that the sediments of many Arctic water bodies are laden with carbon and that this carbon can be activated by changes in the surface water extent. Annual maps of surface water extent at the

continental scale could be used to supply carbon models with information about short term and long term exposure of sediments. The models can incorporate surface water extent with other information to refine estimates of emissions and reductions in carbon within surface water in the Arctic.

Changes in surface water extent can have major impacts on migratory waterfowl through the disruption of access to breeding grounds and food sources (Hinzman et al. 2005; Hinzman et al. 1991). Waterfowl can have ecological and economic impacts on the region from subsistence living to tourism (Bromley 1996) with far reaching effects well south of the Arctic as the birds migrate to the south in the winter.

5.4.2: Future directions

I am currently leading a synthesis paper for the hydrology and permafrost working group for ABoVE. The results from my study region will be used as one of four case studies in this synthesis paper. One goal of this synthesis paper is to determine the best available method for quantifying historical surface water dynamics for the entire ABoVE study domain. A second goal is to identify information gaps that can be filled by ABoVE scientists in the next round of projects. One component of this will certainly be the collection of additional information about subsurface composition at broad scales. The lack of suitable subsurface information on soil composition, soil depth, permafrost type and bedrock type are significant information gaps that need to be filled to truly understand the hydrology of the region.

Historical data from Landsat 5 and 7 have been used to generate maps of land cover, including surface water, because of their capabilities in terms of spatial and

spectral resolution. New satellites have come online over the past few years including more optical instruments such as Landsat 8 and Sentinel 2 A/B. Optical instruments such as Landsat and Sentinel 2 provide moderate resolution maps of land cover features. Using multiple instruments, a continuous record of observations exists that spans 1984 to present. Taking full advantage of this record requires accurate characterization and calibration of each instrument and the differences between the instruments. Work has begun to produce a “harmonized Landsat – Sentinel” data product that accurately calibrates the measurements from the disparate sensors (<https://hls.gsfc.nasa.gov/>). The optical data provided by Landsat and Sentinel 2 is limited to daytime cloud free conditions to acquire data. In regions with pervasive clouds, such as the Arctic, optical instruments with infrequent repeat coverage can struggle to get repeat measurements over short time steps. Radar instruments have the ability to record measurements of the ground even when clouds are present, however the historical record of radar data is sparse. Since their launch in 2014 and 2016, Sentinel 1 A and B have been collecting large area coverage of the Earth using C-band radar which is sensitive to surface water including water with emergent vegetation. Hence, supplementing the optical data with information from radar instruments such as Sentinel 1 A/B can provide not only additional observations but also previously unavailable information about inundated vegetation (Huang et al. 2018). These instruments bring a depth and breadth of information that will enhance our ability to discriminate between wetlands, open water, and just vegetation. As our understanding of the surface improves we will be able to develop a better understanding of the processes that are at work.

Finally, all of the data that is collected must be processed into actionable information which makes research in the Arctic a big data problem that can only be solved through integrated research with high end computing. The use of HEC is necessary to support the creation of novel data products as well as the creation and analysis of modeled synthesis outputs. Simply collocating the large volumes of input data with a compute environment is a significant advance that reduces the burden on the researcher in terms of time needed to download and maintain those data. HEC resources will become even less expensive and easier to access in the coming years. Big data analytics using HEC, long time series of satellite observations, and advanced computing techniques such as machine learning will drive future scientific advances and enable researchers to complete our understanding of the Earth system. In northern Alaska, alone, there is a current need to build on existing maps of wetlands to support the National Wetland Inventory and to understand how exploration for resources, such as oil and gas, is impacting these wetlands. The methods described in this dissertation can be applied more broadly to build the inventory of water bodies, including wetlands, and to facilitate long term monitoring of these sites for changes, both natural and anthropogenic.

Appendix

Chapter 3: Supplemental Material

Remotely sensed data are often used to quantify surface features such as cover type (water, forest, crop, grassland, etc.). The advantage of remotely sensed data lies in the routine coverage at various spatial resolutions even in vast remote areas that are difficult to reach and/or expensive to collect field measurements. Surface features that are detectable with remotely sensed data can be used to infer information about ecological function (Cohen and Goward 2004; King et al. 2005). For example, land cover classes such as forest, cropland, and urban/impervious identify specific parameters and thresholds governing the hydrological cycle. Although Arctic tundra is usually referred to as a single large ecosystem, the landscape is particularly complex with many small water bodies as well as different vegetation types which make it a challenging ecosystem to map. This problem becomes more challenging at coarse spatial resolution (> 250 m spatial resolution) due to the heterogeneity of cover type in each grid cell at coarse resolution. Below we provide a full account of surface parameters data suite assembled in this project.

A.1: Land cover data

Terrestrial Ecoregions of the World (TEOW) is a static single-time map at coarse spatial resolution that depicts land cover qualified with climate data to explain certain niches especially in the tropics (Olson et al. 2001). Three TEOW ecoregions are

represented in the study region, including Northern Arctic, Southern Arctic and Western Taiga Shield.

MODIS Vegetation Continuous Fields – Percent Bare (VCF): produced annually (2000 – 2015) at 250 m spatial resolution to depict the amount of non-vegetated surface (includes bare soil, rock, and permanent ice/snow) (DiMiceli et al. 2011). Data from 2011 – 2015 were used in this analysis.

Landsat ETM+ mosaic of northern Canada - Land Cover: a static map of land cover in Canadian tundra circa 2000 at 90 m spatial resolution. This is an intermediate product between the Landsat ETM+ mosaic and the Landcover of Northern Canada (Olthof et al. 2014) however it provides a class describing bare ground (Lichen-Barren) that filled a need in this analysis and has no direct correlate in the final product.

Earth Observation for Sustainable Development of Forests (EOSD): The EOSD dataset (Wulder et al. 2008) was evaluated but found to be too inconsistent for use in this region with significant and obvious errors between adjacent Landsat scenes that were used as input.

Land Cover of Queen Maud Gulf Bird Sanctuary (LCQMGBS): The LCQMGBS (Didiuk and Ferguson 2005) is derived from Landsat 5 Thematic Mapper data, provides the necessary detail but does not cover the entire study domain. In addition, these data are outdated as they are based solely on Landsat data from 1986 – 1992.

Globeland30: The Globeland30 (Zheng 2014) dataset is derived from Landsat at 30 m spatial resolution circa 2010 but provides a single class for “tundra” which covers most of the study domain, hence did not provide anything to the analysis.

A.2: Vegetation Cover and Condition

Circumpolar Arctic Vegetation Map (CAVM): static map circa 1995 produced at 1:4,000,000 scale to depict Arctic vegetation communities (Walker et al. 2005). This dataset is derived from coarse resolution Advanced Very High Resolution Radiometer (AVHRR) data but is relevant because it was created with expert knowledge and field data providing delineations of the vegetation that are unmatched in contemporary and finer spatial resolution products.

MODIS Evapotranspiration (ET): produced annually at 500 m spatial resolution to depict the sum of evapotranspiration for the year (Mu et al. 2011). While we would have preferred to test evaporation and transpiration separately to determine the influence on surface water in the region, only combined evapotranspiration product exists.

Landsat NDVI vegetation fraction: median Normalized Difference Vegetation Index (NDVI) from August Landsat overpasses in the study region from 2011 – 2015 at 30 m spatial resolution. Vegetation fraction has been shown to be related to NDVI (Quarmby et al. 1993). Here a median NDVI over 5 years of peak growing season data is used as a surrogate to detect relationships between vegetation cover to surface water trend. Landsat NDVI is used because it provides information at finer spatial resolution than is possible with either the CAVM or the MODIS vegetation products.

