REMOTE SENSING OF RECENT CHANGES IN PERMAFROST-INFLUENCED WETLANDS

Investigations in thermokarst lake surface areal trends at multiple spatiotemporal scales

Young Tun Jan

St. Edmund's College

Scott Polar Research Institute

University of Cambridge

This dissertation is submitted for the degree of

Master of Philosophy

13 June 2013

DECLARATION FROM AUTHOR.

To the best of my knowledge, this dissertation represents my own work and conforms to the accepted standards of citation in those instances in which I have availed myself of the works of others. This dissertation is submitted according to the requirements of the Board of Graduate Studies, Department of Geography, and the Scott Polar Research Institute. This dissertation does not exceed the regulations on length and has not been submitted to any other university or institution for any degree, diploma, or similar qualification.

Young Tun Jan

Cambridge, United Kingdom
13 June 2013

ABSTRACT.

Changes in thermokarst lakes have been identified across many high-latitude ecosystems at different spatial and temporal scales. With the declassification of the Landsat archive in 2008, it is now possible to conduct a near-complete yearly time series of homogenous geospatial imagery to assess the trends seen in thermokarst lake surface area. By implementing an automated land-cover classification algorithm, this study examined the dynamics of lake surface area over different spatial and temporal scales in the Tuktoyaktuk Peninsula, Northwest Territories, Canada. Due to the presence of a statistically significant structural temporal break between the years 1997 and 1998, lake areal trends were estimated at two temporal scales, a longer term scale (1985 to 2011) and its component shorter scales (1985 to 1997 and 1998 to 2011). Large lakes saw the greatest changes both in lake areal increase and decreases at all temporal scales, and was suggested to drive the overall changes in surface area. In addition, regional differences were observed in the spatial distribution of individual lake area trends. On the broad scale, a latitudinal divide bifurcated the peninsula into two regions of approximately equal area, where the northern region exhibited general trends of lake areal decline, while the southern region exhibited general trends of lake areal increase. Within these regions, meso- and local hot and cold spots were identified, some that exhibited trends in concordance with local surroundings, while others represented local spatial heterogeneity in areal trends. The spatiotemporal trends in lake area were suggested to be influenced at varying scales by atmospheric and climate variables, and by ground characteristics such as coincident permafrost and surficial geology. As the Arctic continues to warm, a continued observation of thermokarst lake evolution over both broad landscapes and localised regions will be increasingly valuable to future studies that investigate the resulting transformation of the Arctic permafrost region.

TABLE OF CONTENTS.

LIST OF FIGURES.

LIST OF TABLES.

LIST OF EQUATIONS.

ACKNOWLEDGEMENTS.

This dissertation would not have been possible without the continual support and encouragement of my supervisor, Poul Christoffersen, who provided extensive advice and kept me calm throughout the course of this supervision—for that, I am very grateful and cannot thank you enough.

Thanks also to Toby Benham and Ruth Mugford, who helped me get the right foot in to the more quantitative aspects of this project. Zhe Zhu from Boston University gave advice on running and optimising the Fmask algorithm, and was gracious enough to let me use a beta version for testing purposes. Evan Miles and Allen Pope also provided help with image processing and various errors with the [now mostly debugged] code.

Across the pond, thanks go to Kayi Tsang, who stayed up with me through many late nights; and to my brother CJ Young, who helped me proofread the entire dissertation.

Lastly, to my SPRI MPhil mates—thanks for the great year, and through various ups and downs, we all came out of this together [mostly] in one piece.

1. INTRODUCTION.

1.1. Rationale.

Arctic tundra ecosystems have experienced substantial warming in recent decades, and are projected to continue warming into the future (Serreze et al. 2000). One of the most direct manifestations of polar climate warming on land is the thawing of permafrost, which in turn directly alters the landscape terrain of the tundra ecosystem, for example, through the changes seen in thermokarst thaw lake cover (Osterkamp et al. 2000, Jorgenson et al. 2010). In particular, thermokarst lakes are a primary mechanism of permafrost modification as they change in size, coalesce with nearby lakes, or migrate across the landscape (Kozlenko and Jeffries 2000).

A variety of remote sensing approaches have been previously implemented to monitor lake change in Arctic permafrost regions (e.g. Yoshikawa and Hinzman 2003, Smith et al. 2005, Grosse et al. 2006, Plug et al. 2008); with the declassification of military satellite imagery such as the Landsat Archive (Woodcock et al. 2008), these approaches have become increasingly automated due to the increase in the amount of input imagery. However, many of these studies have not adequately accounted for intra- and inter-annual variability that almost certainly affect change measured at multiple scales. Furthermore, many of these studies were conducted over several decades with little reference to intra- and inter-annual variation in surface processes, or the large-scale atmospheric processes that likely drive such processes. Understanding the response of thermokarst lakes to changes in permafrost, and to the direct and indirect effects of climate change in a geospatial context will be important in understanding and predicting the relationship between permafrost, thermokarst lakes, and the situated climate system (Schuur et al. 2007).

1.2. Purpose.

The overall goal of this study was to investigate trends in thermokarst lake changes in regions of continuous permafrost using new techniques for land cover classification. The study area chosen for this study was the Tuktoyaktuk Peninsula, described further in Chapter 3. The study was made possible by the declassification of the Landsat archive in 2008, which provided new

opportunities for a near-continuous analysis of land and lake cover trends. In this regard, an unsupervised automated technique on land cover classification was devised and implemented to process substantial amounts of Landsat satellite imagery to provide a multidecadal time series of lake coverage. The overall study goal can be separated into three component objectives:

1. To quantify changes in the direction and magnitude of thermokarst lake surface areal trends over a multidecadal period;
2. To determine the spatiotemporal distribution of lake surface areal trends at various scales; and
3. To investigate the influence of meteorological conditions in order to best explain the mapped changes in lake area.

Because thermokarst lakes change in size and number due to regional climatic fluctuations, the trends seen in thermokarst lake area and number may be an important signal of more widespread changes throughout the Arctic. The results of the project will therefore hopefully add to the body of work concerning the interspecific relationships between thermokarst lakes and their surrounding environment, and aid in developing methods in processing and classifying large amounts of satellite imagery.

2. BACKGROUND.

2.1. Properties and dynamics of permafrost.

2.1.1. Description of permafrost.

Permafrost, which is defined as the frozen ground that remains at or below 0 °C for two or more consecutive years (Zhang et al. 1999), forms a nearly continuous layer in the northern hemisphere, covering 25.5 million hectares, or 23% of the total land area (Brown et al. 1997). This specified area is traditionally divided into four zones based on the percentage of land area underlain by permafrost (Figure 2.1): the continuous permafrost zone (90 to 100%), the discontinuous permafrost zone (50 to 90%), the sporadic permafrost zone (10 to 50%), and the isolated permafrost zone (0 to 10%; Brown et al. 1997). The temperature, thickness, and geographic continuity of permafrost are controlled to a large extent by the surface energy balance, and therefore vary strongly with latitude (Schuur et al. 2008). Permafrost thickness spans a wide range, where the typical thickness ranges between 100 to 800 metres (m), 25 to 100 m, and 10 to 50 m in the continuous, discontinuous, and sporadic zones, respectively (Anisimov and Reneva 2006).

Overlying the permafrost is the active layer, the uppermost layer of soil that undergoes seasonal thawing each summer and freezing the following winter. The depth of the active layer ranges from a few centimetres in the sporadic permafrost zone to over a metre along the Arctic coastline and varies in thickness depending on the season (Figure 2.2; Brown et al. 2000). The permafrost table separates the active layer from the permafrost column, and delineates the upper boundary of perennially frozen soil depending on seasonality. Therefore, permafrost exhibits little to no annual temperature cycle (Anisimov et al. 1997). As the permafrost is essentially impermeable to water infiltration, the active layer is often saturated despite the often arid climate of the Arctic. On warming tundra terrain, the thawing of the various ground layers and deposits of underground ice cause the affected land surface to subside and collapse to form pits and hummocks, known as thermokarst (Davis 2003). Although the active layer freezes predominantly from the surface downwards, unfrozen water may still be present in the layer at any time during the year (Sturm et al. 2005). This process establishes major thermal and hydrological implications,

notably forming a liquid water phase system that provides a fast conduit for convective heat and water flow to and from the surface (Boike et al. 1998), a phenomenon apparent in the thermal regime of thermokarst lakes (explained in Section 2.2.2).

Figure 2.1. Spatial depiction of the latitudinal zonation of permafrost. From Brown et al. (1997).

Due to the prolonged cold of the Arctic winter, the active layer is further insulated by layers of drift snow largely trapped by tundra shrubs that blanket the tundra in significant amounts for two-thirds of the year (Sturm et al. 2005). Thicker snow produces higher winter soil temperatures and provides a source for soil moisture recharge during the spring when runoff from melt is highest (Sturm et al. 2001). The insulation provides a favourable thermal environment for microbes that aid plant survivability through the winter (Buckeridge and

Grogan 2008, Gouttevin et al. 2012). Because Arctic shrubs and other vegetation classes were found to exert control on both soil temperatures and the amount of water reaching the active layer (via evapotranspiration and the presence of a surface organic layer), vegetation dynamics may influence permafrost temperatures by altering the composition and thickness of the active layer (Runyan and O'Dorico 2012).

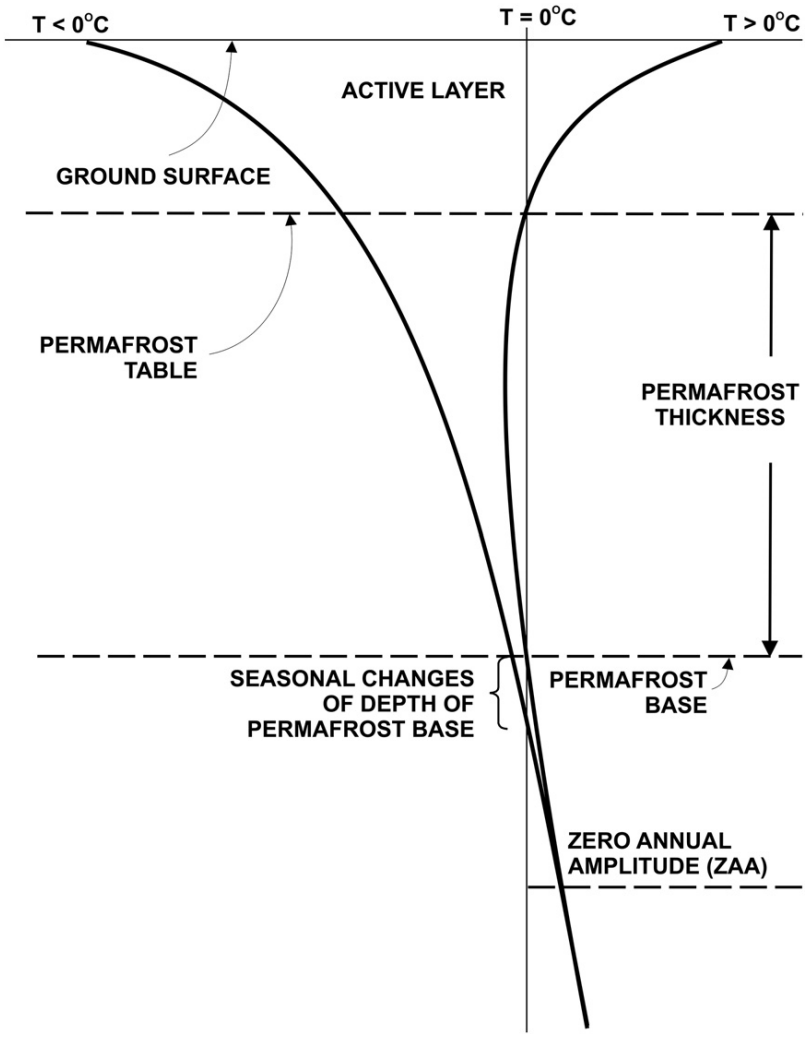


Figure 2.2. Thermophysical model of permafrost in a periglacial environment, relative to 0°C. From Dobinski (2011).

2.1.2. Permafrost and climate warming.

As permafrost is a thermal system with a delayed response to climate forcing due to the thermal and compositional properties of the frozen ground, the present state of permafrost is best described as a combination of former climatic conditions and present-day climate (Harris et al.

2009). Borehole measurements of ground temperature have produced a unidirectional trend of permafrost warming in the past several decades across the Arctic (ACIA 2005, Osterkamp and Jorgenson 2006, Osterkamp et al. 2009). The response of tundra permafrost to longer-term temperature change is dependent upon a multitude of variables, most importantly thermal conductivity, soil ice content, and present permafrost thickness (Dobinski 2011). The process of permafrost degradation is inherently complicated, because once initiated, permafrost may either persist in disequilibrium with the surrounding warmer climate, or degrade rapidly due to saturation along the perimeter of the permafrost table (Vitt et al. 2000).

Permafrost stability is a condition that is principally a product of climate change over various periods and magnitudes, causing permafrost to persist for up to thousands of years or to be newly formed during cold, snow-poor winters and persist only for several years (Jorgenson et al. 2010). Although air temperatures increased by as much as 20 °C during the Pleistocene-Holocene transition, as indicated by oxygen isotope analysis of Greenland ice cores (ACIA 2005), permafrost temperatures have warmed 2 – 4 °C from the Little Ice Age and around 3 °C in the last 20 years of the 20th century (Lachenbruch and Marshall 1986, Osterkamp and Romanovsky 1999, Jorgenson et al. 2010), with the highest warming trends located over central Siberia and the northwestern American continent (Chapman and Walsh 1993, Serreze et al. 2000). Because much of the permafrost around the world (especially zones of discontinuous and sporadic permafrost) is within a few degrees of thaw (e.g. Osterkamp 2007), recent Arctic temperature warming is expected to have a large impact on these periglacial areas.

It is important to note that the impact of climate change on permafrost is indirect, as permafrost is part of a complex geo-ecological system that also plays an important role in the dynamics of the pedosphere, the outermost soil layer of the Earth (Shur and Jorgenson 2007). The thermal regime of permafrost is mediated by factors such as topography, surface and ground water, soil properties, vegetation, and snow, all of which interact with each other to form positive and negative feedbacks to affect permafrost stability (Jorgenson et al. 2010).

2.2. Thermokarst basin and lake formation.

2.2.1. Thermokarst basin formation and development.

The thawing of permafrost often results in thermokarst terrain, which is characterized by irregular basins that forms pits and hummocks over the tundra landscape. Thermokarst basins range in size from 0.1 to 15 km in diameter and from 3 to 40 m in depth (Desyatkin et al. 2009), and often remain frozen for up to 9 months of the year with an ice thickness of up to 2 m (Hinkel et al. 2012). Typical features of thermokarst basins are a steep depression surrounding a flattened bottom, grassland vegetation around the perimeter, and often a central lake (French 2007, Katamura et al. 2007).

Thermokarst development involves several processes: 1. ice thaw and terrain subsidence; 2. ponding; 3. surface and subsurface drainage; 4. surface and subsurface subsidence and stabilization; and 5. associated surface and subsurface erosion (Yoshikawa and Hinzman 2003). As described in Toniolo et al. (2009), thermokarst subsidence occurs when the surface energy balance is altered by any of these above processes. Changes in the energy balance increases the heat flux downwards into the subsurface, initiating thaw of ice wedges and forming depressions in the terrain. The shorelines stabilise as the ground subsides, while the depressions are further shaped by erosion and sedimentation.

2.2.2. Thermokarst lake formation and dynamics.

Ponding of thermokarst depressions can occur when conditions are favourable for surface water retention. Most lakes formed in thermokarst basins are sufficiently shallow to have a permafrost layer beneath their littoral terraces (Burn 2002, 2003). Often, ice-rich permafrost as well as a high permafrost table acts as an aquiclude, retaining water at the surface and restricting the ability of soil thaw melt to percolate to deeper groundwater zones (Yoshikawa and Hinzman 2003). When the lake depth exceeds the maximal winter ice cover thickness, an unfrozen reservoir of water persists throughout the year, maintaining a net positive heat flow into the basal sediments (Smith 1976), and disrupts the thermal stability of near-surface permafrost (Hinkel et al. 2012).

At high latitudes, thermokarst lakes present the greatest thermal departure at the ground surface from systematic and spatially coincident climate patterns (Lachenbruch et al. 1962). The findings

by Burn (2005), who investigated the thermal regimes of tundra lakes on Richards Island in the Mackenzie Delta, Northwest Territories, clearly illustrate the extreme warming influence of lakes on surrounding permafrost ground by attributing the thermal phenomenon to within-lake heat exchange processes. Although the lakes are uniformly well-mixed and reach similar lake-bottom temperatures regardless of depth during the summer, a vertical temperature gradient exists during the winter that is dependent on the thickness of the lake ice. This thermal gradient increases with depth, where deeper water was found to be substantially warmer (at about 3 °C) than on shallow terraces within the same lake (at about -2 °C). Because the active layer in terrestrial permafrost exhibits a similar thermal regime to the seasonal maximum ice thickness in an lacustrine environment, a comparison of these two environments throughout their vertical profiles reveals three processes of lake influence on surrounding permafrost: 1. the ratio between winter and summer heat transfer coefficients; 2. storage of accumulated summer heat beneath lake snow cover; and 3. an asymmetry of lake-ice freezing and thawing geometries (Smith and Riseborough 1996, Riseborough 2006). Consequently, the large thermal effect of water bodies in permafrost regions may be attributed to the heat exchange processes associated with them (Riseborough 2006), and the presence of thermokarst lakes represents a local process of local climate warming.