Landsat NDVI trend: trend in maximum NDVI from July/August Landsat overpasses from 1985 – 2015 at 30 m spatial resolution (Ju and Masek 2016). Numerous studies have shown that there is a mostly positive trend in NDVI in large areas of the tundra (Goetz et al. 2010; Ju and Masek 2016; McManus et al. 2012;

Neigh et al. 2008). Here the trend in NDVI is tested against the trend in surface water extent.

A.3: Other Surface and Subsurface Features

Soil Landscapes of Canada (SLCWG): provides a static representation of soil regions at 1:1,000,000 scale circa 1991 (SLCWG 2010). This dataset provides generalized regions that relate to likely subsurface conditions.

Canadian Digital Elevation Model (CDEM): provides a static representation of elevation above sea level at 1:50,000 spatial resolution (Natural Resources Canada 2012). This is used to show the topography of the region and to calculate the slope, both of which help to determine the flow of surface water.

National Hydro Network (NHN): is derived from the CDEM and provides delineation of sub watersheds and connectivity at 1:50,000 scale (Natural Resources Canada 2007). The watershed delineations provide a hydrologically based subdivision of the study region to explore localized relationships with the surface water trend. An attribute of connectivity is also available for individual water bodies shown in the data set, this was linked to the water body data used in this study to provide an additional descriptive layer.

ArcticDEM: provides a static representation of elevation above sea level relative to the EGM96 geoid at 5 m spatial resolution. This product is generated by stereoscopy with image pairs from Digital Globe WorldView satellites (PGC 2017). At the time of this writing the current release is version 1 and it covers areas north of 60° N. However, there are many omissions due to unavailable data or poor correlation due to steep terrain or cloud cover. Future releases are supposed to solve these problems but

for the study area of this project the ArcticDEM was found to have too many holes to be useful for analysis.

A.4: Weather/Climate Data

Temperature and Precipitation Observations: Canadian Weather Service maintains weather stations throughout Canada and provides the data through publicly accessible portals. The data from four weather stations in and around study domain were used to determine temperature and precipitation for the study domain over the period of interest (CCN 2017).

Daymet: The Daymet version 3 (Thornton et al. 2017) is a raster dataset derived from interpolation of weather station data to a spatial resolution of 1 km. This is an exciting data set to use because it provides information on temperature and precipitation data at a much finer resolution than is usually available. Since the data set is an interpolation of point data it is better in places where there is a high density of weather stations. In the Arctic the network of weather stations is sparse which yields a poor interpolation. This was found to be the case in our study region, hence this data set was not used for the analysis of weather.

Modern Era Retrospective Reanalysis (MERRA): Reanalysis data are generated by assimilating all available observations into a modeling framework to provide consistent estimates of conditions on a global basis (Rienecker et al. 2011). The output from MERRA-2 is ~ 60 km resolution which yields 100 pixels over our study region. The advantage of reanalysis data is that they use all available data to initiate the model which is then run globally. This mitigates, to some extent, the lack of specific observations in certain locations. However, the limitation is that localized

variability cannot be captured based on the generalized model. There are hundreds of variables available through MERRA describing various aspects of the atmosphere and near surface conditions. Investigation of the usefulness of these variables is too much to be included in this manuscript and will be the topic of future work.

From the initial analysis of spatial patterns of change in surface water extent it is clear that not only are there distinct groupings of change but also there is a north west to south east directionality to the change. This is clearer when the data are arranged with the watersheds showing net change for each watershed. The watershed delineation provides the finest logical delineation of the study area while also providing an ecologically and specifically hydrologically based boundary. The DEM shows that the study region is largely flat with only a few rolling ridges in the south west and the slope over most of the region is $< 2^\circ$. Connectivity of individual water bodies was assessed using the attribute from the water bodies identified in the NHN and intersected with water bodies used in this study. This connectivity is determined based on the 30 m CDEM which may not have enough detail to capture micro topography that have been explored in other studies (Yang and Chu 2013). The current version (v2.0) of the 5 m spatial resolution Arctic DEM (PGC 2017) is insufficient to produce new watershed delineations or relative height because of extensive “no data” areas. The next version of the ArcticDEM at 5 m spatial resolution should have more complete coverage which will enable further analysis of connectivity in this region.

Datasets describing the subsurface characteristics of the region are scarce, with only the Soil Landscapes of Canada (SLCWG 2010) and the Circum Arctic

Permafrost Map (Brown et al. 1998) providing complete coverage of the entire region. The permafrost map indicates the entire region is underlain with continuous permafrost and the soils dataset provides only a broad overview of “soil regions” rather than a comprehensive description of variations in subsurface soils. It is beyond the scope of this study to do the fieldwork necessary to develop a comprehensive dataset of subsurface conditions, therefore proxy or surrogate data must be used to infer subsurface characteristics based on surface characteristics.

A composite of NDVI values from Landsat was created using data for the month of August from 5 years (2010 – 2014) by median value composite. The median NDVI value per watershed was extracted, using zonal statistics. The values were plotted against the percent surface water change with an $R^2=0.11$. Though the relationship is weak it does suggest that with decreasing NDVI there is a decrease in surface water extent.

The trend in NDVI was determined from a time series of maximum annual NDVI (from July/August observations) using Landsat data from Landsat 5 and 7 from 1985 to 2015. An increasing NDVI trend is associated with enhanced vegetation productivity and/or an expansion of photosynthetically active vegetation (more green plants result in higher NDVI). A decreasing NDVI trend is associated with decreased vegetation productivity or a decrease in coverage of photosynthetically active vegetation (more bare soil/rock or surface water results in lower NDVI). When NDVI trend is plotted against percent surface water change the resulting correlation is poor with an $R^2=0.04$. It should be noted that the NDVI trend data used here (Ju and Masek 2016) used an internally derived water mask that differed significantly from

the water bodies delineated for this analysis. This may have resulted in a slight bias in this measure since NDVI values near water bodies were masked, however this is mitigated by taking the mean trend per watershed rather than a pixel by pixel analysis.

Annual evapotranspiration from MODIS was related to watersheds for each year from 2000 to 2015 by taking the sum of ET for the watershed. Ordinary least square regression was applied to the time series of ET to determine a trend in ET over the available data. The trend in ET per watershed was then related to percent surface water change per watershed with an $R^2=0.15$.

Building on the relationship between change and the CAVM – Cryptogram Barren Complex shown in figure 3.7, several datasets were tested that describe the amount of bare ground on the surface. The VCF percent bare describes each 250 m pixel as a fraction of the surface that is non-vegetated which means that it could be bare ground, rock, ice, snow, or water. The correlation between VCF percent bare and percent surface water change per watershed was poor with $R^2 < 0.05$.

Bibliography

- ACIA (2005). Impacts of a Warming Arctic. In, *Arctic Climate Impact Assessment (ACIA)* (p. 144). New York, NY USA
- Adam, J.C., & Lettenmaier, D.P. (2003). Adjustment of global gridded precipitation for systematic bias. *Journal of Geophysical Research: Atmospheres*, 108
- Adrian, R., O'Reilly, C.M., Zagarese, H., Baines, S.B., Hessen, D.O., Keller, W., Livingstone, D.M., Sommaruga, R., Straile, D., Van Donk, E., Weyhenmeyer, G.A., & Winder, M. (2009). Lakes as sentinels of climate change. *Limnology and Oceanography*, 54, 2283-2297
- AMAP (2017). Snow, water, ice and permafrost in the Arctic (SWIPA). In (p. 269). Oslo, Norway: Arctic Council
- Andresen, C.G., & Lougheed, V.L. (2015). Disappearing Arctic tundra ponds: Fine-scale analysis of surface hydrology in drained thaw lake basins over a 65 year period (1948-2013). *Journal of Geophysical Research: Biogeosciences*, n/a-n/a
- Arctic Council. (2018). <http://www.arctic-council.org/index.php/en/about-us>. In Decker, M., Brunke, M.A., Wang, Z., Sakaguchi, K., Zeng, X., & Bosilovich, M.G. (2012). Evaluation of the Reanalysis Products from GSFC, NCEP, and ECMWF Using Flux Tower Observations. *Journal of Climate*, 25, 1916-1944
- Arp, C.D., Jones, B.M., Grosse, G., Bondurant, A.C., Romanovsky, V.E., Hinkel, K.M., & Parsekian, A.D. (2016). Threshold sensitivity of shallow Arctic lakes and sublake permafrost to changing winter climate. *Geophysical Research Letters*, 43, 6358-6365
- Arp, C.D., Jones, B.M., Liljedahl, A.K., Hinkel, K.M., & Welker, J.A. (2015). Depth, ice thickness, and ice-out timing cause divergent hydrologic responses among Arctic lakes. *Water Resources Research*, 51, 9379-9401
- Barber, D.G., Lukovich, J.V., Keogak, J., Baryluk, S., Fortier, L., & Henry, G.H.R. (2008). The Changing Climate of the Arctic. *Arctic*, 61, 7-26
- Barnett, T.P., Adam, J.C., & Lettenmaier, D.P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438, 303-309
- Bartholome, E., & Belward, A.S. (2005). GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*, 26, 1959-1977

- Bates, B.C., Kundzewicz, Z.W., Wu, S., & Palutikof, J.P. (2008). Climate Change and Water: Technical Paper VI. In I.p.o.c. change (Ed.), *Technical paper of the Intergovernmental Panel on Climate Change* (p. 210). IPCC Secretariat, Geneva: IPCC
- Billings, W.D., Luken, J.O., Mortensen, D.A., & Peterson, K.M. (1982). Arctic Tundra - a Source or Sink for Atmospheric Carbon-Dioxide in a Changing Environment. *Oecologia*, *53*, 7-11
- Bonan, G.B. (1995). Sensitivity of a GCM Simulation to Inclusion of Inland Water Surfaces. *Journal of Climate*, *8*, 2691-2704
- Bosilovich, M.G., Lucchesi, R., & Suarez, M.J. (2016). MERRA-2: File Specification. In GMAO (Ed.) (p. 73). http://gmao.gsfc.nasa.gov/pubs/office_notes. : NASA Goddard Space Flight Center
- Bouchard, F., Turner, K.W., MacDonald, L.A., Deakin, C., White, H., Farquharson, N., Medeiros, A.S., Wolfe, B.B., Hall, R.I., Pienitz, R., & Edwards, T.W.D. (2013). Vulnerability of shallow subarctic lakes to evaporate and desiccate when snowmelt runoff is low. *Geophysical Research Letters*, *40*, 6112-6117
- Bowling, L.C., & Lettenmaier, D.P. (2010). Modeling the Effects of Lakes and Wetlands on the Water Balance of Arctic Environments. *Journal of Hydrometeorology*, *11*, 276-295
- Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5-32
- Briggs, M.A., Walvoord, M.A., McKenzie, J.M., Voss, C.I., Day-Lewis, F.D., & Lane, J.W. (2014). New permafrost is forming around shrinking Arctic lakes, but will it last? *Geophysical Research Letters*, *41*, 1585-1592
- 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*, 621-649
- Bromley, R.G. (1996). Characteristics and management implications of the spring waterfowl hunt in the western Canadian arctic, northwest territories. *Arctic*, *49*, 70-85
- Brown, J., Ferrians, O.J.J., Heginbottom, J.A., & Melnikov, E.S. (1998). Circum-arctic map of permafrost and ground ice conditions. In N.S.a.I.D. Center (Ed.). Boulder, CO, USA: National Snow and Ice Data Center
- Brown, L., & Young, K.L. (2006). Assessment of three mapping techniques to delineate lakes and ponds in a Canadian High Arctic wetland complex. *Arctic*, *59*, 283-293

- Brown, R.D., & Brasnett, B. (2010). Canadian Meteorological Centre (CMC) Daily Snow Depth Analysis Data, Version 1. In C.M. Centre (Ed.). Boulder, CO, USA: National Snow and Ice Data Center
- Bunn, A.G., Goetz, S.J., Kimball, J.S., & Zhang, K. (2007). Northern High-Latitude Ecosystems Respond to Climate Change. *Eos, Transactions American Geophysical Union*, 88, 333-335
- Camill, P. (2005). Permafrost thaw accelerates in boreal peatlands during late-20th century climate warming. *Climatic Change*, 68, 135-152
- Carroll, M., & Loboda, T. (2017). Multi-Decadal Surface Water Dynamics in North American Tundra. *Remote Sensing*, 9, 497
- Carroll, M., Wooten, M., DiMiceli, C., Sohlberg, R., & Kelly, M. (2016a). Quantifying Surface Water Dynamics at 30 Meter Spatial Resolution in the North American High Northern Latitudes 1991–2011. *Remote Sensing*, 8, 622
- Carroll, M.L., & Loboda, T.V. (2018). The sign, magnitude and potential drivers of change in surface water extent in Canadian tundra. *Environmental Research Letters*
- Carroll, M.L., Townshend, J., Hansen, M., DiMiceli, C., Sohlberg, R., & Wurster, K. (2011a). Vegetative Cover Conversion and Vegetation Continuous Fields. In Ramachandran, B., C.O. Justice, & M. Abrams (Eds.), *Land Remote Sensing and Global Environmental Change: NASA's Earth Observing System and the Science of ASTER and MODIS*. New York: Springer Verlag
- Carroll, M.L., Townshend, J.R., DiMiceli, C.M., Noojipady, P., & Sohlberg, R.A. (2009). A new global raster water mask at 250 m resolution. *International Journal of Digital Earth*, 2, 291-308
- Carroll, M.L., Townshend, J.R.G., DiMiceli, C.M., Loboda, T., & Sohlberg, R.A. (2011b). Shrinking lakes of the Arctic: Spatial relationships and trajectory of change. *Geophysical Research Letters*, 38
- Carroll, M.L., Wooten, M.R., Dimiceli, C.M., Sohlberg, R.A., & Townshend, J.R.G. (2016b). ABoVE: Surface Water Extent, Boreal and Tundra Regions, North America, 1991-2011. In. Oak Ridge, Tennessee, USA: ORNL DAAC, Oak Ridge, Tennessee, USA.
- CCN (2017). Canadian Climate Normals data. In G.o. Canada (Ed.). http://climate.weather.gc.ca/climate_normals: Government of Canada
- Chapin, F.S., 3rd, Sturm, M., Serreze, M.C., McFadden, J.P., Key, J.R., Lloyd, A.H., McGuire, A.D., Rupp, T.S., Lynch, A.H., Schimel, J.P., Beringer, J., Chapman, W.L.,