Because thermokarst lakes are sensitive to local processes and regional climate variability, they may exhibit: 1. changes in the water balance and level of the lake through three-dimensional expansion; 2. increased groundwater outflow due to permafrost melt; and 3. erosion and melting of the shoreline by thermokarst processes (Marsh et al. 2008). Lake size is largely dependent on ground ice volume, regional relief, and surface age (Sellmann et al. 1975), and may grow or shrink in the timespan of between several years to decades. Lakes with an unstable permafrost base may be more susceptible to lateral permafrost erosion, which increases the surface area of the lake (Jorgenson and Osterkamp 2005). Sub-lacustrine permafrost may also degrade vertically through talik expansion, which can proceed until the lake bottom intersects the sub-permafrost groundwater systems.

Once the permafrost underneath lakes thaws, the interactions between surface water and groundwater drainage networks become largely controlled by the composition of the surrounding soil and the presence of topographical and hydraulic gradients (Grosse et al. 2013), resulting in either lake shrinkage through subsurface drainage or growth through groundwater recharge (Yoshikawa and Hinzman 2003). Coarse-grained sandy soil may promote lake drainage,

while fine-grained silty soil may act as an aquiclude similar to permafrost, promoting surface water ponding (Burn 2002, Roach et al. 2013). In addition, the presence of ice structures is often found in permafrost with silty soil, and the degradation of these structures can promote net increases in lake area, particularly in regions of continuous permafrost (Smith et al. 2005, Riordan et al. 2006). Sub-lacustrine permafrost degradation can lead to a reduction in lake area over annual and decadal periods by increasing water infiltration into the deepening active layer, reducing the amount of run-off available to recharge these lakes (Barber et al. 2000, Roach et al. 2013). In contrast to the above noted gradual declines in lake area, thermokarst lakes have also been noted to drain rapidly due to melting of an outlet channel into shallow superpermafrost groundwater systems (Marsh et al. 2008, Marsh et al. 2009). The mechanisms of lake water loss through these conduits have been explained by Mackay (1981, 1997), who investigated permafrost growth by artificially draining selected lakes around the Mackenzie Delta, Northwest Territories; after drainage, the resulting lake basins experienced a reduction in active layer depth from water loss and freeze-thaw consolidation, and was accompanied by frost heave and the growth of aggradational ice.

2.3. Spatial and temporal trends in high-latitude lake cover.

Thermokarst lake changes across the Arctic tundra region have been observed by experimental field methodologies at the local scale, and confirmed by remote sensing at the regional and the circumpolar scale. These observations are reported and summarized in Table 2.1.

A majority of studies have reported unidirectional decreases in lake number and surface area worldwide over the entire time series of modern satellite data; these study areas are typically located in regions of discontinuous and sporadic permafrost (Yoshikawa and Hinzman 2003, Smith et al. 2005, Riordan et al. 2006, Kirpotin et al. 2009, Sannel and Kuhry 2011, Roach et al. 2013). These changes have all been suggested as ramifications of a warming climate. Conversely, areas located in regions of continuous permafrost have experienced decreases in lake surface area only within the last three decades (Plug et al. 2008, Labrecque et al. 2009). These trends are furthermore considerably weaker than that of the more southerly areas, as other studies in continuous permafrost regions have found negligible or even increasing trends (Smith et al. 2005, Riordan et al. 2006). Therefore, areas situated within the zone of continuous permafrost are

instead better characterised by spatial heterogeneity (Duguay et al. 1999, Smith et al. 2005, Riordan et al. 2006), where lake increases have been observed along with lake decreases in the same area, highlighting a potential for resiliency at finer spatial scales (Roach et al. 2013).

Changes in the inter-annual variability in lake extent have been attributed to a range of potential atmospheric and climate-related drivers, all of which are related to a warming Arctic. A rise in Arctic surface temperatures is thought to be the primary source that alters the thermokarst lake equilibrium, affecting thermokarst processes and permafrost degradation (Yoshikawa and Hinzman 2003, Smith et al. 2005, Jepsen et al. 2013a), increased potential evapotranspiration (Riordan et al. 2006, Smol and Douglas 2007), atmospheric teleconnection patterns (Labrecque et al. 2009), increased terrestrialisation (Payette 2004, Roach et al. 2011), and/or hydrogeomorphic characteristics (Arp et al. 2011). Additionally, because the recharge of Arctic lakes is an important component in the lake water balance budget (Gibson and Edwards 2002), precipitation has a considerable effect on variations in the areal extent of thermokarst lakes (Plug et al. 2008, Tarasenko 2013). The strength of these relationships highlights the dynamic coupling between thermokarst lakes and its surrounding ecosystem.

Table 2.1. Recent reported studies of thermokarst lake changes in high-latitude ecosystems.

Region	Study	Site [permafrost zone]	Image source [count]	Time period	Observed changes
Alaska	Yoshikawa and Hinzman (2003)	Council, Seward Peninsula [Discontinuous]	Aerial photography, IKONOS	1950 – 2000	Thermokarst developed in areas with high alluvium deposits and ice-wedge polygonal terrain from 1981 to 2000. Thermokarst ponds over ice-wedge terrain decreased in area over the entirety of the last century through internal drainage in taliks.
	Riordan et al. (2006)	Arctic Coastal Plain [continuous]; Boreal forest regions [discontinuous]	Aerial photography [35], Landsat TM/ETM+ [14]	1951 – 2002	Shallow closed-basin ponds decreased in area (-27%) and number (-49%) in all boreal forest regions. Arctic Coastal Plain closed-basin ponds had little change in area.
	Jones et al. (2011)	Northern Seward Peninsula [continuous]	Aerial photography [25], IKONOS [2]	1950 – 2007	Lakes increased in number (+10.7%) but decreased in area (-14.9%). Large lakes (>40 ha) decreased in number (-24%) and area (-26%). Drainage of lakes was due to lateral breaching rather than subterranean infiltration.
	Rover et al. (2012) Jepsen et al. (2013b)	Yukon Flats, central Alaska [discontinuous]	Landsat MSS [3], TM [12], ETM+ [5]	1979 – 2009	3.4% of lakes decreased in area while 86% of lakes exhibited no change in extent. 59.9% of lakes decreased in area through the growing season. Shallow permafrost dynamics have a high degree of control over lake area through groundwater flow.
	Roach (2011) Roach et al. (2011) Nicol et al. (2013) Roach et al. (2013)	National Wildlife Refuges, central Alaska [discontinuous]	Aerial photography [82], Landsat TM/ETM [78]	1948 – 2009	Decreases in lake area in the Yukon Flats (-2.96%/year), Innoko NWF (-1.54%/year), and Kanuti NWR (-1.50%/year). Slight increases in lake sarea in the Yukon Flats West (0.31%/year). No significant changes in lake area in Tetlin and Yukon Flats East areas. Lakes with decreasing areas were most likely found in burned area, on coarser, well-drained soils, and farther from rivers. Spatial heterogeneity in lake areal trends were seen in all study areas.

Table 2.1 (continued).

Region	Study	Site [permafrost zone]	Image source [count]	Time period	Observed changes
Canada	Duguay et al. (1999)	Old Crow Flats, northern Yukon Territory [continuous]	Landsat TM	1986 – 1999	While there were both increases and decreases in individual lakes, there was an insignificant change in overall lake area.
	Plug et al. (2008)	Tuktoyaktuk Peninsula, Northwest Territories [continuous]	Landsat MSS [2], TM [6], ETM+ [6]	1978 – 2001	Increases in lake area occurred primarily between 1978 – 1992, and decreases between 1992 – 2001, and were dependent on total precipitation for the 12 months preceding the growing season.
	Labrecque et al. (2009)	Old Crow Basin, northern Yukon Territory [continuous]	Aerial photography [2], Landsat ETM+ [1]	1951 – 2001	General increases in lake area occurred primarily between 1951 to 1972, and decreases between 1972 – 2001. Overall decrease in lake area (-3.5%) occurred during entire study period. These decreases typically occurred in large lakes, whereas smaller lakes generally increased in area.
	Marsh et al. (2009)	Eastern Mackenzie Delta and Southern Tuktoyaktuk Peninsula, Northwest Territories [continuous]	Aerial photography	1950 – 2000	Highest rate of lake drainage occurred between 1950 – 1973 (~1 lake/year). The rate of drainage decreased over the following time periods 1973 – 1985 and 1985 – 2000 (to ~0.3 lakes/year), related to the effect of a warming climate.
	Carroll et al. (2011)	Canadian Arctic (50 °N - 70°N) [various]	MODIS	2000 – 2009	A total of ~0.8 million lakes were mapped to cover the study area. Small gains in lake area occurred in the southern regions were offset by larger losses in the North, which resulted in a net decrease in lake area during the entire time period (-6,700 km^2).

Table 2.1 (continued).

Region	Study	Site [permafrost zone]	Image source [count]	Time period	Observed changes
Fennoscandia	Sannel and Kuhry (2011)	Hudson Bay Lowlands, Canada; Tavvavuoma, Sweden; Rogovaya, European Russia [various]	Aerial photography (5), Quickbird (2), IKONOS (1)	1954 – 2003	High rate of lake drainage occurred in the peat plateau of Tavvavuoma (sporadic permafrost zone), along with infilling of fen vegetation. Low rates of lake drainage in the Hudson Bay Lowlands and Rogovaya (continuous and discontinuous permafrost zones).
Siberia	Smith et al. (2005)	Yamalo-Nenets Autonomous Okrug, Siberia [various]	Landsat MSS (38), MODIS	1973 – 2004	Decreases in large lake area (-11%) and total lake area (-6%) between 1972 – 1998. Increases in total lake area (+12%) and number (+4%) in areas of continuous permafrost, with decreases in total lake area (-9%, -5%, -6%) and number (-13%, -12%, -11%) in areas of discontinuous, sporadic, and isolated permafrost zones, respectively.
	Kirpotin et al. (2008) Kirpotin et al. (2009)	Yamalo-Nenets Autonomous Okrug, Siberia [various]	Landsat MSS, TM, ETM+, Resurs-F2, Spot-5, ERS-2, ALOS	1973 – 2007	Increases in lake area in areas of continuous permafrost (+10 – +12%), but decreases in lake area in areas of discontinuous permafrost (-21 – -29%).
	Kravtsova and Bystrova (2009) Kravtsova and Tarasenko (2010) Tarasenko (2013)	Central Yakutia, Siberia [continuous]	Landsat TM/ETM+	2000 – 2009	Decreases in lake area during summer periods (June to August). Increases in large lake area from 1976 – 2000 due to the water phase cycle from low to high during this period.

2.4. Remote Sensing of land cover and the advent of free satellite imagery.

2.4.1. History of land cover remote sensing.

A key component to large-scale global change research is the use of satellite imagery and remote sensing techniques to assess and monitor the state of the changing surface of the Earth. Remote sensing is an increasingly attractive option to depict and classify land cover, as it represents the Earth's surface in a consistent and continuous approach over large scales (Foody 2002). Importantly, remote sensing provides an efficient means of obtaining spatial and temporal trends in dangerous or inaccessible areas, such as the majority of the Arctic region.

Past studies that examined lake change through time have used a combination of historical aerial photography and modern satellite imagery, and often have determined change in lake surface area by a combination of manually delineating lake extent from these sources (Yoshikawa and Hinzman 2003, Riordan et al. 2006). Although manual lake delineation of satellite imagery generally produces accurate results, it is inherently subjective and labour intensive, and increasingly time consuming depending on the amount of data (Bolstad et al. 1990). Therefore, automated methods of terrain and lake cover classification are becoming increasingly favoured, given two primary pretexts: 1. the launch of several new satellite platforms within the turn of the new millennium; and 2. the recent declassification of Landsat satellite imagery by the United States Geological Survey (USGS). The latter pretext is important to the context of the study, and is discussed further in the following section (Section 2.4.2.)

While lakes and other water bodies may be obvious to the image interpreter, the process of land cover classification using automated methods is difficult because automated classification techniques do not possess the pattern recognition capabilities of the human brain (Hudak and Brockett 2004).

2.4.2. The Landsat Program

The Landsat Program, a series of Earth-observing satellite missions run jointly between the United States Geological Survey (USGS) and the National Aeronautics and Space Administration (NASA), has been running continuously since 1972 and provides the longest time series of remote sensing imagery data on record. With a spatial resolution of 30 metres and a 16-day

sampling frequency, Landsat observations provide an often ideal balance between coarse but high temporal resolution data for global monitoring and sub-metre observations for localized studies (Goward et al. 2001). As a result, these images are a crucial dataset for characterizing and detecting land cover and land use change at multiple spatial scales (e.g. Cohen and Goward 2004, Wulder et al. 2012, Goodwin et al. 2013).

Throughout the history of the Landsat Program, the high cost and limited access of image acquisition and processing (between USD 600 to USD 4000) have continually limited the scope and extent of geospatial research, particularly during the Earth Observation Satellite Company (EOSAT) privatisation period from 1984 to 1999. In 2008, NASA and the USGS implemented a new Landsat Data Distribution Policy that provides Level 1 corrected data for the entirety of the United States Landsat Archive held by the USGS and all future data through the Earth Resources Observation and Science (EROS) Center (Woodcock et al. 2008). The new open access policy has subsequently increased the distribution of the images from approximately 25,000 images in 2001 to 2.5 million images in 2010 (Wulder et al. 2012). Free access to Landsat imagery, coupled with technological advancements in computer memory capacity and processing speed have presented an opportunity for long-term time series applications and large area studies (Goodwin et al. 2013).

Previous studies involving land cover classification have shown that Landsat imagery with a spatial resolution at or below 30 metres can accurately classify a large variety of landscapes, from the tropics to the polar regions (e.g. Guerschman et al. 2003, Rozenstein and Karnieli 2011, Zhu et al. 2012). In addition, Landsat imagery has become increasingly used in time series studies investigating thermokarst lake changes (Table 2.1).

3. LOCATION AND STUDY AREA.

The study area comprises the Tuktoyaktuk Peninsula, which represents the wetland regions surrounding the Mackenzie Delta region on the northern coast of the Northwest Territories, Canada. The Tuktoyaktuk Peninsula is located above the Arctic Circle on the Arctic Coastal Plain directly east of the Mackenzie River, and forms part of the southeast coast of the Beaufort Sea (Figure 3.1). These geographical boundaries form the spatial extent of the study area, along with the Arctic treeline ecotone to the south. The Peninsula is around 40 km wide, and extends northeast for 200 km from the hamlet of Tuktoyaktuk to Cape Dalhousie at its northernmost point. The topography of the Tuktoyaktuk Peninsula is characterized by relatively low relief, of which the highest point does not extend past 205 m. The area is dominated by oriented thaw lakes and tundra vegetation, which include sedge, moss, and low-shrub tundra (>40 cm; Côté and Burn 2002, CAVM 2003). The Tuktoyaktuk Peninsula is located within the zone of continuous permafrost, with permafrost thicknesses ranging from less than 100 m to over 740 m (Judge 1986). Ground ice is abundant beneath the soil surface (Mackay 1971, Pollard and French 1980), and in terms of global permafrost, the amount of ground ice in the Tuktoyaktuk Coastlands appears to be excessively high (Murton 1996).

The Tuktoyaktuk Peninsula was selected as the area of study for 3 reasons: 1. the high percentage of thermokarst lakes covering the landscape; 2. a reasonable quantity of previous studies on the region; 3. the relatively high percentage of cloud-free Landsat imagery.

3.1. Thermokarst lakes in the Tuktoyaktuk Peninsula.

An important feature of the Tuktoyaktuk Peninsula is the large number of freshwater lakes, which occupy between 15 to 50 percent of any given landscape subsection. The lakes are frozen and ice-covered from mid-October until June, and reach a maximum ice thickness of between 1.5 to 2.0 metres during late April or May (Côté and Burn 2002). Although ice breakup in this region starts as early as late May, and only start to refreeze in late August (Pienitz et al. 1997), central ice pans may remain in the centre of the lakes until late June or July (Côté and Burn 2002).

Although there exists many different types of lakes in the area, most originate from the flooding of thermokarst basins (Rampton and Bouchard 1975). Thermokarst lakes in the Tuktoyaktuk Peninsula are prone to catastrophic damage, and many lakes have drained in the past, having partial former shorelines, misfit lakes, and/or residual pondage (Mackay 1979, 1988, Murton 1996).

Many of the lake basins on the Tuktoyaktuk Peninsula are characterized by shallow littoral shelves surrounding a deeper central pool (Burn 2005). Many of these lakes are oriented northwards as a result of easterly and westerly unimodal winds across the coastlands, driving perpendicular elongation through a two-cell circulation system (Côté and Burn 2002). Field measurements show these shelves to extend between 0.1 to 1 kilometres from shore proportional to the lake diameter, and the pools to be between 10 to 30 metres deep (Burn 2002). Shallow water sediments are mainly comprised of sand or gravel and/or organic material derived from previous bank collapse; the organic material, which includes peat, sod, and detrital mats, are highly erodible and are often reworked during summer storms (Murton 1996). In contrast, deeper water sediments are mainly covered by fine clastic sediment and organic material that have settled from suspension beneath winter ice (Murton 1996).