Epstein, H.E., Euskirchen, E.S., Hinzman, L.D., Jia, G., Ping, C.L., Tape, K.D., Thompson, C.D., Walker, D.A., & Welker, J.M. (2005). Role of land-surface changes in arctic summer warming. *Science*, *310*, 657-660

Chapin, F.S., McGuire, A.D., Randerson, J., Pielke, R., Baldocchi, D., Hobbie, S.E., Roulet, N., Eugster, W., Kasischke, E., Rastetter, E.B., Zimov, S.A., & Running, S.W. (2000). Arctic and boreal ecosystems of western North America as components of the climate system. *Global Change Biology*, *6*, 211-223

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., & Mills, J. (2015). Global land cover mapping at 30m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, *103*, 7-27

Chen, M., Rowland, J.C., Wilson, C.J., Altmann, G.L., & Brumby, S.P. (2013). The Importance of Natural Variability in Lake Areas on the Detection of Permafrost Degradation: A Case Study in the Yukon Flats, Alaska. *Permafrost and Periglacial Processes*, *24*, 224-240

Cohen, W.B., & Goward, S.N. (2004). Landsat's role in ecological applications of remote sensing. *Bioscience*, *54*, 535-545

Cole, J.J., Prairie, Y.T., Caraco, N.F., McDowell, W.H., Tranvik, L.J., Striegl, R.G., Duarte, C.M., Kortelainen, P., Downing, J.A., Middelburg, J.J., & Melack, J. (2007). Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. *Ecosystems*, *10*, 171-184

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W.J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A.J., & Wehner, M. (2013). Long-term Climate Change: Projections, Commitments and Irreversibility. In T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (Ed.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N., & Vitart, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*, 553-597

Defourny, P., Vancutsem, C., Bicheron, P., Brockmann, C., Nino, F., Schouten, L., & Leroy, M. (2006). GLOBCOVER: a 300 m global land cover product for 2005 using Envisat MERIS time series. In, *Proceedings of the ISPRS Commission VII mid-term symposium, Remote sensing: from pixels to processes* (pp. 8-11): Enschede, the Netherlands

Didiuk, A.B., & Ferguson, R.S. (2005). Land cover mapping of Queen Maud Gulf Migratory Bird Sanctuary, Nunavut. . In C.W. Service (Ed.) (p. 36)

DiMiceli, C.M., Carroll, M.L., Sohlberg, R.A., Huang, C., Hansen, M.C., & Townshend, J.R.G. (2011). Annual Global Automated MODIS Vegetation Continuous Fields (MOD44B) at 250 m Spatial Resolution for Data Years Beginning Day 65, 2000 - 2010, Collection 5 Percent Tree Cover. In. University of Maryland, College Park, MD, USA: <http://landcover.org/data/vcf>

Downing, J.A. (2010). Emerging global role of small lakes and ponds: little things mean a lot. *Limnetica*, 29, 9-23

Downing, J.A., Prairie, Y.T., Cole, J.J., Duarte, C.M., Tranvik, L.J., Striegl, R.G., McDowell, W.H., Kortelainen, P., Caraco, N.F., Melack, J.M., & Middelburg, J.J. (2006). The global abundance and size distribution of lakes, ponds, and impoundments. *Limnology and Oceanography*, 51, 2388-2397

Easterling, D.R., Horton, B., Jones, P.D., Peterson, T.C., Karl, T.R., Parker, D.E., & Salinger, M.J. (1997). Maximum and Minimum Temperature Trends for the Globe. *Science*, 277, 364-367

Feng, M., Sexton, J.O., Channan, S., & Townshend, J.R. (2015). A global, high-resolution (30-m) inland water body dataset for 2000: first results of a topographic-spectral classification algorithm. *International Journal of Digital Earth*, 1-21

Francis, J.A., Chan, W.H., Leathers, D.J., Miller, J.R., & Veron, D.E. (2009a). Winter Northern Hemisphere weather patterns remember summer Arctic sea-ice extent. *Geophysical Research Letters*, 36

Francis, J.A., White, D.M., Cassano, J.J., Gutowski, W.J., Hinzman, L.D., Holland, M.M., Steele, M.A., & Vorosmarty, C.J. (2009b). An arctic hydrologic system in transition: Feedbacks and impacts on terrestrial, marine, and human life. *Journal of Geophysical Research-Biogeosciences*, 114

French, N.H.F., Jenkins, L.K., Loboda, T.V., Flannigan, M., Jandt, R., Bourgeau-Chavez, L.L., & Whitley, M. (2015). Fire in arctic tundra of Alaska: past fire activity, future fire potential, and significance for land management and ecology. *International Journal of Wildland Fire*, 24, 1045-1061

- Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114, 168-182
- Gelaro, R., McCarty, W., Suarez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G.K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., & Zhao, B. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, 30, 5419-5454
- Gibson, J.J., & Edwards, T.W.D. (2002). Regional water balance trends and evaporation-transpiration partitioning from a stable isotope survey of lakes in northern Canada. *Global Biogeochemical Cycles*, 16, 10-11-10-14
- Globe, D. (2016). World View 2 Multi-Spectral Data. In D. Globe (Ed.)
Goetz, S.J., Epstein, H.E., Bhatt, U.S., Jia, G.J., Kaplan, J.O., Lischke, H., Yu, Q., Bunn, A., Lloyd, A.H., Alcaraz-Segura, D., Beck, P.S.A., Comiso, J.C., Reynolds, M.K., & Walker, D.A. (2010). Recent Changes in Arctic Vegetation: Satellite Observations and Simulation Model Predictions, 9-36
- Goward, S., Arvidson, T., Williams, D., Faundeen, J., Irons, J., & Franks, S. (2006). Historical Record of Landsat Global Coverage. *Photogrammetric Engineering & Remote Sensing*, 72, 1155-1169
- Grosse, G., Romanovsky, V., Jorgenson, M.T., Anthony, K.W., Brown, J., & Overduin, P.P. (2011). Vulnerability and feedbacks of permafrost to climate change. *EOS Transactions*, 92, 73-80
- Grosse, G., Romanovsky, V., Walter, K., Morgenstern, A., Lantuit, H., & Zimov, S. (2008). Distribution of Thermokarst Lakes and Ponds at Three Yedoma Sites in Siberia. In D.L.K.a.K.M. Hinkel (Ed.), *Proceedings of the 9th International Conference on Permafrost* (pp. 551-556). Fairbanks, Alaska: Institute of Northern Engineering, University of Alaska, Fairbanks, Alaska.
- Hansen, M., DeFries, R.S., Townshend, J.R.G., Carroll, M., DiMiceli, C., & Sohlberg, R. (2003). Global percent tree cover at a spatial resolution of 500m: First results of the MODIS vegetation continuous fields algorithm. *Earth Interactions*, 7, 15
- Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., Dimiceli, C., & Carroll, M. (2002). Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. *Remote Sensing of Environment*, 83, 303-319

Hansen, M.C., Stehman, S.V., & Potapov, P.V. (2010). Quantification of global gross forest cover loss. *Proc Natl Acad Sci U S A*, 107, 8650-8655

Hernandez, P.A., Graham, C.H., Master, L.L., & Albert, D.L. (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, 29, 773-785

Hinkel, K.M., Jones, B.M., Eisner, W.R., Cuomo, C.J., Beck, R.A., & Frohn, R. (2007). Methods to assess natural and anthropogenic thaw lake drainage on the western Arctic coastal plain of northern Alaska. *Journal of Geophysical Research-Earth Surface*, 112

Hinkel, K.M., Paetzold, F., Nelson, F.E., & Bockheim, J.G. (2001). Patterns of soil temperature and moisture in the active layer and upper permafrost at Barrow, Alaska: 1993–1999. *Global and Planetary Change*, 29, 293-309

Hinzman, L.D., Bettez, N.D., Bolton, W.R., Chapin, F.S., Dyurgerov, M.B., Fastie, C.L., Griffith, B., Hollister, R.D., Hope, A., Huntington, H.P., Jensen, A.M., Jia, G.J., Jorgenson, T., Kane, D.L., Klein, D.R., Kofinas, G., Lynch, A.H., Lloyd, A.H., McGuire, A.D., Nelson, F.E., Oechel, W.C., Osterkamp, T.E., Racine, C.H., Romanovsky, V.E., Stone, R.S., Stow, D.A., Sturm, M., Tweedie, C.E., Vourlitis, G.L., Walker, M.D., Walker, D.A., Webber, P.J., Welker, J.M., Winker, K., & Yoshikawa, K. (2005). Evidence and implications of recent climate change in northern Alaska and other arctic regions. *Climatic Change*, 72, 251-298

Hinzman, L.D., Kane, D.L., Gieck, R.E., & Everett, K.R. (1991). Hydrologic and Thermal-Properties of the Active Layer in the Alaskan Arctic. *Cold Regions Science and Technology*, 19, 95-110

Hobbie, J.E. (1984). The ecology of tundra ponds of the Arctic coastal plain: A community profile. In F.a.W. Service (Ed.) (p. 64). Washington, DC: Fish and Wildlife Service

Huang, W., DeVries, B., Huang, C., Lang, M., Jones, J.W., Creed, I., & Carroll, M. (2018). Automated Extraction of Surface Water Extent from Sentinel-1 Data. *Remote Sensing*, *In review*

IARPC (2018). <https://www.iarpcollaborations.org/about.html>. In. <https://www.iarpcollaborations.org/about.html>: IARPC

IPCC (2014). IPCC (2014) Fifth Assessment Report. Cambridge: Cambridge University Press. In M.C. Parry, OF; Palutikof, JP; van der Linden, PJ; Hanson, CE (Ed.). Cambridge

Jaynes, E.T. (1957). Information theory and statistical mechanics. *The Physical Review*, 106, 620-630

- Jepsen, S.M., Voss, C.I., Walvoord, M.A., Minsley, B.J., & Rover, J. (2013). Linkages between lake shrinkage/expansion and sublacustrine permafrost distribution determined from remote sensing of interior Alaska, USA. *Geophysical Research Letters*, *40*, 882-887
- Jepsen, S.M., Walvoord, M.A., Voss, C.I., & Rover, J. (2016). Effect of permafrost thaw on the dynamics of lakes recharged by ice-jam floods: case study of Yukon Flats, Alaska. *Hydrological Processes*, *30*, 1782-1795
- Jones, B.M., Arp, C.D., Hinkel, K.M., Beck, R.A., Schmutz, J.A., & Winston, B. (2009). Arctic lake physical processes and regimes with implications for winter water availability and management in the National Petroleum Reserve Alaska. *Environmental Management*, *43*, 1071-1084
- Jones, J.W. (2015). Efficient wetland surface water detection and monitoring via Landsat: Comparison with in situ data from the Everglades Depth Estimation Network (EDEN). *Remote Sensing*, *7*, 12503-12538
- Jones, P.D., New, M., Parker, D.E., Martin, S., & Rigor, I.G. (1999). Surface air temperature and its changes over the past 150 years. *Reviews of Geophysics*, *37*, 173-199
- Jorgenson, M.T., & Grosse, G. (2016). Remote Sensing of Landscape Change in Permafrost Regions. *Permafrost and Periglacial Processes*, *27*, 324-338
- Jorgenson, M.T., Romanovsky, V., Harden, J., Shur, Y., O'Donnell, J., Schuur, E.A.G., Kanevskiy, M., & Marchenko, S. (2010). Resilience and vulnerability of permafrost to climate change. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, *40*, 1219-1236
- Jorgenson, M.T., Shur, Y.L., & Pullman, E.R. (2006). Abrupt increase in permafrost degradation in Arctic Alaska. *Geophysical Research Letters*, *33*
- Ju, J., & Masek, J.G. (2016). The vegetation greenness trend in Canada and US Alaska from 1984–2012 Landsat data. *Remote Sensing of Environment*, *176*, 1-16
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R., & Joseph, D. (1996). The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, *77*, 437-471
- Karlsson, J.M., Lyon, S.W., & Destouni, G. (2012). Thermokarst lake, hydrological flow and water balance indicators of permafrost change in Western Siberia. *Journal of Hydrology*, *464*, 459-466

Kaufman, D.S., Schneider, D.P., McKay, N.P., Ammann, C.M., Bradley, R.S., Briffa, K.R., Miller, G.H., Otto-Bliesner, B.L., Overpeck, J.T., & Vinther, B.M. (2009). Recent warming reverses long-term arctic cooling. *Science*, 325, 1236-1239

King, R.S., Baker, M.E., Whigham, D.F., Weller, D.E., Jordan, T.E., Kazyak, P.F., & Hurd, M.K. (2005). Spatial Considerations for Linking Watershed Land Cover to Ecological Indicators in Streams. *Ecological Applications*, 15, 137-153

Kokelj, S.V., & Jorgenson, M.T. (2013). Advances in Thermokarst Research. *Permafrost and Periglacial Processes*, 24, 108-119

Kortelainen, P., Pajunen, H., Rantakari, M., & Saarnisto, M. (2004). A large carbon pool and small sink in boreal Holocene lake sediments. *Global Change Biology*, 10, 1648-1653

Lam, N.S.-N., & Quattrochi, D.A. (1992). On the Issues of Scale, Resolution, and Fractal Analysis in the Mapping Sciences*. *The Professional Geographer*, 44, 88-98

Langer, M., Westermann, S., Boike, J., Kirillin, G., Grosse, G., Peng, S., & Krinner, G. (2016). Rapid degradation of permafrost underneath waterbodies in tundra landscapes-Toward a representation of thermokarst in land surface models. *Journal of Geophysical Research: Earth Surface*, 121, 2446-2470

Lawrence, D.M., Slater, A.G., Tomas, R.A., Holland, M.M., & Deser, C. (2008). Accelerated Arctic land warming and permafrost degradation during rapid sea ice loss. *Geophysical Research Letters*, 35

Lehner, B., & Doll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, 296, 1-22

Liljedahl, A.K., Boike, J., Daanen, R.P., Fedorov, A.N., Frost, G.V., Grosse, G., Hinzman, L.D., Iijma, Y., Jorgenson, J.C., Matveyeva, N., Necsoiu, M., Reynolds, M.K., Romanovsky, V.E., Schulla, J., Tape, K.D., Walker, D.A., Wilson, C.J., Yabuki, H., & Zona, D. (2016). Pan-Arctic ice-wedge degradation in warming permafrost and its influence on tundra hydrology. *Nature Geoscience*, 9, 312-318

Linacre, E.T. (1992). *Climate data and Resources*. Routledge

McGuire, A.D., Chapin, F.S., Walsh, J.E., & Wirth, C. (2006). Integrated regional changes in arctic climate feedbacks: Implications for the global climate system. *Annual Review of Environment and Resources*, 31, 61-91