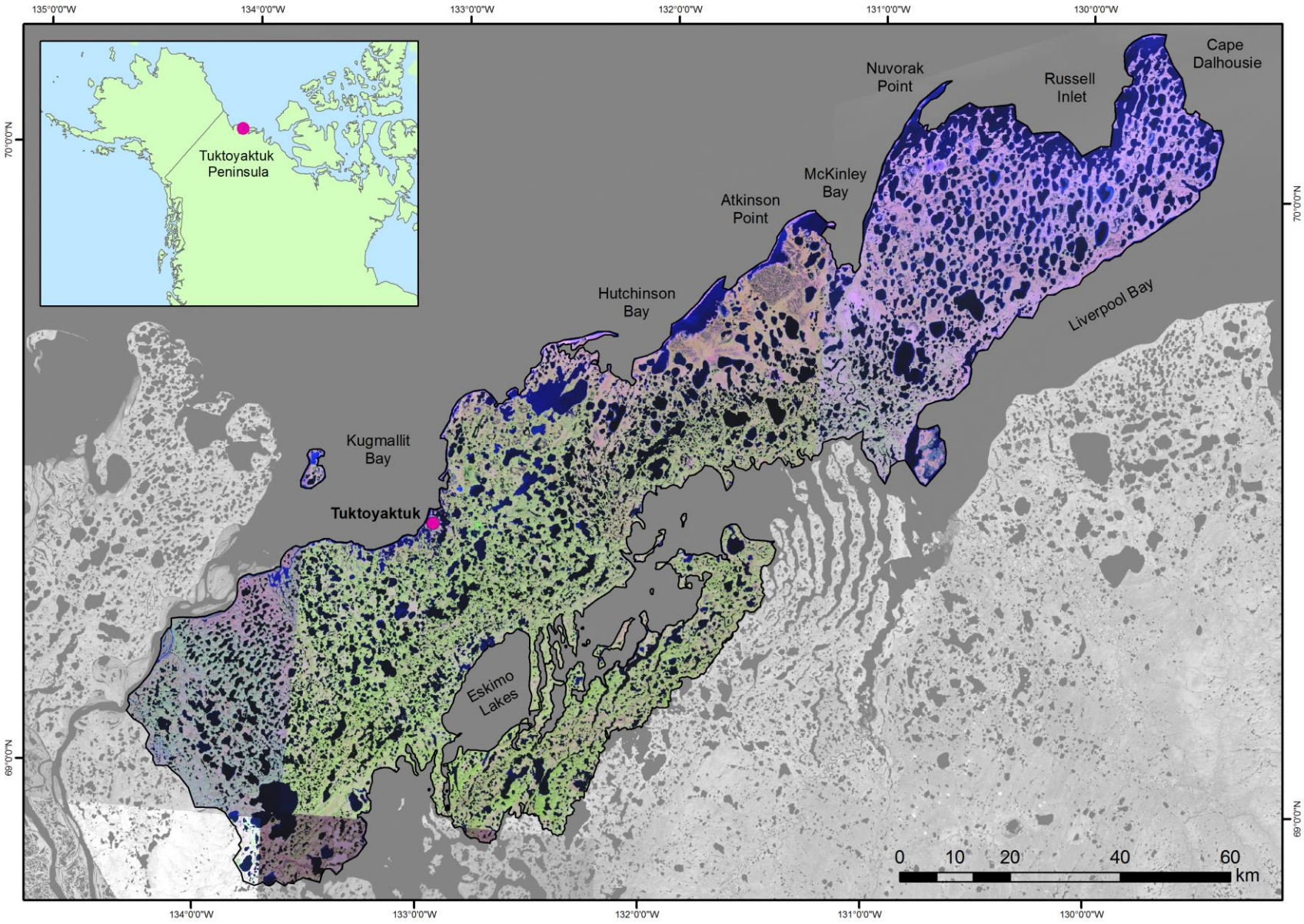


Figure 3.1. Landsat 5 TM image mosaic of the study area (band combination 5-4-3), which comprises the Tuktoyaktuk Peninsula, Northwest Territories, Canada. Notable landmarks, as well as the hamlet of Tuktoyaktuk, N. T., are spatially located on the map.

3.2. Climate of the Tuktoyaktuk Peninsula.

Based on the Köppen Climate Classification scheme (Kottek et al. 2006), the Tuktoyaktuk Peninsula has a continental sub-Arctic climate, characterized by short, cool summers and very long, cold winters. Historical climate averages (1971 – 2000) recorded at the Tuktoyaktuk weather station are summarized in Table 1 (Environment Canada 2013). While summer air temperatures hover up to 10 °C above freezing, mean annual air temperatures measure around -10 °C.

The hydrology of the lakes in Tuktoyaktuk Peninsula is largely controlled by a combination of low precipitation and permafrost melt, with snowmelt runoff during the spring being the largest influx of water during the year (Anema et al. 1990). The annual precipitation in Tuktoyaktuk averages 140 mm, of which between 35 to 50 percent is manifested as snow. Rainfall in the area is highest during late July and August, occurring mostly as sustained low-intensity storms (Plug et al. 2008). In the summer, evaporation from the lake surface (~230 mm) is generally greater than that of precipitation (~60 mm; Marsh and Bigras 1988, Bigras 1990).

Table 3.1. Climate averages for Tuktoyaktuk for the years 1979 to 2010. Temperature and precipitation statistics were obtained from the National Climate Data and Information Archive (Environment Canada 2013). Wind statistics were obtained from the Aurora Research Institute, Aurora College (Pinard 2011).

Climate	June	July	August	September	Yearly
Daily Average (°C)	5.9	10.9	9.2	3.0	-10.2
(Standard Deviation)	(1.7)	(2.1)	(1.8)	(2.0)	(2.0)
Daily Minimum (°C)	1.2	6.4	5.5	0.4	-6.6
Daily Maximum (°C)	10.7	15.3	12.8	5.6	-13.9
Rainfall (mm)	8.5	17.9	25.5	14.8	70.2
Snowfall (mm)	0.6	0.2	0.5	4.4	69.2
Average Wind Speed and Direction (km/h)	4.34 W	4.29 W	4.28 NW	4.40 W	4.28 W

4. METHODS.

The study was based on lake cover classification and analysis of 352 multi-spectral Landsat satellite images for the study area (Section 4.1). Systematic object-based classification of various land and water classes were applied for each pixel in all images using the Fmask algorithm (Version 2.2 Windows Executable, Section 4.2; Zhu and Woodcock 2012), and image objects were segmented and combined into yearly mosaics (Section 4.3). To account for intra-annual variability and seasonality in lake surface area, membership using fuzzy logic was then calculated for each pixel by applying three thresholds on a membership function to determine upper, middle, and lower bounds of lake rasters for each year within the temporal domain (1985 – 2011; Section 4.4). The surface area of lake bodies were extracted from the output lake image, and various statistics were calculated to determine trends at varying spatiotemporal scales (Section 4.5). Finally, these trends were correlated with various climatic variables to discern any possible relationships between areal trends and weather forcings (Section 4.6).

4.1. Landsat data and pre-processing.

Multi-spectral Landsat images were acquired from the United States Geological Survey (USGS) in Reston, Virginia for the study area, for dates between 1985 and 2011 (http://earthexplorer.usgs.gov). The images were derived from the Landsat 5 Thematic Mapper (TM) instrument for the years 1984 through 2011 and the Landsat 7 Enhanced Thematic Mapper Plus (ETM+) instrument for the years 1999 through 2013. Images acquired from the Landsat 7 ETM+ included those produced after 31 May 2003, which all contain an anomaly caused by a hardware component failure of the Scan Line Corrector (SLC). The images containing this data loss are hereafter referred to as SLC-off images. The SLC was used to compensate for the forward motion of the spacecraft to align all the scans in parallel orientation (Chander et al. 2009). Without a functional SLC, the satellite line of sight traces a zigzag pattern, producing wedge-shaped gaps on all subsequent images derived after the incident date varying in width from one pixel near the centre of the image to 14 pixels along the east and west edges of the image. The resulting missing data corresponds to a loss of approximately 22 percent of the target area for each image (Storey et al. 2005). Note that the failure of the SLC mirror assembly has no impact on the radiometric performance with the valid pixels.

Table 4.1. Landsat 5 TM and Landsat ETM+ spectral band properties.

Band number	Spectral response	Band width (λ, μm)	Spatial resolution (m)
Landsat 5 TM			
1	Blue	0.45 – 0.52	30
2	Green	0.52 – 0.60	30
3	Red	0.63 – 0.69	30
4	Near Infrared	0.76 – 0.90	30
5	Short-wave Infrared	1.55 – 1.75	30
6	Thermal (Long) Infrared	10.4 – 12.5	120
7	Mid Infrared	2.08 – 2.35	30
Landsat 7 ETM+			
1	Blue-Green	0.45 – 0.515	30
2	Green	0.525 – 0.605	30
3	Red	0.63 – 0.69	30
4	Near Infrared	0.75 – 0.90	30
5	Short-wave Infrared	1.55 – 1.75	30
6	Thermal Infrared	10.4 – 12.5	60
7	Mid Infrared	2.09 – 2.35	30
8	Panchromatic	0.52 – 0.90	15

All multispectral Landsat images covered a visual area of 185 km x 172 km, and have a spatial resolution of 30 m x 30 m for six different reflective bands (Bands 1 – 5 and 7; Table 4.1). In addition, the spatial resolution of Band 6 was improved from 120 m to 60 m, and a 15 m-panchromatic band (Band 8) was introduced with the launch of the Landsat 7 satellite.

A total of 352 Landsat-5 and Landsat-7 images were collected between the years 1985 to 2012 (Appendix 1). For the study area comprising the Tuktoyaktuk Peninsula, the images consisted of Paths 61 and 62 Rows 11 and 12, Path 63 Rows 10, 11, and 12, and Paths 64 and 65 Row 11. Images were selected based on availability, lack of obscuring cloud cover, and season of acquisition. Initially, all Landsat images that contained less than 40 percent of clouds covering the landscape were downloaded, and then images were individually screened based on lack of obscuring cloud cover over the study area. Only images that that lacked snow and substantial ice cover were used; temporally, these images spanned from mid-June to late September and represented the summer growing season.

All Landsat images were previously processed through the Level 1 Product Generation System (LPGS) with a Standard Terrain Correction (Level 1T). After image acquisition, all Landsat images were re-projected to a common Universal Transverse Mercator (UTM) Zone (WGS 1984 UTM Zone 9). Atmospheric correction was conducted as part of the Fmask algorithm for land cover classification (Section 4.2).

4.2. Land cover classification of pixels.

The procedure for determining lake presence required several stages of robust image segmentation and classification, as detailed below. These stages collectively identified image pixels that were and were not representative of lakes (terrain, cloud, cloud shadow, snow/ice, and other water bodies) and excluded them from further data processing downstream. Image classifications were conducted using both supervised and unsupervised techniques.

In order to spatially determine lake coverage over the scope of a Landsat image, a systematic object-based classification of various terrain and water classes must first be applied for each image pixel. From these cover types, only those that are not classified as cloud and cloud shadow should be applied for further data processing. Cloud and cloud shadow contamination of Landsat imagery has continually been a significant obstacle for spatial analyses of land cover (Ju and Roy 2008, Goodwin et al. 2013, Melaas et al. 2013). Optically thick clouds prevent panchromatic and infrared wavelength remote sensing of the covered surfaces and their associated shadows reduce the surface reflectance of its projected areas (Kaufman 1987). The subsequent effect of clouds and its shadows on surface illumination and reflectance observations in affected areas can bias calculations in areal classification and surface change (Zhu and Woodcock 2012).

In this light, the published Function of mask (Fmask) approach (Version 2.2, Windows Executable) was applied to all obtained Landsat imagery. Fmask is a newly-developed cloud and cloud shadow detection algorithm for the Landsat TM and ETM+ instruments that builds on the results of previous screening approaches. The Fmask algorithm is outlined in this section; a more detailed description can be found in Appendix 2 and a complete approach and evaluation can be found in Zhu and Woodcock (2012). Digital Number (DN) values were converted to top-of-atmosphere (TOA) reflectances and brightness temperatures (BT; °C) using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) atmospheric correction

procedure (Masek et al. 2006, Vermote and Saleous 2007). Then, various spectral tests were conducted to identify and extract the land cover classes of terrain, water, cloud, cloud shadow, and snow/ice. The algorithm can be summarized into four distinct "passes":

Pass 1: Identification of potential cloud pixels. A series of 4 spectral tests were applied to input images to flag pixels that may represent clouds, termed "Potential Cloud Pixels (PCPs)". The spectral tests applied thresholds that identified cloud pixels based on the temperature, reflectivity, and whiteness characteristics normally exhibited by clouds. As well as identifying thick cumulus clouds, Pass 1 accounted for cirrus clouds and haze. Pass 1 also attempted to discriminate clouds from other terrain bodies that exhibit similar spectral characteristics, such as rocks, turbid waters, bare soil, and sand. Pixels not included in the first pass are considered as confident in representing clear-sky pixels. Pixels were labelled as PCPs if they satisfied all of the spectral tests used within this pass.

Pass 2: Extraction of a potential cloud layer. Using absolutely clear-sky pixels (pixels not identified as PCPs in Pass 1), a potential cloud layer was extracted by computing potential cloud probability for all pixels in input images. Cloud probability was calculated separately for PCPs over terrain and water due to variations in temperature and reflectance distributions over each input image. This required discrimination between terrain and water pixel classes, which was calculated by thresholding the Normalised Difference Vegetation Index (NDVI) and the top-of-atmosphere (TOA) reflectance values in the upper visible spectrum. By combining the results from Passes 1 and 2 (PCPs and potential cloud probability), a potential cloud layer was generated. To err on cloud pixel commission rather than omission, the cloud mask was dilated by one pixel depending on the results of its neighbouring pixels.

Pass 3: Identification and extraction of a potential cloud shadow layer. Due to the darkening effect of cloud shadows in the near-infrared (NIR) spectrum, a potential shadow layer was generated by applying a flood-fill transformation to the NIR band. The potential cloud layer extracted in Pass 2 was transformed into three dimensions, and the resulting cloud objects were matched with their respective shadow projections by a series of geometric and trigonometric calculations relating the spatial location of the cloud morphometrics, the respective cloud shadow location, and the sun and satellite positions. Because the resulting cloud shadow mask may have contained spatial gaps, the shadow

pixels were dilated over a 3-by-3 pixel neighbourhood in 8-connected directions to fill these gaps.

Pass 4: Identification and extraction of a potential snow and ice layer. The spectral tests used to detect snow and ice-contaminated pixels were derived from the MODIS snow mapping algorithm (Hall et al. 2001). The pixels identified as representing snow and ice bodies formed the potential snow and ice layer.

In addition, terrain and water cover types were identified in Pass 2, and were designated for each pixel after identification of all other cover types. Because more than one class may have existed within any one pixel, the algorithm assigned a class hierarchy in the following order from highest to lowest: clouds, cloud shadows, snow and ice, terrain, and water (Figure 4.1).

Each output image was visually inspected to identify misclassified pixels, which once identified, were masked from further image processing. The misclassification of haze, shaded cloud, or cloud-contaminated shadows had been particularly problematic when classifying water bodies using Landsat imagery (Zhu and Woodcock 2012). Clouds and their shadows that had been misclassified as terrain or water were subsequently masked and considered as missing data. If the image contained misclassified data over large areal extents, the image was discarded and not used for further image processing.