McGuire, A.D., Walsh, J.E., Kimball, J.S., Klein, J.S., Euskirchen, S.E., Drobot, S., Herzfeld, U.C., Maslanik, J., Lammers, R.B., Rawlins, M.A., Vorosmarty, C.J., Rupp,

- T.S., Wu, W., & Calef, M. (2008). The western Arctic Linkage Experiment (WALE): Overview and synthesis. *Earth Interactions*, *12*, 1-13
- McManus, K.M., Morton, D.C., Masek, J.G., Wang, D.D., Sexton, J.O., Nagol, J.R., Ropars, P., & Boudreau, S. (2012). Satellite-based evidence for shrub and graminoid tundra expansion in northern Quebec from 1986 to 2010. *Global Change Biology*, *18*, 2313-2323
- Merow, C., Smith, M.J., & Silander, J.A. (2013). A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, *36*, 1058-1069
- Miller, G.H., Brigham-Grette, J., Alley, R.B., Anderson, L., Bauch, H.A., Douglas, M.S.V., Edwards, M.E., Elias, S.A., Finney, B.P., Fitzpatrick, J.J., Funder, S.V., Herbert, T.D., Hinzman, L.D., Kaufman, D.S., MacDonald, G.M., Polyak, L., Robock, A., Serreze, M.C., Smol, J.P., Spielhagen, R., White, J.W.C., Wolfe, A.P., & Wolff, E.W. (2010). Temperature and precipitation history of the Arctic. *Quaternary Science Reviews*, *29*, 1679-1715
- Minsley, B.J., Abraham, J.D., Smith, B.D., Cannia, J.C., Voss, C.I., Jorgenson, M.T., Walvoord, M.A., Wylie, B.K., Anderson, L., Ball, L.B., Deszcz-Pan, M., Wellman, T.P., & Ager, T.A. (2012). Airborne electromagnetic imaging of discontinuous permafrost. *Geophysical Research Letters*, *39*, n/a-n/a
- Moratto, Z.M., Broxton, M.J., Beyer, R.A., Lundy, M., & Husmann, K. (2010). Ames Stereo Pipeline, NASA's open source automated stereogrammetry software. In, *Lunar and Planetary Science Conference*. The Woodlands, Texas
- Mu, Q., Zhao, M., & Running, S.W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, *115*, 1781-1800
- Muster, S., Heim, B., Abnizova, A., & Boike, J. (2013). Water Body Distributions Across Scales: A Remote Sensing Based Comparison of Three Arctic TundraWetlands. *Remote Sensing*, *5*, 1498-1523
- Natural Resources Canada, C.f.T.I. (2007). National Hydro Network [digital data]. In N.R. Canada (Ed.). Available online: <http://www.geobase.ca/geobase/en/data/nhn/index.html>
- Natural Resources Canada, C.f.T.I. (2012). Canadian Digital Elevation Model (CDEM). In C.f.T.I. Natural Resources Canada (Ed.). online
- Neigh, C., Tucker, C., & Townshend, J. (2008). North American vegetation dynamics observed with multi-resolution satellite data. *Remote Sensing of Environment*, *112*, 1749-1772

Neigh, C.S.R., Masek, J.G., & Nickeson, J.E. (2013). High-Resolution Satellite Data Open for Government Research. *Eos, Transactions American Geophysical Union*, *94*, 121-123

Nitze, I., Grosse, G., Jones, B., Arp, C., Ulrich, M., Fedorov, A., & Veremeeva, A. (2017). Landsat-Based Trend Analysis of Lake Dynamics across Northern Permafrost Regions. *Remote Sensing*, *9*, 640

NOAA (2015). Using satellites to predict malaria risk. In *NESDIS New Archive*. http://www.nesdis.noaa.gov/news_archives/using_satellite_to_predict_malaria_risk.html NOAA/NESDIS

NRC, C. (2015). Principal mineral areas of Canada. In N.R.C. Minerals and metals sector (Ed.): Natural Resources Canada

O'Reilly, C.M., Sharma, S., Gray, D.K., Hampton, S.E., Read, J.S., Rowley, R.J., Schneider, P., Lenters, J.D., McIntyre, P.B., Kraemer, B.M., Weyhenmeyer, G.A., Straile, D., Dong, B., Adrian, R., Allan, M.G., Anneville, O., Arvola, L., Austin, J., Bailey, J.L., Baron, J.S., Brookes, J.D., de Eyto, E., Dokulil, M.T., Hamilton, D.P., Havens, K., Hetherington, A.L., Higgins, S.N., Hook, S., Izmest'eva, L.R., Joehnk, K.D., Kangur, K., Kasprzak, P., Kumagai, M., Kuusisto, E., Leshkevich, G., Livingstone, D.M., MacIntyre, S., May, L., Melack, J.M., Mueller-Navarra, D.C., Naumenko, M., Noges, P., Noges, T., North, R.P., Plisnier, P.-D., Rigosi, A., Rimmer, A., Rogora, M., Rudstam, L.G., Rusak, J.A., Salmaso, N., Samal, N.R., Schindler, D.E., Schladow, S.G., Schmid, M., Schmidt, S.R., Silow, E., Soylyu, M.E., Teubner, K., Verburg, P., Voutilainen, A., Watkinson, A., Williamson, C.E., & Zhang, G. (2015). Rapid and highly variable warming of lake surface waters around the globe. *Geophysical Research Letters*, *42*, 10,773-710,781

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., & Wulder, M.A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, *148*, 42-57

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., & Kassem, K.R. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth. *Bioscience*, *51*, 933

Olthof, I., & Fraser, R. (2014). Detecting Landscape Changes in High Latitude Environments Using Landsat Trend Analysis: 2. Classification. *Remote Sensing*, *6*, 11558-11578

Olthof, I., Latifovic, R., & Pouliot, D. (2014). Development of a circa 2000 land cover map of northern Canada at 30 m resolution from Landsat. *Canadian Journal of Remote Sensing*, *35*, 152-165

Pavelsky, T.M., & Smith, L.C. (2008). Remote sensing of hydrologic recharge in the Peace-Athabasca Delta, Canada. *Geophysical Research Letters*, *35*

Pearson, R.G., Raxworthy, C.J., Nakamura, M., & Townsend Peterson, A. (2006). Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, *34*, 102-117

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*, 418-422

PGC (2017). Arctic DEM. In U.o.M.S.P.M. Polar Geospatial Center (Ed.). St. Paul, Minnesota, USA: Polar Geospatial Center, University of Minnesota St. Paul Minnesota

Phillips, S.J. (2005). A Brief Tutorial on Maxent. In (pp. 1-38). Princeton, NJ: AT&T Research Labs, Princeton University

Pietroniro, A., Prowse, T., & Peters, D. (1999). Hydrologic assessment of an inland fresh water delta using multi-temporal satellite remote sensing. *Hydrological Processes*, *13*

Post, W.M., Peng, T.H., Emanuel, W.R., King, A.W., Dale, V.H., & Deangelis, D.L. (1990). The Global Carbon-Cycle. *American Scientist*, *78*, 310-326