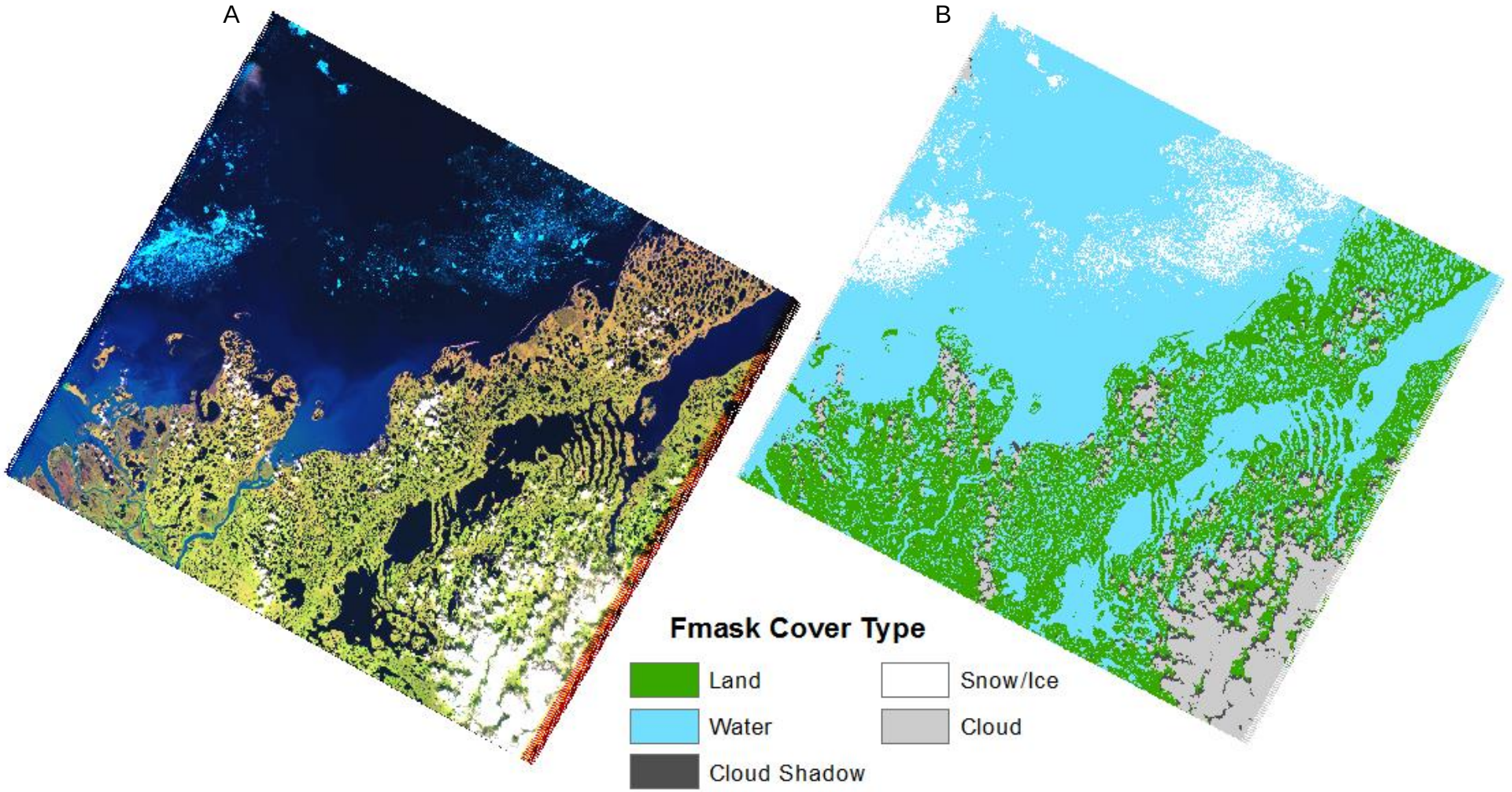


Figure 4.1. Example Fmask result for Landsat scene (Path 63 Row 11, 4 July 2011). **(A)** Landsat 5 TM scene (band combination 5-4-3). **(B)** Fmask output of the example Landsat scene.

4.3. Segmentation of land cover classed images.

After running the Fmask algorithm for each Landsat image, a simple algorithm was established to hierarchically rank lake detections to reduce false positives. This algorithm was implemented using Python 2.7 in the ArcPy module from ArcGIS (Version 10.0, ESRI, Redlands, CA), and is detailed in Appendix 2.1. In this study, pixels were grouped into two broad classes: water and terrain. In general, a pixel was designated as representative of lakes and was assigned a value of 1 if it satisfied all of the following conditions:

1. The pixel was flagged as water.
2. The pixel was not flagged as terrain.
3. The pixel was not flagged as cloud.
4. The pixel was not flagged as cloud shadow.
5. The pixel does not lie in areas of missing data.

To distinguish lakes from other water bodies such as rivers, seas, or wetlands, a river mask was obtained from Natural Earth public domain map dataset at the 1:10m scale (NACIS 2013; http://www.naturalearthdata.com). Pixels that satisfied conditions 1 through 5, but were located in these regions of non-lacustrine water bodies were internally flagged and excluded from further data processing. All lakes that were previously misconstrued as rivers did not meet the requirements of this condition, and were therefore excluded from further data processing. Similarly, wetland areas and frequently-inundated coastal terrain that have displayed connectivity to the ocean in at least one or more scenes were discarded from further analyses.

As areas in northern Canada are characterized by below-freezing temperatures for the majority of the year, snow and ice are often present during the early summer months. The presence of snow and/or ice, particularly on the periphery of lake shorelines, may potentially bias lake areal data by increasing individual lake size. The study accounted for this limitation by restricting the intra-annual period to the summer growing season (June to September). However, if snow and/or ice were present in any input Landsat image, affected pixels were flagged for further review if:

6. The pixel satisfied conditions 2 through 5.
7. The pixel was flagged as snow or ice.

By manually examining a number of lakes with ice and snow cover at varying concentrations, it was found that a large majority of pixels flagged as snow or ice were coincident with known areas of water ($p < 0.05$). Therefore, if the concerned lake pixel was flagged as snow or ice, it was included in the final lake assignment and assigned a value of 0.5.

Otherwise, a pixel was designated as representative of terrain and assigned a value of -1 if it satisfied the following conditions:

8. The pixel satisfied conditions 3 through 5.
9. The pixel was not flagged as snow or ice.
10. The pixel was flagged as terrain.

After applying the ranking algorithm to each Fmask-classified image, each image was reduced to containing only terrain, water, and snow/ice pixels. The range of values in each image was consequently reduced to -1, 0.5, and 1 respectively, to describe the range of land cover classes as an ordinal set. The resulting objects classified as lakes are believed to be a rather conservative but accurate representation of lake cover within the image.

To determine inter-annual changes in lake dynamics, and to increase the confidence of lake presence, Landsat images were combined into yearly raster mosaics, as applied to each pixel by the equation:

$$\frac{\sum_{i=1}^{N} 100 x_i}{N}, \qquad \text{[Eq. 4.1.]}$$

where N represents the total number of times a specific pixel is present within a given year, and x_i is the pixel measurement vector in the domain $1 \leq i \leq N$ (Appendix 2.1). The resulting mosaics create a spectrum where positive numbers represent increasing confidence in lake pixel presence, negative numbers represent decreasing confidence in lake pixel presence, and 0 represents the crossover point (Figure 4.2). Pixels were normalized to represent a percentage of confidence where a value of 100 represents a pixel uniformly being flagged as water in all instances, a value of -100 representing a pixel uniformly being flagged as terrain in all instances, and a value of 0

representing the crossover point between terrain and water. For instance, a pixel that was represented as water 6 times in a total of 6 images will have a value of 100, a pixel that was represented as water 2 times and as snow 2 times in a total of 4 images will have a value of 75, a pixel that was represented as water 1 time, as snow 2 times, and as terrain 5 times in a total of 8 images will have a value of -30, and a pixel that was represented as terrain 1 time in a total of 1 image will have a value of -100.

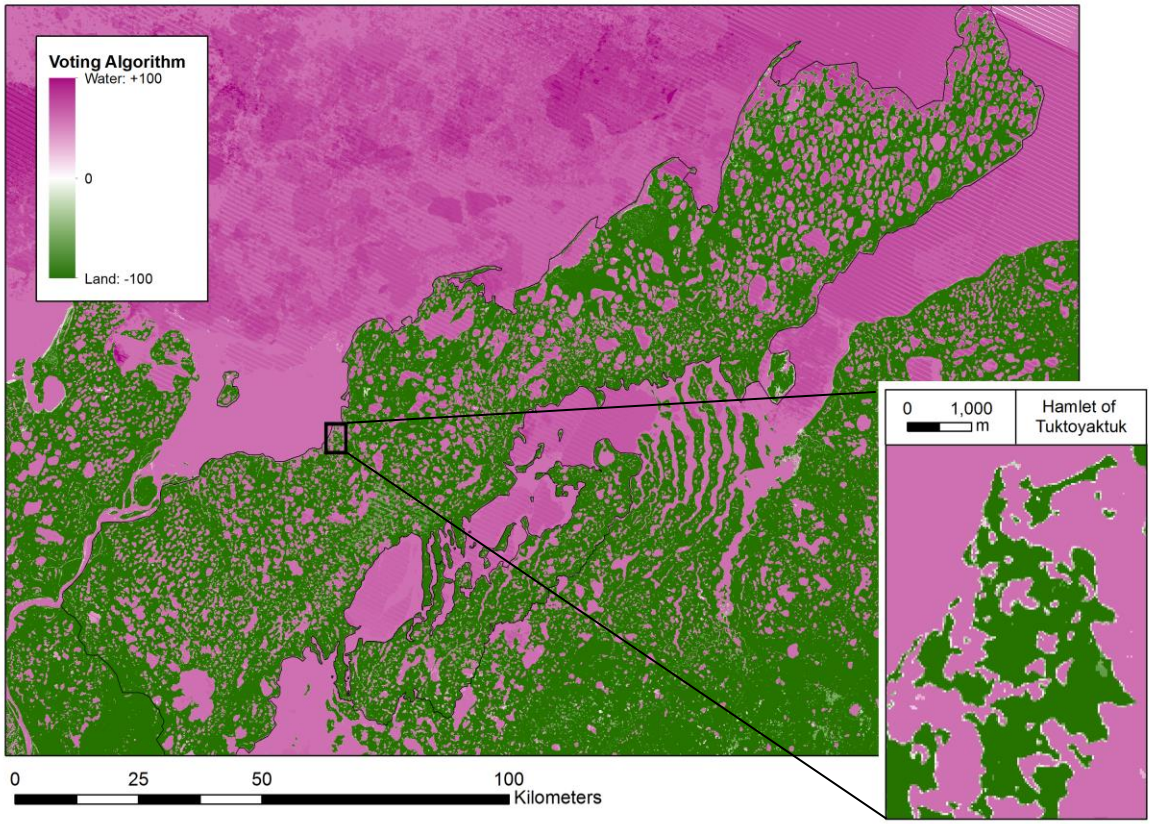


Figure 4.2. Example of output yearly raster mosaic composite for the year 2011. Pixels representative of water were mapped to the magenta channel, and pixels representative of terrain were mapped to the green channel. Where more votes were recorded, the hues of each pixel were modulated by increased components. Water bodies that were not representative of lacustrine environments (e.g. seas, rivers) were internally flagged and excluded from further data processing. Diagonal striping patterns illustrate the effects of use of Landsat 7 ETM+ SLC-off mode imagery with missing wedge-shaped data stripes. The inset image of the hamlet of Tuktoyaktuk, N.T., portrays the gradient between water and terrain due to the influence of seasonal change on local water levels. This variability was accounted for using a fuzzy membership algorithm (detailed in Section 4.4.).

4.4. Uncertainty of pixel-classified images.

When classifying an image, generally two kinds of problems are faced: 1. the spatial boundary between two land cover classes is not defined by a Boolean domain, but instead represents a gradient between the classes; and 2. a single pixel may contain more than one type of land cover class (Sharma et al. 2011). In the study, a third problem arose from the combination of images into yearly mosaics: 3. the loss of seasonal variability. These problems are primarily due to a limitation in available spectral resolution, and ambiguity induced by point-spread functions at the sub-pixel level (Townshend et al. 2000, Foody and Atkinson 2002). Figure 4.3 exemplifies how there is no definite boundary between terrain and lake as a result of these problems.

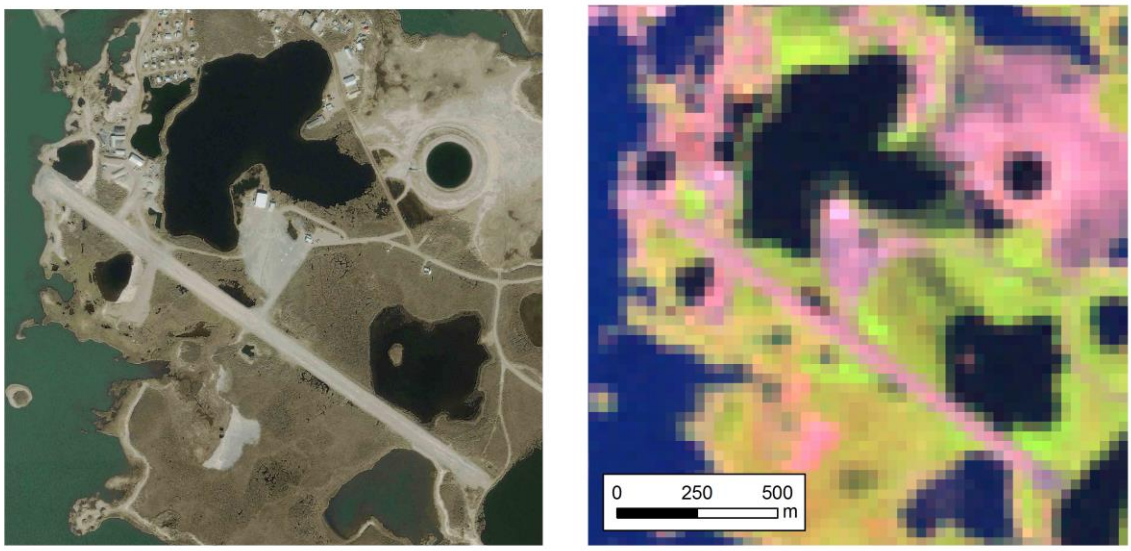


Figure 4.3. A spatial comparison of **(A)** vertical aerial photography and **(B)** Landsat imagery (band combination 5-4-3) of the hamlet of Tuktoyaktuk, N.T. Because Landsat imagery has a resolution of 30 m by 30 m, the boundary between terrain and lake is represented by a gradient instead of a strict delineation. Aerial photography was obtained from the Northwest Territories Department of Transportation (http://www.dot.gov.nt.ca).

A fuzzy classification model (McBratney and Moore 1985, Wang 1990, Moon 1998) was implemented to account for such vagueness in both space and attributes (Williamson 1994, Fisher 2000, Robinson 2003). Fuzzy sets, originally proposed by Zadeh (1965), are widely used in GIS and are increasingly applied to identifying and monitoring water bodies (e.g. Yin et al. 1999, Carter et al. 2007, Fisher 2010, Sharma et al. 2011).

4.4.1. Description of fuzzy membership

In a fuzzy classification model, all objects within the fuzzy set are characterized by a membership function $f_A(x)$ which associates with the object a real number in the range [0, 1]. In the set A, a value of $f_A(x) = 0$ represents an absence of the object in the set, and a value of $f_A(x) = 1$ represents full confidence of the object in the set. Values between $f_A(x) = 0$ and 1 indicate the relative strength of the object containing properties relevant to the set in a continuum. In general, a value of 0.5 suggests that the object has equal probability of inclusion into or exclusion from the set. In a geospatial context, fuzzy representation is partitioned as a spectral space composed of a family of fuzzy sets F_1, F_2, …, F_l on the domain X such that for all values x which is an element of X (Wang 1990):

$$\begin{aligned} &0 \le f_{F_i(x)} \le 1 \\ &\sum_{x \in X} f_{F_i(x)} > 0, \\ &\sum_{i=1}^{m} f_{F_i(x)} = 1 \end{aligned} \qquad \text{[Eq. 4.2.]}$$

recorded in a fuzzy partition matrix:

$$\begin{bmatrix} f_{F_1(x_1)} & f_{F_1(x_2)} & \cdots & f_{F_1(x_n)} \\ f_{F_2(x_1)} & f_{F_2(x_2)} & \cdots & f_{F_2(x_n)} \\ \vdots & \vdots & \ddots & \vdots \\ f_{F_n(x_1)} & f_{F_n(x_2)} & \cdots & f_{F_n(x_n)} \end{bmatrix}, \qquad \text{[Eq. 4.3.]}$$

where F_i represents the spectral classes, X represents the domain of all pixels in the dataset, m is the number of predefined classes, n is the number of pixels, x_i is a the pixel measurement vector in the domain $1 \le i \le m$, and f_{F_i} is the membership function of the fuzzy set $F_i(1 \le i \le m)$.

As stated in Wang (1990), a fuzzy partition of a spectral space allows for an abundance of spectral information to be incorporated for analysis. Membership values are used to describe the gradient of intermediate and multiple cover classes, and classifies stray pixels and pixels isolated between classes.

4.4.2. Application of fuzzy membership.

Although the fuzzy set has a range of memberships varying from 0 to 1, applying a threshold of any given value allows for the conversion to a Boolean set, yielding an alpha-cut (α-cut). α-cuts are a method to describe a vague interpretation of a landscape (the fuzzy set) as a hard or crisp composition (the Boolean set), but doing so results in the loss of much spectral information (Fisher 2010). To alleviate this concern, several α-cuts were selected to represent the fuzzy set as an ordinal series of Boolean sets. For practical purposes, the study limits the series of α-cuts to 0.3, 0.5, and 0.7 to represent the lower, middle, and upper bounds, respectively. These values were previously used in Bijker et al. (2011) to map the spatial distribution and uncertainty of Lake Naivasha, in Kenya. The value range of yearly mosaics of Landsat images (Section 4.3) were transformed to a range of [0, 10], where 5 represents the midpoint between terrain and lake representation, and a simple right sigmoidal membership function was applied to input images, given by the expression (Tsoukalas and Uhrig 1997):

$$f(x,c) = \frac{1}{1+\left(\frac{x}{c}\right)^{a}}, \qquad \text{[Eq. 4.4.]}$$

where *a*, the spread, is a standardized factor that controls the slope at the crossover point $x = c$. The resulting curve describes a fuzzy membership function where the larger input values exhibit membership closer to $f_A(x) = 1$. The study applied a fixed crossover value of $c = 5$ and a fixed spread value of $a = 5$ (Appendix 3.2), implemented using Python 2.7 in the ArcPy module from ArcGIS 10.0.

4.5. Statistical analyses of spatiotemporal trends.

To perform spatio-temporal analyses that examined both individual and overall lake variability over different time periods, yearly lake raster mosaics were converted to polygon features, and each lake was assigned a unique ID number to track individual lake change. To summarise lake polygons that may have coalesced or separated through time, lake area was defined as either the polygon or group of polygons within the combined maximum extent of that individual water body throughout the entire time series, following the methods of Roach et al. (2013). Lakes that were smaller than 1,350 m^2 (the midpoint between 1 and 2 pixels) in all images were excluded

from further data processing in order to reduce omission and commission errors (Roach et al. 2012). If a specific lake polygon was smaller than 1,350 m^2 in at least one of the input mosaicked images, the lake was removed from analysis.

The analysis focused on two major components: first, the net change in lake surface area and count was calculated over the entire study area regardless of regional specifications; and second, a series of spatial clustering statistical tests were applied to trends in individual lake morphometrics to identify and examine any patterns in spatial heterogeneity over the study area.

4.5.1. Overall variability in lake morphometrics.

Lakes were designated into five area classes of increasing lake size (Table 4.2). The classification scheme models a geometric distribution to normalise the left-skewed distribution in lake count. To test for structural breaks in lake area and count time series, a Chow test was implemented on linear regressions of ice cover in each regional time series to test for a structural break in correlation. Specifically, the Chow test is a series of F-tests to determine if regression coefficients are structurally different between the two subsets (Chow 1960).

Table 4.2. Specifications for lake surface area classes. Note that lakes representing a single (1) pixel were not included in analysis due to high degrees of sub-pixel uncertainty.

Area class	Area range (m^2)	Pixel range
1	901 – 9,000	2 – 10
2	9,001 – 90,000	11 – 100
3	90,001 – 900,000	101 – 1,000
4	900,001 – 9,000,000	1,001 – 10,000
5	9,000,000 +	10,000 +

4.5.2. Spatial heterogeneity in lake morphometrics.

General linear models (GLMs) were conducted for each identified lake body to estimate trends in individual lake area over time using the R Project for Statistical Computing (Version 2.13.0, R Development Core Team 2006). As all time series under different fuzzy membership α-cut thresholds reported similar trends (e.g. Figure 5.1; see Sections 6.1 and 6.2), only lakes determined using the α-cut threshold of 0.5 were investigated. The predictor (independent)

variable used was time (in years), and the response (dependent) variable used was the surface area for each individual lake body (in m^2). The regression estimate (slope) represented the rate at which each water body increases or decreases in surface area over time.

To examine spatial autocorrelation over the entire study area, the Global Moran's *I* statistic (Moran 1950) was applied to the set of lake polygons which exhibited a significant ($p < 0.05$) trend in areal increases or decreases. Given a set of features, the Moran's *I* evaluates whether the spatial pattern expressed was clustered, dispersed, or randomly distributed across the study area. The Moran's *I* statistic for spatial autocorrelation is given as follows:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j} z_i z_j}{\sum_{j=1}^{n} z_j^2}, \qquad \text{[Eq. 4.5.]}$$

where z_j is the deviation of an attribute for the feature j from its mean $\left(x_i - \bar{X}\right)$, N is the number of spatial units indexed by i and j, and $w_{i,j}$ is an element of a matrix of spatial weights between features i and j. The Moran's *I* tests the null statistic that trend classes in lakes were randomly distributed without any noticeable spatial pattern. Calculation of the Moran's *I* statistic was conducted in ArcGIS 10.0, using an inverse-weighted function that conceptualizes spatial relationships by distance from a target feature with a fixed distance band. The band distance was determined by first calculating the Moran's *I* statistic using a starting band distance of 5000 m, then subsequently increasing the distance by 1000 m increments until an optimal minimum in *p*-value was achieved.

To examine spatial autocorrelation at more local scales, the Getis-Ord's Local G_i^* statistic (Getis and Ord 1992) was applied to the same lake polygons that exhibited significant areal trends. The Local G_i^* identifies statistically significant hot and cold spots (cluster of high and low values, respectively) in the spatial distribution of lake trends. The Getis-Ord's Local G_i^* statistic is given as follows:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j - \bar{X}\sum_{j=1}^{n} w_{i,j}}{\sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\bar{X}\right)^2}\sqrt{\frac{n\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2}{n-1}}}, \qquad \text{[Eq. 4.6.]}$$

using the same variables as the Moran's *I* statistic. The local G_i^* statistic is a Z-score, which is used to quantitatively determine the confidence with which the clusters have been identified. Calculation of the Getis-Ord's G_i^* was similarly conducted in ArcGIS using a variable distance band, which optimises the cut-off distance between target features to ensure inclusion of at least one neighbour for all target features. A general identification of hot and cold spots is potentially more useful than reducing the scope of the test to favour only the highest or lowest values; therefore, input slope values were log-transformed to reduce the weight of the largest extremes. The Getis-Ord's G_i^* can be applied to transformed values because it is asymptotically normal (Getis and Ord 1992). As the local G_i^* statistic measures the magnitude of clustering, it should only be run if other global tests also predict significant clustering patterns (Carroll et al. 2011).

4.6. Correlation to weather variables.

Monthly climate data for weather stations across Canada are available for free acquisition from the National Climate Data and Information Archive in Fredericton, New Brunswick (Environment Canada 2013). Unfortunately, the Northwest Territories has few stations with a long-term weather record, and the closest weather station, located at the Tuktoyaktuk/James Gruben Airport (IATA: YUB) in Tuktoyaktuk, NWT (69° 26' 00" N, 133° 01' 35" W), is entirely missing data observations for the period spanning from January 1994 to September 1999. As a result, this study instead used ERA-Interim Reanalysis data (Dee et al. 2011) to examine atmospheric forcing on the study area.

2-metre ground temperature and total precipitation data were obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) in Reading, United Kingdom. The data were derived from synoptic monthly means of the ERA Interim global atmospheric reanalysis project for the period 1985 – 2011 (Uppala et al. 2005). The dataset was produced using a sequential data assimilation scheme, advancing in time by 12-hourly analysis cycles that combine available observations with prior information from a forecast model to estimate the changes in

the global earth, ocean, and atmosphere (Dee et al. 2011). The data used in this study has a grid cell resolution set at 0.5° x 0.5°. Monthly data from June to September were averaged for each year to produce yearly summer means.

For each climatic dataset, GLMs were conducted to examine trends and mean conditions that may influence overall lake areal change patterns. The predictor (independent) variable used was the specific climatic variable, and the response (dependent) variable used was the surface area for each individual lake body. The regression estimate (slope) represented the relative influence of each climatic variable on lake areal trends.

GLMs were also conducted for each individual lake body to estimate trends in individual lake area over time. Only lakes that showed a significant increasing or decreasing trend in surface area, as determined in Section 4.5.2, were regressed. To test for spatial autocorrelation, the Global Moran's *I* statistic was furthermore applied to the set of lake polygons which exhibited a significant ($p < 0.05$) trend both in areal increases or decreases over time (as identified in Section 4.5.2) and in areal increases or decreases over each climatic variable.

5. VALIDATION OF METHODOLOGY.

In order to ensure robust results, a quantitative validation of the methodology was first conducted. Specifically, two aspects of the methodology were validated: the accuracy of the Fmask algorithm in detecting land cover classes of terrain, water, and snow/ice (Section 5.1); and the correlation of weather variables between ground observations and ERA-Interim Reanalysis data (Section 5.2).

5.1. Validation of the Fmask algorithm.

5.1.1. Methods for the Fmask algorithm validation.

Because maps are deliberate generalisations of reality, and that all transformations of these maps introduce error (Maling 1989), quality assessment of remote-sensing image classification systems for land cover mapping is essential to evaluate the product accuracy (Smits et al. 1999). In this light, virtual "ground truthing" was conducted on images processed through the Fmask algorithm following a procedure modified from Zhu and Woodcock (2012) and Goodwin et al. (2013). First, two images from each year were randomly selected using a random number generator, with the condition that all WRS Path/Row combinations for all years are equally represented. This corresponded to a selection of 48 from a total of 352 scenes. A sample of between 60 and 90 pixels (depending on the amount of cloud cover projected on the image) was generated from each input image using a random generator and stored as a set of point shapefiles. These points were spatially constrained to overlay only areas classified by Fmask as pixels with data. Each pixel was visually interpreted and recorded whether the pixel contained: terrain, water, snow/ice, or 'ambiguous'. Because there does not exist enough aerial photography imagery for each Landsat imagery assessed, the accuracy of image classifications were directly assessed by visually interpreting each Landsat image and identifying whether or not a certain pixel was correctly classified. The accuracy of image classifications was assessed overall regardless of class type:

$$Overall\ Accuracy = \frac{Agreement}{Total\ Pixels}, \qquad \text{[Eq. 5.1.]}$$

and separately for each class using the producer's and user's accuracy:

$$Producer's\ Accuracy = \frac{Agreement\ for\ class}{Agreement\ for\ class + omission\ of\ class} \text{ and} \quad \text{[Eq. 5.2.]}$$

$$User's\ Accuracy = \frac{Agreement\ for\ class}{Agreement\ for\ class + comission\ of\ class}, \quad \text{[Eq. 5.3.]}$$

where an agreement represents the same classification between visual interpretation and the Fmask algorithm. Because thorough assessments were conducted for cloud cover and cloud shadow cover accuracy in previous studies (Zhu and Woodcock 2012, Goodwin et al. 2013), the validation assessment conducted for this study examined only points coinciding with terrain, water, and snow/ice classes.

5.1.2. Accuracy assessment of the Fmask algorithm.

At the pixel scale, the overall accuracy of the Fmask algorithm on terrain, water, and snow/ice cover classes is high with a score of 97.2%. When partitioned into separate classes, accuracy assessments revealed high percentages for all land cover classes, ranging from 89.4% to 99.4% representing both user's and producer's accuracies.

Because the intra-annual study period was constrained to the growing season (June to September), the large majority of Landsat scenes used for the study were largely snow and ice free. Therefore, the number of input points was comparatively lower than those used for other land cover classes. Regardless, results for snow and ice classification were determined to be the least accurate with similar producer's and user's accuracies; upon further visual inspection, nearly all of the points that were misclassified to and from snow and ice were found to reside in areas comprising sparse snow cover and scattered ice floes. Because pixels classified as snow and/or ice were assigned a value representing half of water, the process of image segmentation, where additional images representing a snow- and ice-free landscape has been designed to reduce the impact of this discrepancy.

Table 5.1. Accuracy of the Fmask algorithm on land cover classification, by land cover class. The overall accuracy of the Fmask algorithm for all cover classes is 0.972, with an input of 3112 random points. The number of total points input for each land cover class represents the total amount of points assigned to the specified land cover class by the Fmask algorithm. The producer's and user's accuracies were defined in Equations 4.2 and 4.3, and are decimal representations of percentage. Accuracy assessments for cloud and cloud shadow (indicated by sharps) were obtained from Zhu and Woodcock (2012).

Land cover class	Total points	Producer's accuracy	User's accuracy
Terrain	1708	0.960	0.994
Water	1246	0.996	0.952
Snow/Ice	158	0.911	0.993
Cloud#	142 scenes	0.921	0.894
Cloud Shadow#	142 scenes	>0.7	~0.5

A spatial examination of the accuracy assessments for terrain and water pixels revealed that almost all misclassified pixels resided either in transition zones between the two land cover classes or were coincident with small streams, where both entities were between 1 to 2 pixels wide (30 to 60 metres; Figure 5.1).This tendency was true regardless of the fuzzy membership α-cut threshold used, which indicates high, average, and low bounds of lake extents. When faced with such ambiguous situations, the Fmask algorithm was found to assign coincident pixels to terrain approximately 80% of the time, and to water approximately 20% of the time (Table 5.2). The preference of terrain to water by the Fmask algorithm implies that lake surface area projections in this study were conservative in estimates. However, because the exact same algorithm and methods were applied to all images for all years, the outputs of lake surface area was believed to be an unambiguous representation of lakes.

Table 5.2. Preference of Fmask algorithm towards terrain and water cover classes when presented with ambiguous situations. Two ambiguous locations were identified: the terrain-water transition zone and small streams. The number of total points input for each situation represents the total amount of points coincident with the locations of the specified situation. The preference ratio are decimal representations of percentages.

Land cover class	Total points	Terrain points	Terrain preference ratio	Water points	Water preference ratio
Transition Zone	167	134	0.802	33	0.198
Small Streams	23	19	0.826	4	0.174

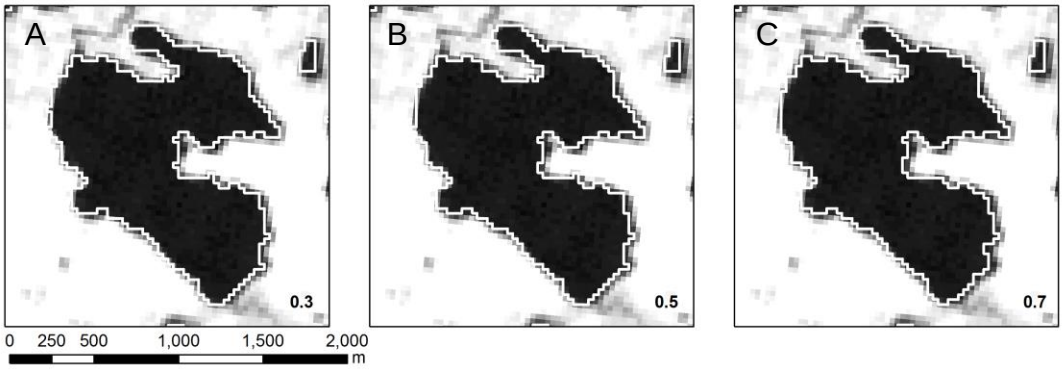


Figure 5.1. Error in lake polygon delineation methodology with respect to two ambiguous situations: the terrain-water transition zone and small streams. Imagery represents polygons and Landsat scenes from 4 August 2010 (Band 5 uncorrected). Polygons were delineated with three fuzzy membership α-cut thresholds of (A) 0.3; (B) 0.5; and (C) 0.7. Note that even in the lowest α-cut threshold, algorithms used in the method were more likely to classify pixels within these situations as terrain rather than water.

5.2. Evaluation of ERA-Interim Reanalysis data with ground observations.

Following the methods of Mugford et al. (2012), the ERA-Interim Reanalysis data were compared with observations from the Tuktoyaktuk weather station to check for systematic bias between variables in the two datasets. The resulting time series and corresponding correlation scatterplots from 1979 to 2013 for ERA-Interim and concurrent ground observations are shown in Figure 5.2.

The ERA-Interim monthly mean air temperatures corresponded well with observations at the Tuktoyaktuk Weather Station, capturing both inter-annual variability and seasonal temperatures well (Figure 5.2A). When removing estimated and interpolated data from ground observations, the correlation between the two datasets was nearly identical ($R^2 = 0.995$).

Monthly total precipitation was not as well represented between the two datasets. This was represented by a lower correlation between the two datasets, likely amplified by larger differences during months with extremely high precipitation levels ($R^2 = 0.754$). Gauge measurements of precipitation in the Arctic inherently contain large uncertainties due to wind-induced effects, differences in gauge types, difficulty in distinguishing various snow types, treatment of trace amounts, and other difficulties in cold environments (Legates and Willmott 1990, Walsh et al. 1998). Similar to Mugford et al. (2012), it was not surprising that the ERA-Interim data was not as tightly correlated with ground observations. Nonetheless, both precipitation datasets captured

a seasonal cycle with peak precipitations during the summer months. Bias-correction procedures are available to correct for precipitation uncertainties, in some cases increasing the total precipitation by up to 70% (Bowling and Lettenmaier 2010); however, correction factors are found to be smallest during the summer months (up to 10%; Yang et al. 2005). Additionally, these methods are beyond the scope of the study, and as such, the ERA-Interim data was used in the study with caution.