Poulter, B., Bousquet, P., Canadell, J.G., Ciais, P., Pregon, A., Saunois, M., Arora, V.K., Beerling, D.J., Brovkin, V., Jones, C.D., Joos, F., Gedney, N., Ito, A., Kleinen, T., Koven, C.D., McDonald, K., Melton, J.R., Peng, C., Peng, S., Prigent, C., Schroeder, R., Riley, W.J., Saito, M., Spahni, R., Tian, H., Taylor, L., Viovy, N., Wilton, D., Wiltshire, A., Xu, X., Zhang, B., Zhang, Z., & Zhu, Q. (2017). Global wetland contribution to 2000–2012 atmospheric methane growth rate dynamics. *Environmental Research Letters*, *12*, 094013

Prigent, C., Papa, F., Aires, F., Jimenez, C., Rossow, W.B., & Matthews, E. (2012). Changes in land surface water dynamics since the 1990s and relation to population pressure. *Geophysical Research Letters*, *39*

Putkonen, J., Grenfell, T.C., Rennert, K., Bitz, C., Jacobson, P., & Russell, D. (2009). Rain on Snow-Little Understood Killer in the North. *Eos, Transactions American Geophysical Union*, *90*, 221-222

- Quarmby, N.A., Milnes, M., Hindle, T.L., & Silleos, N. (1993). The use of multi-temporal NDVI measurements from AVHRR data for crop yield estimation and prediction. *International Journal of Remote Sensing*, *14*, 199-210
- Rawlins, M.A., Steele, M., Holland, M.M., Adam, J.C., Cherry, J.E., Francis, J.A., Groisman, P.Y., Hinzman, L.D., Huntington, T.G., Kane, D.L., Kimball, J.S., Kwok, R., Lammers, R.B., Lee, C.M., Lettenmaier, D.P., McDonald, K.C., Podest, E., Pundsack, J.W., Rudels, B., Serreze, M.C., Shiklomanov, A., Skagseth, Ø., Troy, T.J., Vörösmarty, C.J., Wensnahan, M., Wood, E.F., Woodgate, R., Yang, D., Zhang, K., & Zhang, T. (2010). Analysis of the Arctic System for Freshwater Cycle Intensification: Observations and Expectations. *Journal of Climate*, *23*, 5715-5737
- Reichle, R.H., Liu, Q., Koster, R.D., Draper, C.S., Mahanama, S.P.P., & Partyka, G.S. (2017). Land Surface Precipitation in MERRA-2. *Journal of Climate*, *30*, 1643-1664
- Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R.D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C.R., Reichle, R., Robertson, F.R., Ruddick, A.G., Sienkiewicz, M., & Woollen, J. (2011). MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*, *24*, 3624-3648
- Roach, J.K., Griffith, B., & Verbyla, D. (2013). Landscape influences on climate-related lake shrinkage at high latitudes. *Glob Chang Biol*, *19*, 2276-2284
- Rover, J., Ji, L., Wylie, B.K., & Tieszen, L.L. (2012). Establishing water body areal extent trends in interior Alaska from multi-temporal Landsat data. *Remote Sensing Letters*, *3*, 595-604
- Salomon, J., Hodges, J., Friedl, M., Schaaf, C., Strahler, A., Gao, F., Schneider, A., Zhang, X., El Saleous, N., & Wolfe, R. (2004). Global Land-Water Mask Derived from MODIS Nadir BRDF-Adjusted Reflectances (NBAR) and the MODIS Land Cover Algorithm. In, *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International* (p. 241). Alaska
- Schnase, J.L., Duffy, D.Q., Tamkin, G.S., Nadeau, D., Thompson, J.H., Grieg, C.M., McInerney, M.A., & Webster, W.P. (2017). MERRA Analytic Services: Meeting the Big Data challenges of climate science through cloud-enabled Climate Analytics-as-a-Service. *Computers, Environment and Urban Systems*, *61*, 198-211
- Schneider, P., & Hook, S.J. (2010). Space observations of inland water bodies show rapid surface warming since 1985. *Geophysical Research Letters*, *37*, n/a-n/a
- Screen, J.A., & Simmonds, I. (2010). The central role of diminishing sea ice in recent Arctic temperature amplification. *Nature*, *464*, 1334-1337

- Serreze, M.C., Barrett, A.P., Stroeve, J.C., Kindig, D.N., & Holland, M.M. (2009). The emergence of surface-based Arctic amplification. *Cryosphere*, 3, 11-19
- Serreze, M.C., & Barry, R.G. (2011). Processes and impacts of Arctic amplification: A research synthesis. *Global and Planetary Change*, 77, 85-96
- Serreze, M.C., & Francis, J.A. (2006). The arctic amplification debate. *Climatic Change*, 76, 241-264
- Serreze, M.C., Walsh, J.E., Chapin, F.S., Osterkamp, T., Dyurgerov, M., Romanovsky, V., Oechel, W.C., Morison, J., Zhang, T., & Barry, R.G. (2000). Observational evidence of recent change in the northern high-latitude environment. *Climatic Change*, 46, 159-207
- Sexton, J.O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D.-H., Collins, K.M., Channan, S., DiMiceli, C., & Townshend, J.R. (2013). Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. *International Journal of Digital Earth*, 6, 427-448
- Sharma, S., Gray, D.K., Read, J.S., O'Reilly, C.M., Schneider, P., Qudrat, A., Gries, C., Stefanoff, S., Hampton, S.E., Hook, S., Lenters, J.D., Livingstone, D.M., McIntyre, P.B., Adrian, R., Allan, M.G., Anneville, O., Arvola, L., Austin, J., Bailey, J., Baron, J.S., Brookes, J., Chen, Y., Daly, R., Dokulil, M., Dong, B., Ewing, K., de Eyto, E., Hamilton, D., Havens, K., Haydon, S., Hetzenauer, H., Heneberry, J., Hetherington, A.L., Higgins, S.N., Hixson, E., Izmet'eva, L.R., Jones, B.M., Kangur, K., Kasprzak, P., Koster, O., Kraemer, B.M., Kumagai, M., Kuusisto, E., Leshkevich, G., May, L., MacIntyre, S., Muller-Navarra, D., Naumenko, M., Noges, P., Noges, T., Niederhauser, P., North, R.P., Paterson, A.M., Plisnier, P.D., Rigosi, A., Rimmer, A., Rogora, M., Rudstam, L., Rusak, J.A., Salmaso, N., Samal, N.R., Schindler, D.E., Schladow, G., Schmidt, S.R., Schultz, T., Silow, E.A., Straile, D., Teubner, K., Verburg, P., Voutilainen, A., Watkinson, A., Weyhenmeyer, G.A., Williamson, C.E., & Woo, K.H. (2015). A global database of lake surface temperatures collected by in situ and satellite methods from 1985-2009. *Sci Data*, 2, 150008
- Simmons, A., Uppala, S., Dee, D., & Kobayashi, S. (2006). ERAInterim: New ECMWF reanalysis products from 1989 onwards. In *ECMWF Newsletter* (pp. 25-35). Reading, United Kingdom
- Slater, A.G., Bohn, T.J., McCreight, J.L., Serreze, M.C., & Lettenmaier, D.P. (2007). A multimodel simulation of pan-Arctic hydrology. *Journal of Geophysical Research-Biogeosciences*, 112
- SLCWG (2010). Soil landscapes of Canada. In A.a.a.-f. Canada (Ed.). Canada: Agriculture and agri-food Canada

Smejkalova, T., Edwards, M.E., & Dash, J. (2016). Arctic lakes show strong decadal trend in earlier spring ice-out. *Sci Rep*, *6*, 38449