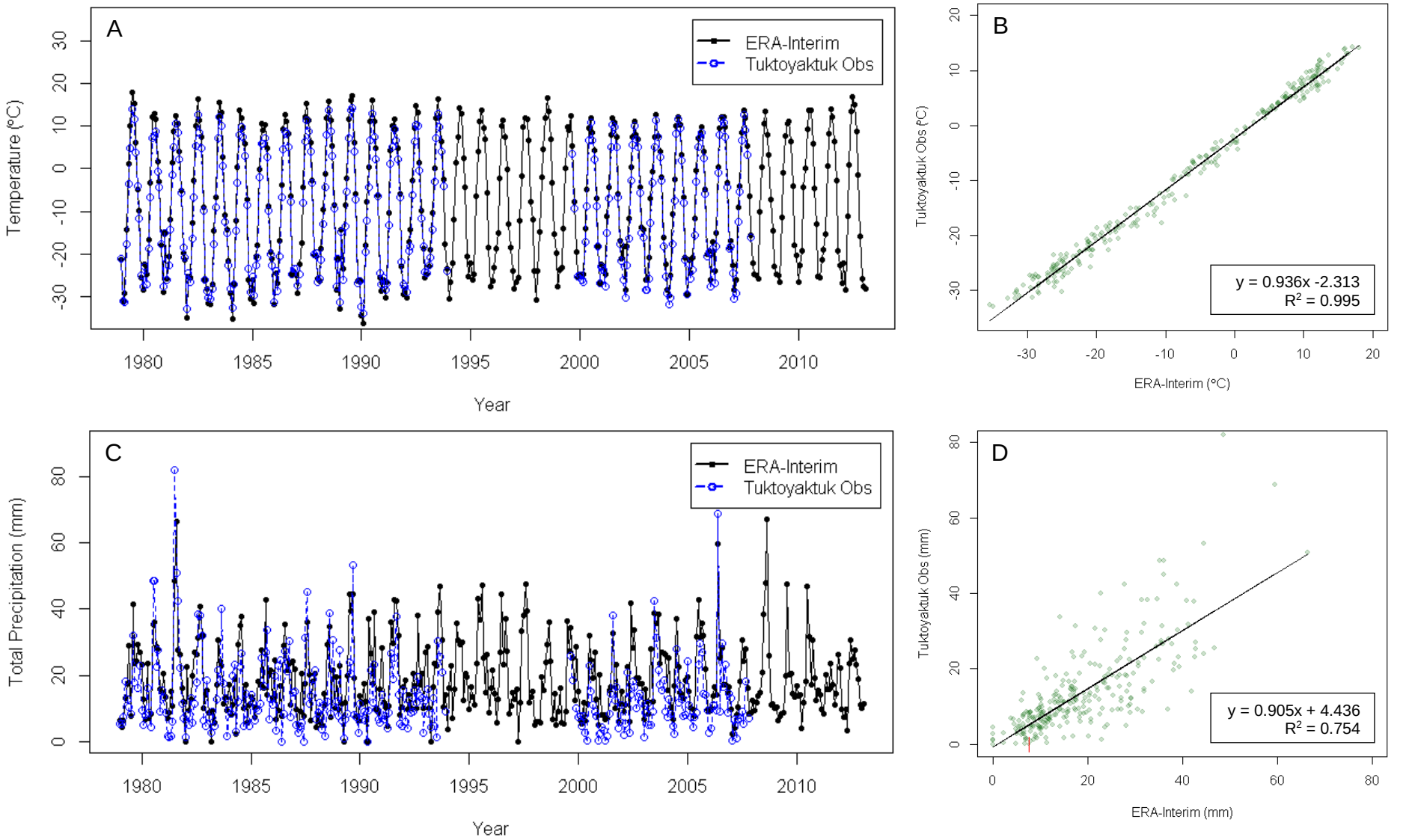


Figure 5.2. Comparative time series and corresponding correlation scatterplots of **(A, B)** monthly means of temperature and **(C, D)** monthly total precipitation for ERA-Interim data and ground observations at Tuktoyaktuk Weather Station for the years 1979 to 2012.

6. RESULTS.

In accordance with the structure and order of the Methods (Section 4), the results of the study are separated into three principal components: the presence of linear and bilinear trends in overall lake dynamics (Section 6.1); the statistical significance of spatiotemporal trends in lake surface area (Section 6.2); and the correlation of lake surface areal trends found in Sections 6.1 and 6.2 with candidate weather variables (Section 6.3). Note that as the majority of available Landsat imagery during the years 1987, 1992, and 1993 were either largely contaminated by cloud cover or could not be run through the Fmask algorithm for unknown reasons, a cloud-free representation of the entire study area proved difficult, and as such, these years were excluded from the time series.

6.1. Overall trends in lake dynamics.

6.1.1. Magnitude and direction.

Lakes occupied between 2,021 km^2 to 2,352 km^2 (22.5% to 26.2%) of the total land area on the Tuktoyaktuk Peninsula (8,990 km^2), and averaged 2,155 km^2 (24.0%) between the years 1985 to 2011 over all three fuzzy membership α-cut thresholds. A total of 16,486 distinct lakes and lake groups were identified over the entirety of the temporal period. Over the entire study area, the total surface areal coverage of lakes did not portray any significant unidirectional trends over the entire temporal domain (1985 – 2011; Figure 6.1A). The overall lake surface area remained between 2,100 km^2 to 2,300 km^2 during this period. This ambiguity was supported by the varying rates of change obtained from implementing three different fuzzy membership α-cut thresholds, which varied from -0.982 km^2/year to +0.586 km^2/year (Table 6.1).

Table 6.1. Linear temporal trends in overall lake surface area over time, for each fuzzy membership α-cut threshold. The regression estimate represented the rate at which lakes collectively increase or decrease in surface area over each year. Significant Chow *p*-values represented presence of a structural break between the years 1997 and 1998. All significant *p*-values are denoted by an asterisk (*).

α-cut	Temporal Domain	Estimate (km^2)	Intercept (km^2)	Adjusted R^2	GLM *p*-value	Chow value (*p*-value)
0.3	1985 – 1997	9.300	-16,293.381	0.151	0.146	6.990 0.004*
	1998 – 2011	-1.963	6,120.509	0.051	0.217	
	1985 – 2011	-0.777	3,749.260	-0.030	0.606	
0.5	1985 – 1997	8.124	-14,004.806	0.229	0.092	10.347 <0.001*
	1998 – 2011	-0.838	3,838.404	-0.069	0.700	
	1985 – 2011	0.882	393.332	-0.024	0.526	
0.7	1985 – 1997	6.805	-11,421.73	0.283	0.066	13.429 0.002*
	1998 – 2011	-3.016	8,172.910	0.232	0.047*	
	1985 – 2011	1.002	119.178	-0.004	0.352	

Table 6.2. Linear temporal trends in overall lake surface area over time for each area class, separated by fuzzy membership α-cut threshold and each area class: **(A)** 0.3; **(B)** 0.5; and **(C)** 0.7.. The regression estimate represented the rate at which lakes collectively increase or decrease in surface area over each year for the temporal domain 1985 to 2011. Significant Chow *p*-values represented presence of a structural break between the years 1997 and 1998. All significant *p*-values are denoted by an asterisk (*).

Table 6.2A. Linear temporal trends in surface area by area class using a fuzzy membership α-cut threshold of 0.3.

Area Class	Temporal Domain	Estimate (km^2)	Intercept	Adjusted R^2	GLM *p*-value	Chow value (*p*-value)
1	1985 – 1997	0.063	-110.726	-0.093	0.643	2.790 (0.083)
	1998 – 2011	-0.051	113.936	-0.005	0.352	
	1985 – 2011	-0.096	205.312	0.245	0.006*	
2	1985 – 1997	-0.151	415.181	0.005	0.299	3.433 (0.051)
	1998 – 2011	-0.188	488.661	0.016	0.293	
	1985 – 2011	-0.151	415.181	0.005	0.300	
3	1985 – 1997	2.434	-4,330.682	0.098	0.197	7.135 (0.004*)
	1998 – 2011	-0.510	1,531.280	0.073	0.180	
	1985 – 2011	0.030	450.330	-0.041	0.950	
4	1985 – 1997	4.693	-8,132.085	0.270	0.071	8.316 (0.002*)
	1998 – 2011	-1.064	3,326.947	0.029	0.26	
	1985 – 2011	-0.575	2,350.270	-0.009	0.386	
5	1985 – 1997	2.327	-4,267.650	0.172	0.129	6.144 (0.008*)
	1998 – 2011	-0.150	659.620	-0.055	0.578	
	1985 – 2011	0.016	328.187	-0.042	0.962	

Table 6.2B. Linear temporal trends in surface area by area class using a fuzzy membership α-cut threshold of 0.5.

Area Class	Temporal Domain	Estimate (km²)	Intercept	Adjusted R^2	GLM *p*-value	Chow value (*p*-value)
1	1985 – 1997	0.111	-206.716	-0.005	0.356	0.017 (0.012*)
	1998 – 2011	-0.059	149.619	0.031	0.256	
	1985 – 2011	-0.057	126.048	0.080	0.088	
2	1985 – 1997	0.240	-367.798	-0.078	0.573	10.180 (<0.001*)
	1998 – 2011	-0.001	110.874	-0.083	0.996	
	1985 – 2011	0.033	43.161	-0.039	0.811	
3	1985 – 1997	2.059	-3,599.193	0.190	0.116	16.556 (<0.001*)
	1998 – 2011	-0.085	672.143	-0.081	0.867	
	1985 – 2011	0.613	-726.736	0.047	0.148	
4	1985 – 1997	3.635	-6,048.446	0.199	0.110	6.762 (0.005*)
	1998 – 2011	-0.620	2,424.257	-0.060	0.618	
	1985 – 2011	0.147	888.561	-0.0394	0.821	
5	1985 – 1997	2.079	-3,782.641	0.285	0.646	7.377 (0.004*)
	1998 – 2011	-0.063	481.543	-0.080	0.844	
	1985 – 2011	0.146	62.262	-0.028	0.581	

Table 6.2C. Linear temporal trends in surface area by area class using a fuzzy membership α-cut threshold of 0.7.

Area Class	Temporal Domain	Estimate (km²)	Intercept	Adjusted R^2	GLM *p*-value	Chow value (*p*-value)
1	1985 – 1997	0.009	-4.692	-0.123	0.904	2.847 (0.079)
	1998 – 2011	-0.119	249.253	0.246	0.041*	
	1985 – 2011	-0.071	154.155	0.280	0.003*	
2	1985 – 1997	-0.136	376.411	-0.108	0.735	4.891 (0.017*)
	1998 – 2011	-0.068	241.589	-0.073	0.744	
	1985 – 2011	-0.020	205.11	-0.033	0.653	
3	1985 – 1997	2.099	-3686.965	0.285	0.065	24.274 (<0.001*)
	1998 – 2011	-0.416	1329.523	0.069	0.186	
	1985 – 2011	0.643	-793.878	0.080	0.088	
4	1985 – 1997	3.612	-6025.212	0.361	0.039*	9.009 (0.001*)
	1998 – 2011	-1.679	4535.321	0.139	0.103	
	1985 – 2011	0.523	121.300	0.002	0.318	
5	1985 – 1997	1.221	-2081.270	0.262	0.075	6.854 (0.005*)
	1998 – 2011	-0.733	1817.259	0.244	0.042*	
	1985 – 2011	-0.042	432.548	-0.039	0.819	

When lakes were portioned into the five defined areal classes (Table 4.2), medium and large lakes (Class 3 and Class 4) saw both the greatest increases and decreases in lake area over the entire temporal domain (1985 – 2011), highlighting its influence in driving overall changes in lake area. As the α-cut threshold was increased from 0.3 to 0.7, these classes changed from a decreasing to an increasing trend, due to lakes switching from a higher to a lower areal class (Table 6.2). This finding was attributed to either intraannual variability or the presence of a structural break (explained further in Section 6.1.2).

In contrast, the total count of lakes had a significant decreasing trend over the entire temporal domain (1985 – 2011; Figure 6.1C). The average rates of change obtained from using the three different fuzzy membership α-cut thresholds all portrayed similar rates of decrease at a loss of 61, 58, and 72 lakes/year for thresholds set at 0.3, 0.5, and 0.7 respectively (Table 6.3).

The smallest class of lakes (Class 1) constituted a large majority within the total amount of lakes (~66%; Figure 6.2, Table 6.4), and subsequently drove the negative trend at -28 to -34 lakes/year, depending on the α-cut thresholds used (Table 6.3). Because the changes in area within Class 1 were also the least influential, as indicated by the low area-to-count ratio (Table 6.4), the large decreasing trend observed within this class only had a small effect on the overall changes in surface area. Correspondingly, the distribution of lake count was determined to be logarithmic, which renders the data difficult to analyse due to its resulting skewed distribution (Figure 6.1D). The distribution of lake surface area, on the other hand, approximately follows a normal distribution (Figure 6.1B). Finally, as it is probable that lake count would change with the use of higher resolution imagery that potentially identifies lakes smaller than the 1-pixel equivalent in this study (900 m^2), lake count is therefore inherently dependent on specific aspects of Landsat data quality. Although the study acknowledged the variability of count data in this section, due to the above reasons, the rest of the study focused on changes in lake surface area rather than count.

Table 6.3. Linear temporal trends in overall lake count over time, for each fuzzy membership α-cut threshold, each area class, and the combination of all classes. The regression estimate represented the rate at which lakes collectively increase or decrease in count over each year for the temporal domain 1985 to 2011. All significant *p*-values are denoted by an asterisk (*).

α-cut	Area Class	Estimate	Intercept	Adjusted R^2	GLM *p*-value
0.3	1	-34.33	72,500.02	0.295	0.003*
	2	-17.610	39,534.82	0.402	0.001*
	3	-7.341	17,278.92	0.189	0.019*
	4	-1.047	33,24.168	-0.033	0.608
	5	-1.633	3,398.052	0.068	0.116
	Total	-62.145	136,352.75	0.392	<0.001*
0.5	1	-28.249	60,237.40	0.242	0.009*
	2	-14.915	34,063.14	0.313	0.003*
	3	-11.033	24,684.73	0.330	0.002*
	4	-2.434	60,69.873	0.029	0.207
	5	-1.735	3,603.085	0.083	0.093*
	Total	-58.373	128658.22	0.359	<0.001*
0.7	1	-31.460	66,505.85	0.420	<0.001*
	2	-18.886	41,924.19	0.564	<0.001*
	3	-13.560	29,787.00	0.360	0.001*
	4	-6.578	14,380.07	0.204	0.015*
	5	-2.123	4,383.690	0.116	0.057
	Total	-72.610	156,981.12	0.536	<0.001*

6.1.2. Comparison of longer-term and recent trends

Over the entire study area, the rates of areal decline for recent temporal records (1998 – 2011) were more negative than when compared to the rates of decline over the entire temporal domain (1985 – 2012; Table 6.1). Chow tests revealed a significant structural break in the total surface areal coverage time series for all three fuzzy membership α-cut thresholds (all $p < 0.004$; Table 6.1), separating the time series into two periods 1985 – 1997 and 1998 – 2011. Consequently, areal trends for all three thresholds were bifurcated into an initial positive correlation and a subsequent negative correlation (Figure 6.2A – C). The bifurcation resulted in an increase in statistical significance (decrease in *p*-value) for all regressions (Table 6.1).

When the total surface areal coverage was partitioned into areal classes, the structural break remained statistically significant (Table 6.2). In general, larger classes were found to have higher Chow statistics, with medium and large lakes (Classes 3 and 4) presented the highest Chow statistics overall. These lakes also saw the largest increases and decreases in lake area, respective of temporal period. Class 1 was the only class that was better represented as a continuous time series, rather than as two distinct periods.

The presence of a significant structural break in lake surface areal trends forms the basis of the next section (Section 6.2), where the period indicating an initial positive correlation in lake areal extent is referred to as the first temporal period, and the period indicating the subsequent decreasing correlation is referred to as the second temporal period.

Table 6.4. Average surface area and total count coverage and percentage, and the area-count percentage ratio for each fuzzy membership α-cut threshold, classified into five area classes (defined in Table 4.2). The area-count percentage ratio represents the weight of lakes in each class in influencing changes in total surface area.

α-cut	Area Class	Average Areal Coverage (km^2)	Percentage of Total Area (%)	Average Total Count	Percentage of Total Count	Area-to-Count Ratio
0.3	1	13.279	0.617	3,918	31.967	0.019
	2	112.651	5.231	4,366	35.615	0.147
	3	510.173	23.690	2,636	21.506	1.102
	4	1,198.994	55.676	1,204	9.828	5.665
	5	359.689	16.702	133	1.085	15.394
	Total	2,153.510	100	12,257	100	1
0.5	1	12.616	0.595	3,785	30.936	0.019
	2	108.631	5.125	4,324	35.348	0.145
	3	498.002	23.496	2,718	22.216	1.058
	4	1,180.323	55.687	1,270	9.829	5.363
	5	353.938	16.699	137	1.117	14.949
	Total	2,119.545	100	12,233	100	1
0.7	1	11.923	0.5433	3,649	30.015	0.018
	2	104.606	4.766	4,267	35.095	0.136
	3	490.499	22.348	2,783	22.895	0.976
	4	1,164.654	53.065	1,317	10.836	4.897
	5	347.863	15.850	141	1.159	13.673
	Total	2,194.786	100	12,157	100	1

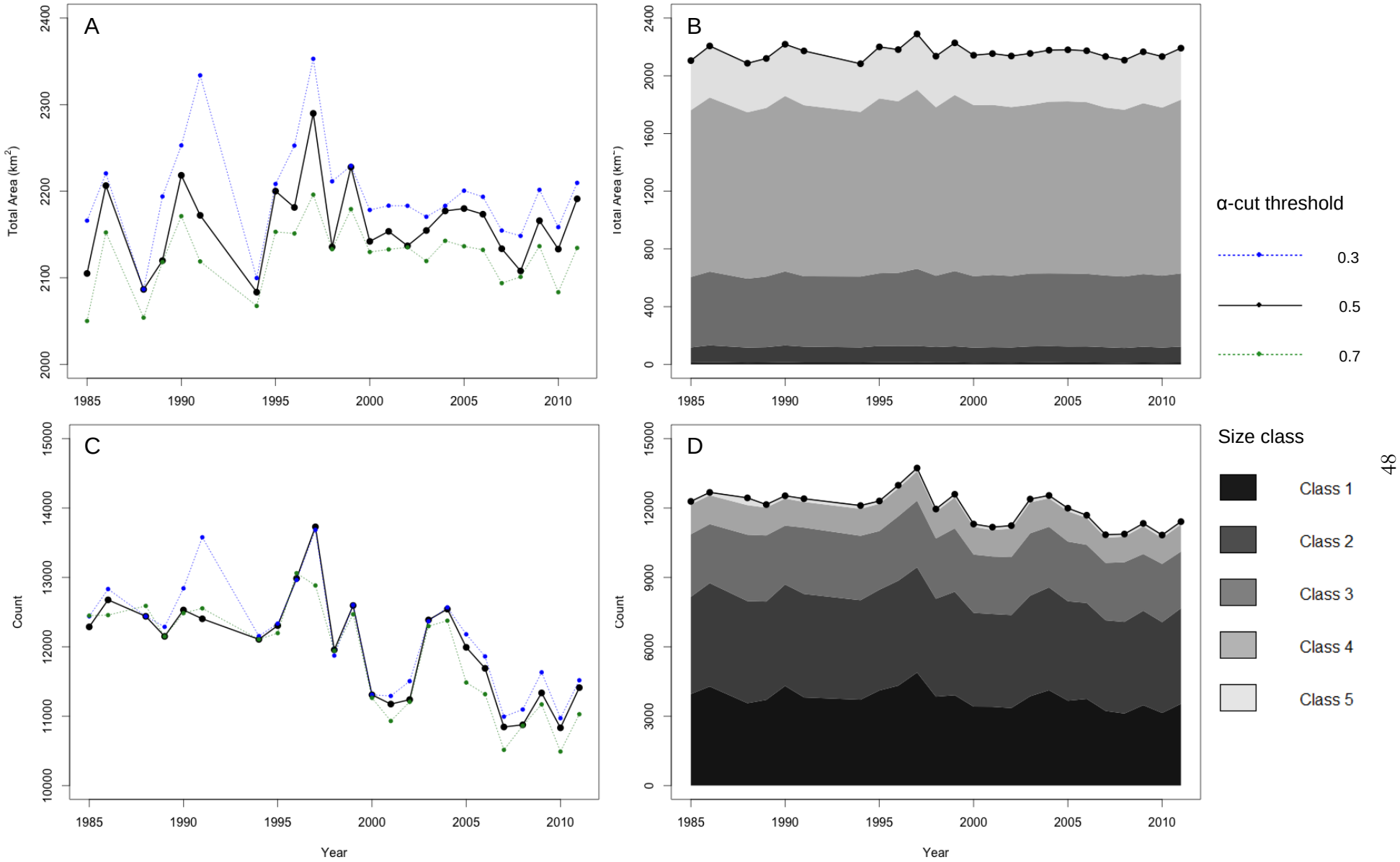


Figure 6.1. Comparative time series of three fuzzy membership α-cut thresholds and five lake surface area classes (using an α-cut threshold of 0.5), respectively, for **(A, B)** total surface area and **(C, D)** lake count for the years 1985 to 2011.

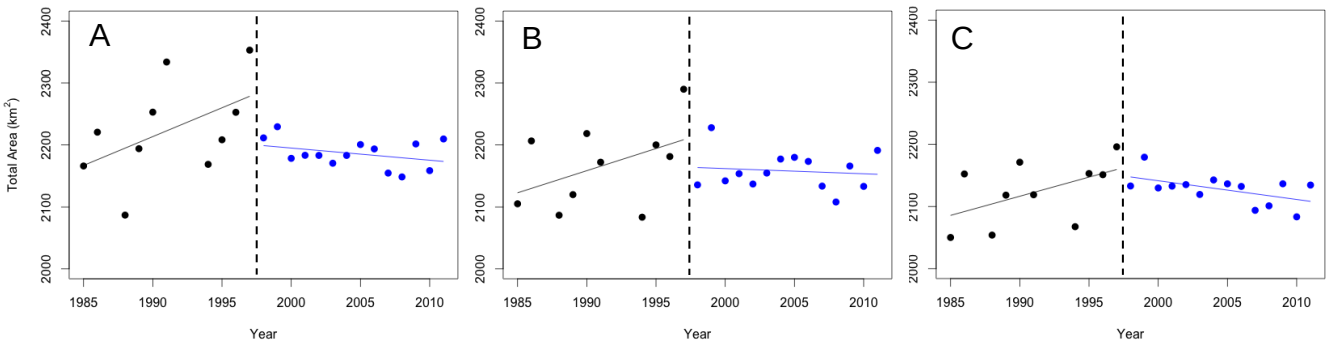


Figure 6.2. Time series of total surface area for three fuzzy membership α-cut thresholds of **(A)** 0.3; **(B)** 0.5; and **(C)** 0.7. Trendlines were calculated using a general linear model regressing total area by year, and are denoted by the solid line. The significant structural breaks, as identified by Chow tests, were denoted by the dotted line between years 1997 and 1998, where all $p < 0.05$. The structural break separated the time series into an initial positive correlation (black) and a subsequent negative correlation (blue).

6.2. Spatial heterogeneity in lake dynamics.

6.2.1. Individual rates of change.

When examining the trends of individual lakes, a wide range in annual trends were found that included both substantial decreases and increases in lake area (Figure 6.3). Over the entire temporal domain, there was on average a very slight increasing trend in individual lake surface area, although the range of values was substantially variable (standard deviation = 5,348 m^2/year; Table 6.5). Lakes displayed a higher positive trend in surface area increase during the first period (2,375 m^2/year) than the second period (474 m^2/year; Table 6.5).

Table 6.5. Summary statistics for GLMs regressing lake area over time. The Global Moran's *I* statistic was used to evaluate whether lakes exhibiting significant increasing or decreasing areal trends were spatially clustered, dispersed, or randomly distributed across the overall study area. The corresponding scale was the optimized distance band that determined the spatial extent of clustering.

Temporal Period	# of Lakes	Mean Estimate (m^2/year)	Estimate Standard Error (m^2/year)	Moran's *I* Value (Scale; m)	Moran's *I* *p*-value
1985 – 1997	527	2,375.747	1,098.375	-0.005 (23,000)	0.045*
1998 – 2011	1293	474.434	346.094	0.016 (22,000)	0.002*
1985 - 2011	1872	23.776	195.804	-0.005 (23,000)	0.045*

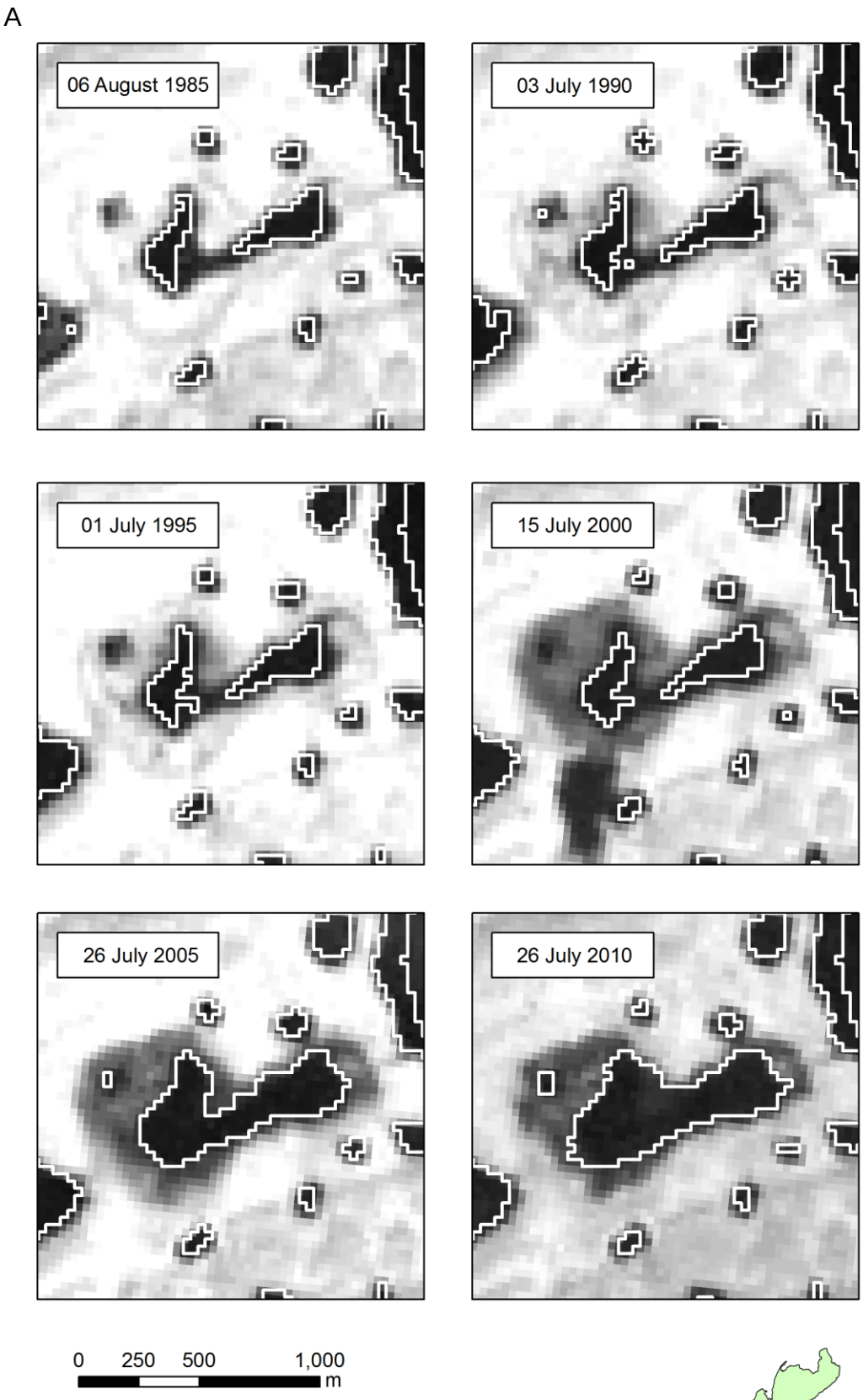


Figure 6.3A. Time series of Landsat images (Band 5 uncorrected) showing a lake in the South Eskimo Lakes region with a significantly **(A)** increasing and **(B)** decreasing trend. Delineations of lake extent represent polygons generated using a fuzzy membership α-cut threshold of 0.5. Individual lake equation was obtained using a GLM as described in Section 4.5.2.

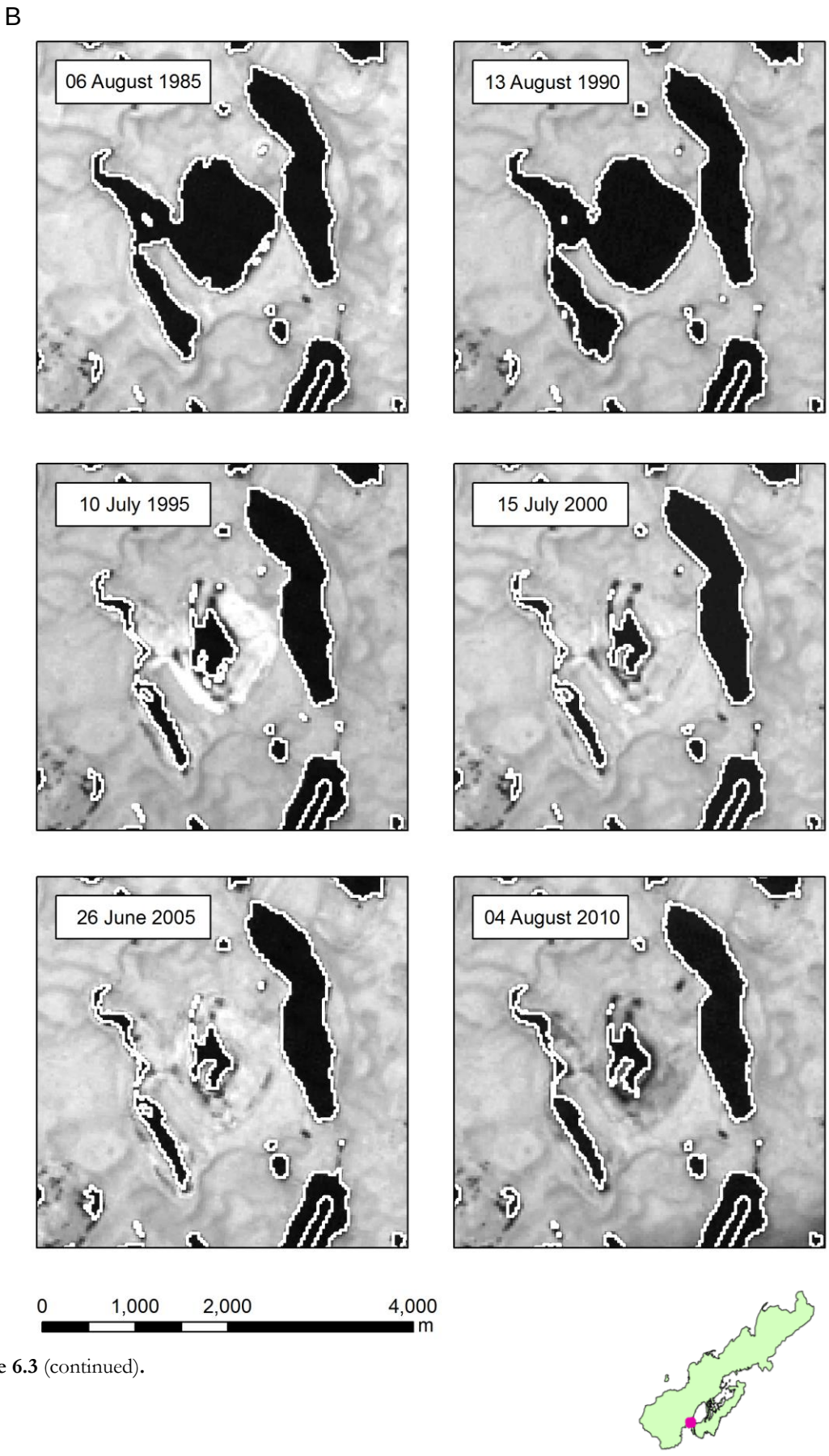


Figure 6.3 (continued).

6.2.2. Spatiotemporal variability in lake area.

When running the Global Moran's *I* test on individual lakes showing a significant area trend, Moran's *I* values were positive and significant (all $p < 0.05$) for all temporal domains, indicating that lakes with high or low areal trends were likely to be spatially clustered (Table 6.5; Figure 6.4). Maximum clustering of individual lake areal trends occurred at scales ranging from 22 to 23 km, suggesting broad-scale heterogeneity in trends (Table 6.5). Over the entire temporal domain, lake trends were overall spatially segregated, with lakes showing an increasing areal trend located at the bottom half of the peninsula, and lakes showing a decreasing trend located at the upper half of the peninsula (Figure 6.4A).

The Getis-Ord's Local G_i^* test determined that lakes showing the most extreme areal trends were likely to be spatially clustered in several hot and cold spots (Figure 6.5). In this study, hot spots were defined as a cluster of lakes collectively showing a significant increase in surface area over time, and cold spots defined the converse, representing a cluster of lakes collectively showing significant decreasing trends. Several hot and cold spot clusters of large lakes were identified over the entire temporal domain, and represented various spatial extents (Figure 6.5A). The presence of a mid-peninsula transition zone (Box 1) was the most noticeable of these clusters, spanning an area of approximately 400 km^2. At the meso-scale, hot and cold spots were also identified spanning the northern coastline of the peninsula (Boxes 2 – 4). A significant hot spot was further identified at the local scale in an area otherwise depicting unidirectional decreases in lake areal extents, indicating spatial heterogeneity at finer spatial extents (Box 5).

During the first temporal period, lakes were seen to cluster most significantly at the Southern region of the Eskimo Lakes Region (Figure 6.5C inset). This meso-scale hot spot was also observed during the second temporal period, forming the southern extent of a corridor of lakes running northwards along the middle of the peninsula, collectively exhibiting increasing areal trends. Regardless of temporal extent, hot spots were located almost exclusively in the lower half of the peninsula, and cold spots were located mostly within the upper half of the peninsula.