Smith, L.C., Sheng, Y., MacDonald, G.M., & Hinzman, L.D. (2005). Disappearing Arctic lakes. *Science*, *308*, 1429

Smol, J.P., & Douglas, M.S. (2007a). Crossing the final ecological threshold in high Arctic ponds. *Proc Natl Acad Sci U S A*, *104*, 12395-12397

Smol, J.P., & Douglas, M.S.V. (2007b). From controversy to consensus: making the case for recent climate change in the Arctic using lake sediments. *Frontiers in Ecology and the Environment*, *5*, 466-474

Stokstad, E. (2004). Defrosting the carbon freezer of the north. *Science*, *304*, 1618-1620

Storey, J., Choate, M., & Lee, K. (2014). Landsat 8 Operational Land Imager On-Orbit Geometric Calibration and Performance. *Remote Sensing*, *6*, 11127-11152

Stow, D.A., Hope, A., McGuire, D., Verbyla, D., Gamon, J., Huemmrich, F., Houston, S., Racine, C., Sturm, M., Tape, K., Hinzman, L., Yoshikawa, K., Tweedie, C., Noyle, B., Silapaswan, C., Douglas, D., Griffith, B., Jia, G., Epstein, H., Walker, D., Daeschner, S., Petersen, A., Zhou, L.M., & Myneni, R. (2004). Remote sensing of vegetation and land-cover change in Arctic Tundra Ecosystems. *Remote Sensing of Environment*, *89*, 281-308

Stroeve, J.C., Serreze, M.C., Holland, M.M., Kay, J.E., Malanik, J., & Barrett, A.P. (2011). The Arctic's rapidly shrinking sea ice cover: a research synthesis. *Climatic Change*, *110*, 1005-1027

Su, F.G., Adam, J.C., Trenberth, K.E., & Lettenmaier, D.P. (2006). Evaluation of surface water fluxes of the pan-Arctic land region with a land surface model and ERA-40 reanalysis. *Journal of Geophysical Research-Atmospheres*, *111*

Subin, Z.M., Riley, W.J., & Mironov, D. (2012). An improved lake model for climate simulations: Model structure, evaluation, and sensitivity analyses in CESM1. *Journal of Advances in Modeling Earth Systems*, *4*, 27

Tang, Q.H., Gao, H.L., Yeh, P., Oki, T., Su, F.G., & Lettenmaier, D.P. (2010). Dynamics of Terrestrial Water Storage Change from Satellite and Surface Observations and Modeling. *Journal of Hydrometeorology*, *11*, 156-170

Tarnocai, C., Kimble, J.M., Swanson, D., Goryachkin, S., Naumov, Y.M., Stolbovoi, V., Jakobsen, B., Broll, G., Montanarella, L., Arnoldussen, A., Arnalds, O., & Yli-Halla, M. (2002). Northern Circumpolar Soils. 1:10,000,000 scale map. In A.a.A.-

F.C. Research Branch (Ed.). Ottawa, Canada: National Snow and Ice Data Center, Boulder, CO.

Thornton, P.E., Thornton, M.M., Mayer, B.W., Wei, Y., Devarakonda, R., Vose, R.S., & Cook, R.B. (2017). Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3. . In O.R. ORNL DAAC, Tennessee, USA. (Ed.). ORNL DAAC, Oak Ridge, Tennessee, USA.: ORNL DAAC, Oak Ridge, Tennessee, USA.

Townshend, J.R., Masek, J.G., Huang, C., Vermote, E.F., Gao, F., Channan, S., Sexton, J.O., Feng, M., Narasimhan, R., Kim, D., Song, K., Song, D., Song, X.-P., Noojipady, P., Tan, B., Hansen, M.C., Li, M., & Wolfe, R.E. (2012). Global characterization and monitoring of forest cover using Landsat data: opportunities and challenges. *International Journal of Digital Earth*, 5, 373-397

Turetsky, M.R., Treat, C.C., Waldrop, M.P., Waddington, J.M., Harden, J.W., & McGuire, A.D. (2008). Short-term response of methane fluxes and methanogen activity to water table and soil warming manipulations in an Alaskan peatland. *Journal of Geophysical Research-Biogeosciences*, 113

USGS (2016). Landsat Surface Reflectance High Level Data Products.

Verpoorter, C., Kutser, T., Seekell, D.A., & Tranvik, L.J. (2014). A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, 41, 6396-6402

Walker, D.A., Reynolds, M.K., Daniëls, F.J.A., Einarsson, E., Elvebakk, A., Gould, W.A., Katenin, A.E., Kholod, S.S., Markon, C.J., Melnikov, E.S., Moskalenko, N.G., Talbot, S.S., & Yurtsev, B.A. (2005). The Circumpolar Arctic vegetation map. *Journal of Vegetation Science*, 16, 267

Walter, K.M., Smith, L.C., & Chapin, F.S., 3rd (2007). Methane bubbling from northern lakes: present and future contributions to the global methane budget. *Philos Transact A Math Phys Eng Sci*, 365, 1657-1676

Walter, K.M., Zimov, S.A., Chanton, J.P., Verbyla, D., & Chapin, F.S., 3rd (2006). Methane bubbling from Siberian thaw lakes as a positive feedback to climate warming. *Nature*, 443, 71-75

Watts, J.D., Kimball, J.S., Jones, L.A., Schroeder, R., & McDonald, K.C. (2012). Satellite Microwave remote sensing of contrasting surface water inundation changes within the Arctic–Boreal Region. *Remote Sensing of Environment*, 127, 223-236

White, D., Hinzman, L., Alessa, L., Cassano, J., Chambers, M., Falkner, K., Francis, J., Gutowski, W.J., Holland, M., Holmes, R.M., Huntington, H., Kane, D., Kliskey, A., Lee, C., McClelland, J., Peterson, B., Rupp, T.S., Straneo, F., Steele, M.,

- Woodgate, R., Yang, D., Yoshikawa, K., & Zhang, T. (2007). The arctic freshwater system: Changes and impacts. *Journal of Geophysical Research-Biogeosciences*, 112
- White, J.C., & Wulder, M.A. (2014). The Landsat observation record of Canada: 1972–2012. *Canadian Journal of Remote Sensing*, 39, 455-467
- Woo, M.-K., & Guan, X.J. (2006). Hydrological connectivity and seasonal storage change of tundra ponds in a polar oasis environment, Canadian High Arctic. *Permafrost and Periglacial Processes*, 17, 309-323
- Wulder, M.A., Cranny, M.M., Hall, R.J., Luther, J.E., Beaudoin, A., White, J.C., Goodenough, D.G., & Dechka, J.A. (2008). Satellite land cover mapping of Canada's forests: the EOSD land cover project. In J.C. Campbell, K.B. Jones, J.H. Smith, & M.T. Koeppe (Eds.), *North America Land Cover Summit* (pp. 21-30). Washington, DC, USA: American Association of Geographers
- Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., & Woodcock, C.E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, 122, 2-10
- Yang, J., & Chu, X. (2013). Effects of DEM Resolution on Surface Depression Properties and Hydrologic Connectivity. *Journal of Hydrologic Engineering*, 18, 1157-1169
- Yoshikawa, K., & Hinzman, L.D. (2003). Shrinking thermokarst ponds and groundwater dynamics in discontinuous permafrost near council, Alaska. *Permafrost and Periglacial Processes*, 14, 151-160
- Zheng, X. (2014). GlobeLand 30. In N.G.C.o. China (Ed.). Beijing, China