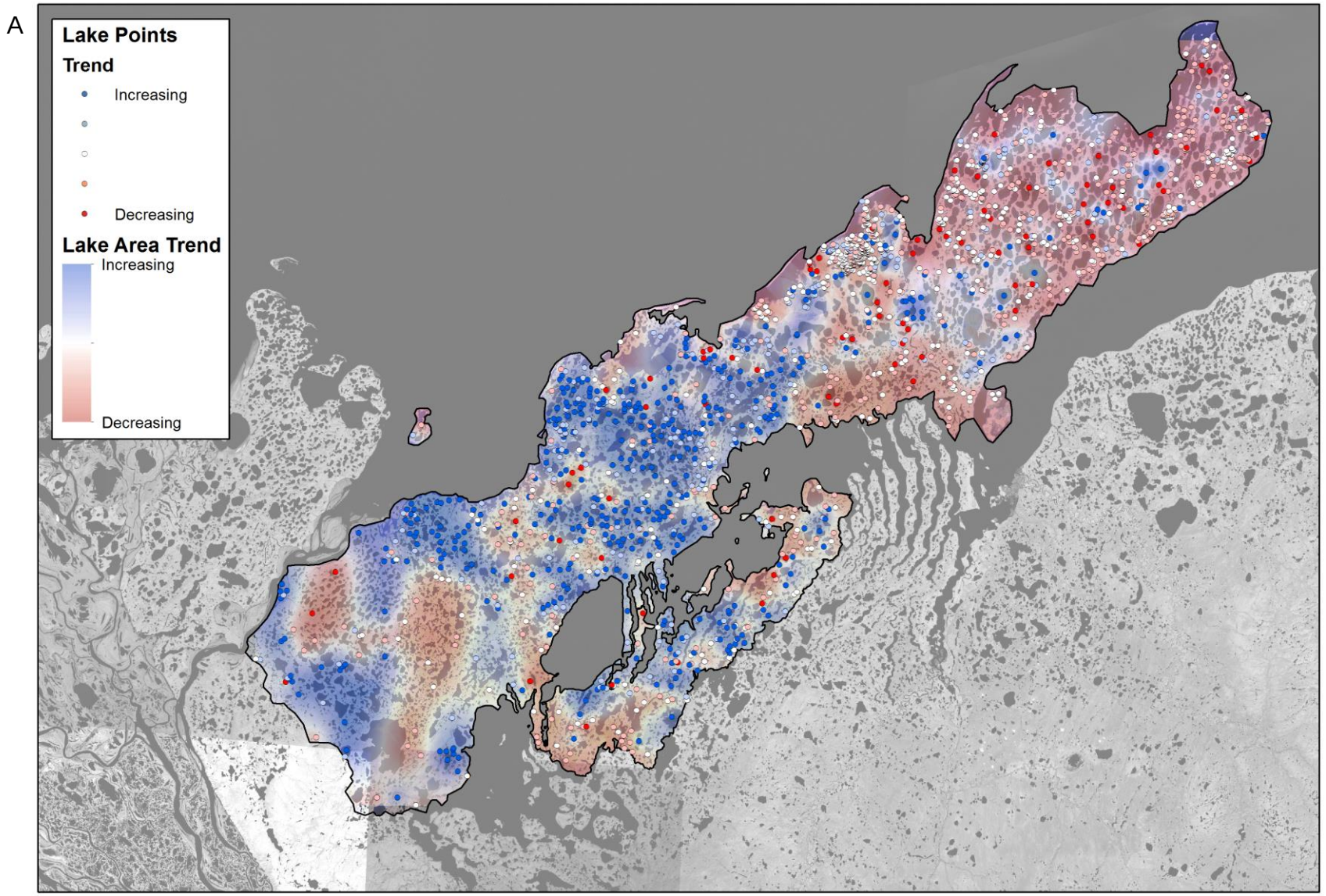
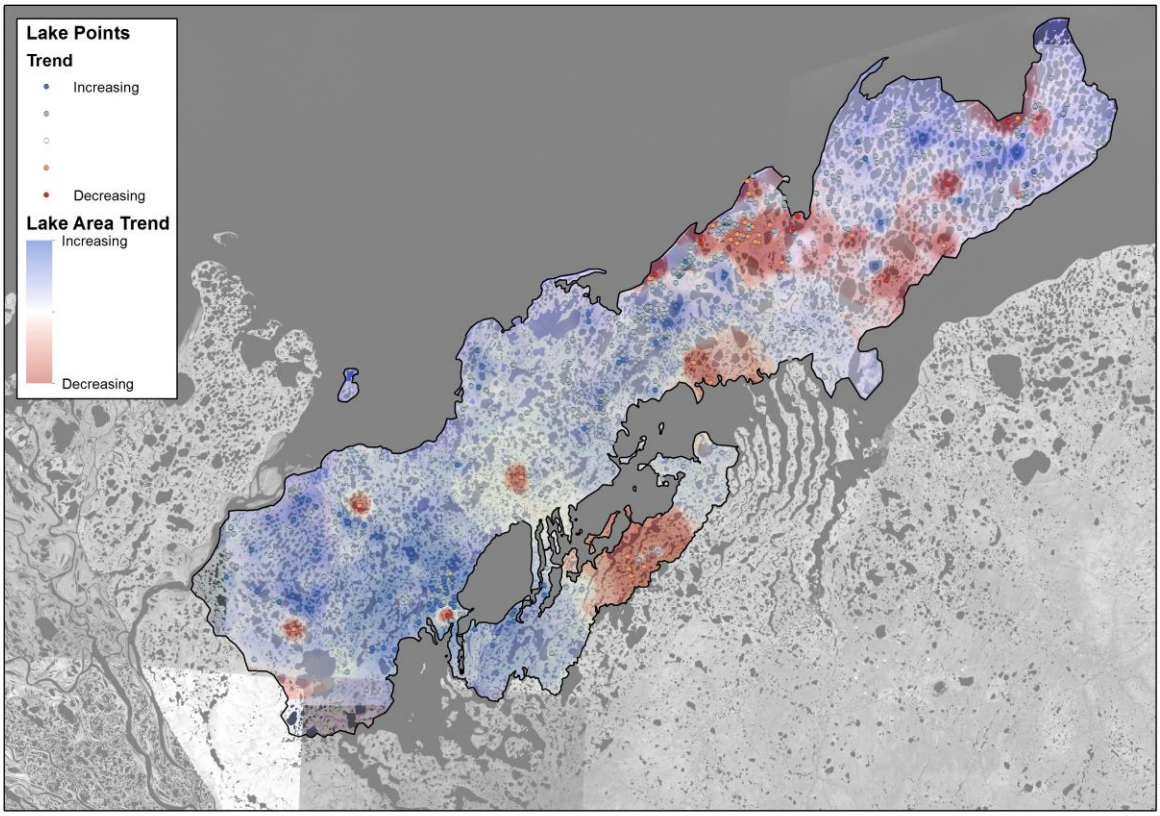


Figure 6.4. Spatial distribution of lake surface areal trends over the Tuktoyaktuk Peninsula for **(A)** the entire temporal domain 1985 to 2011; **(B)** the first temporal period 1985 to 1997; and **(C)** the second temporal period 1998 to 2011. Lake areas were determined using a fuzzy membership α-cut threshold of 0.5. Ordinary kriging was used to provide a visual representation to spatially portray individual lakes and their respective areal trends. Lake area trends were portrayed as points.

B



C

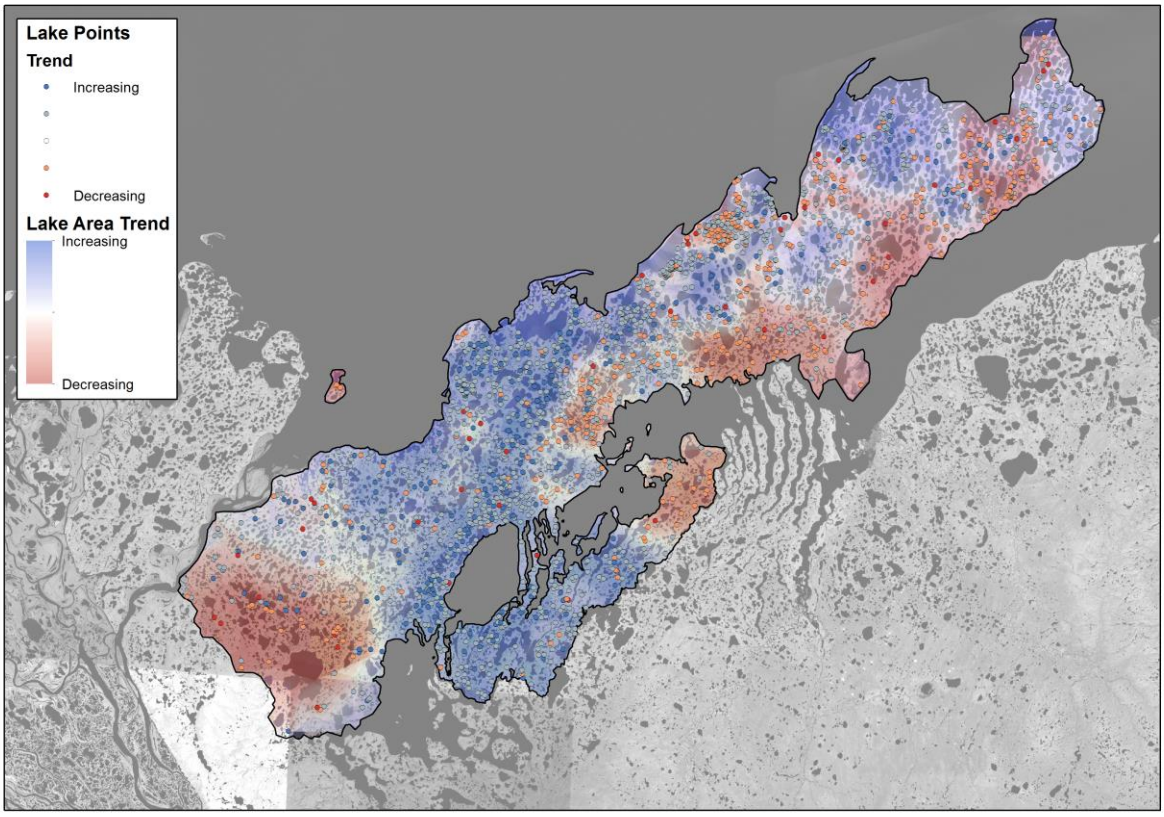


Figure 6.4 (continued).

A

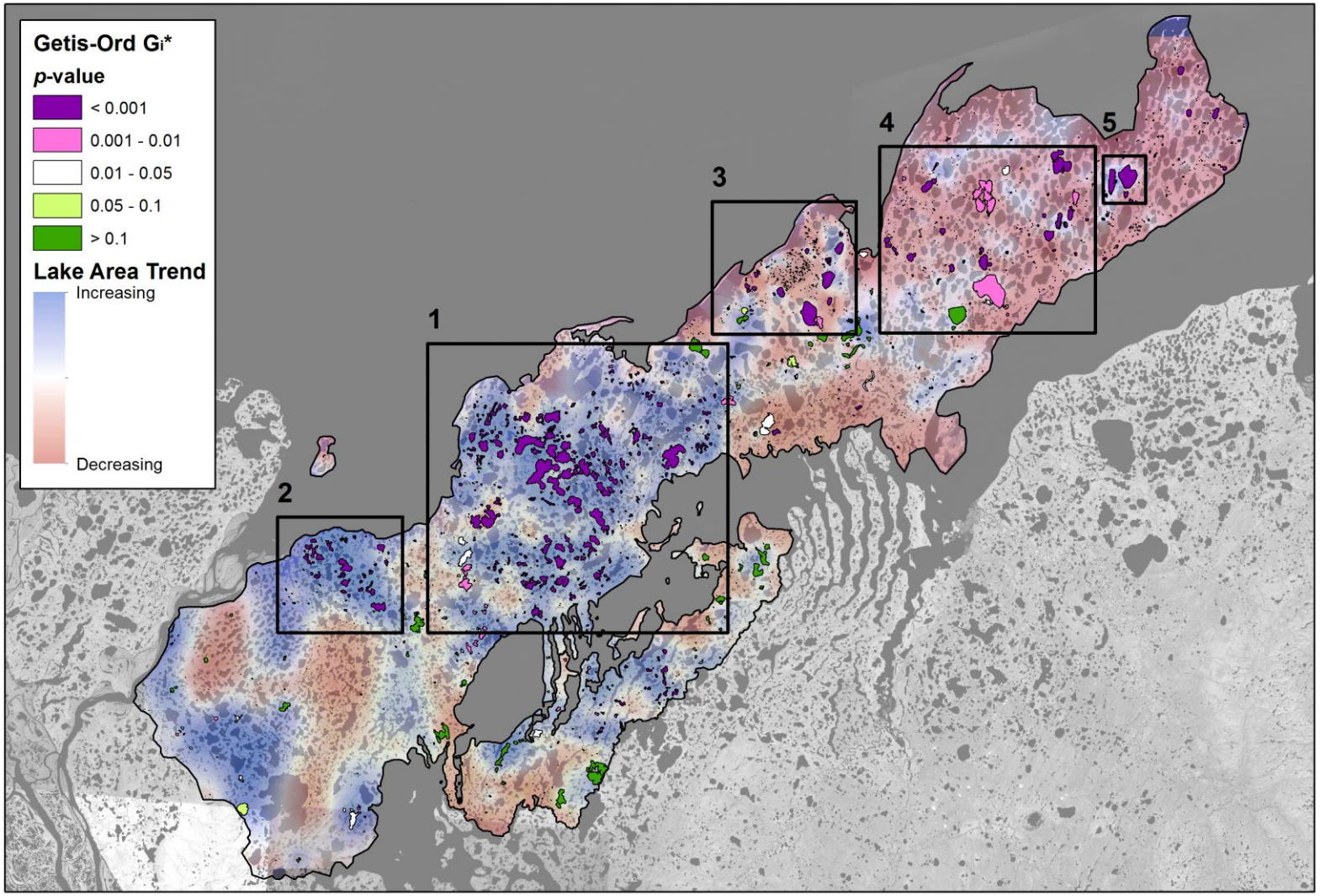
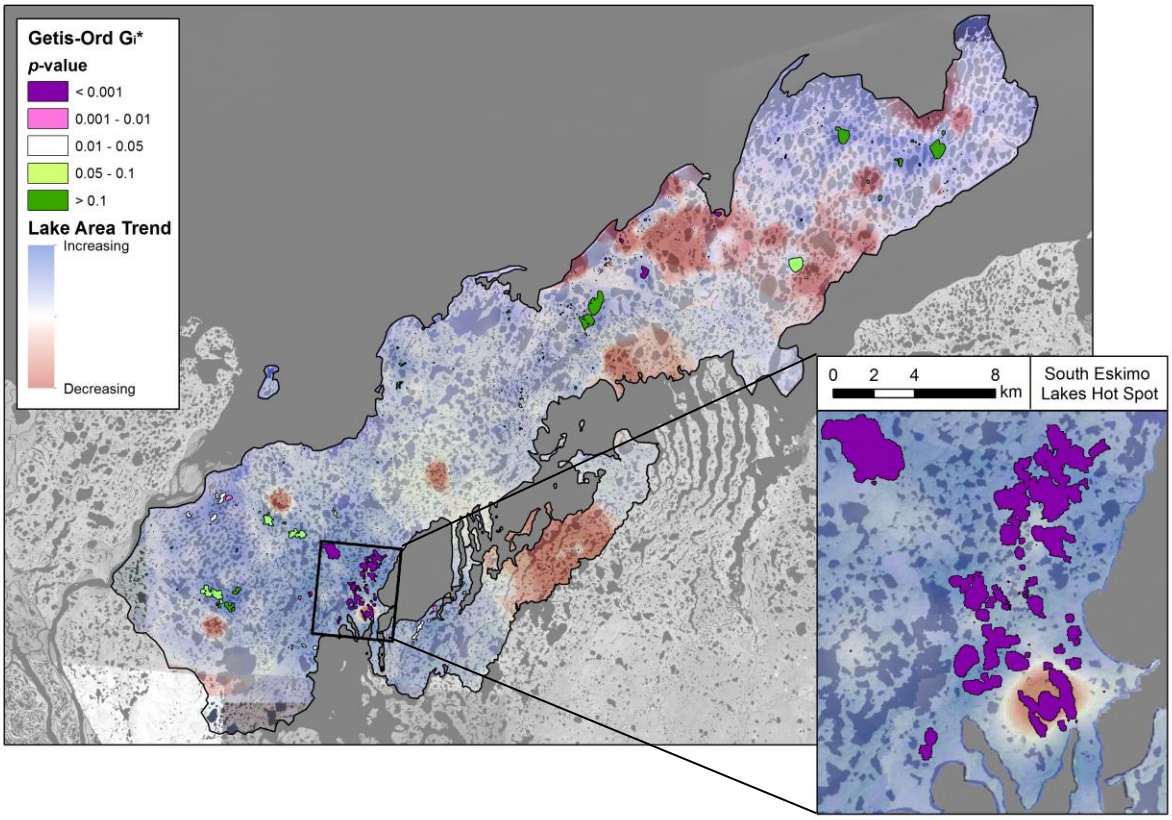


Figure 6.5. Significant hot and cold spot clusters of lake surface areal trends over the Tuktoyaktuk Peninsula for **(A)** the entire temporal domain 1985 to 2011; **(B)** the first temporal period 1985 to 1997; and **(C)** the second temporal period 1998 to 2011. Lake areas were determined using a fuzzy membership α-cut threshold of 0.5. Ordinary kriging was used to provide a visual representation to spatially portray individual lakes and their respective areal trends. Numbers and inset figures provide a closer depiction of transition zones and hot and cold spot clusters.

B



C

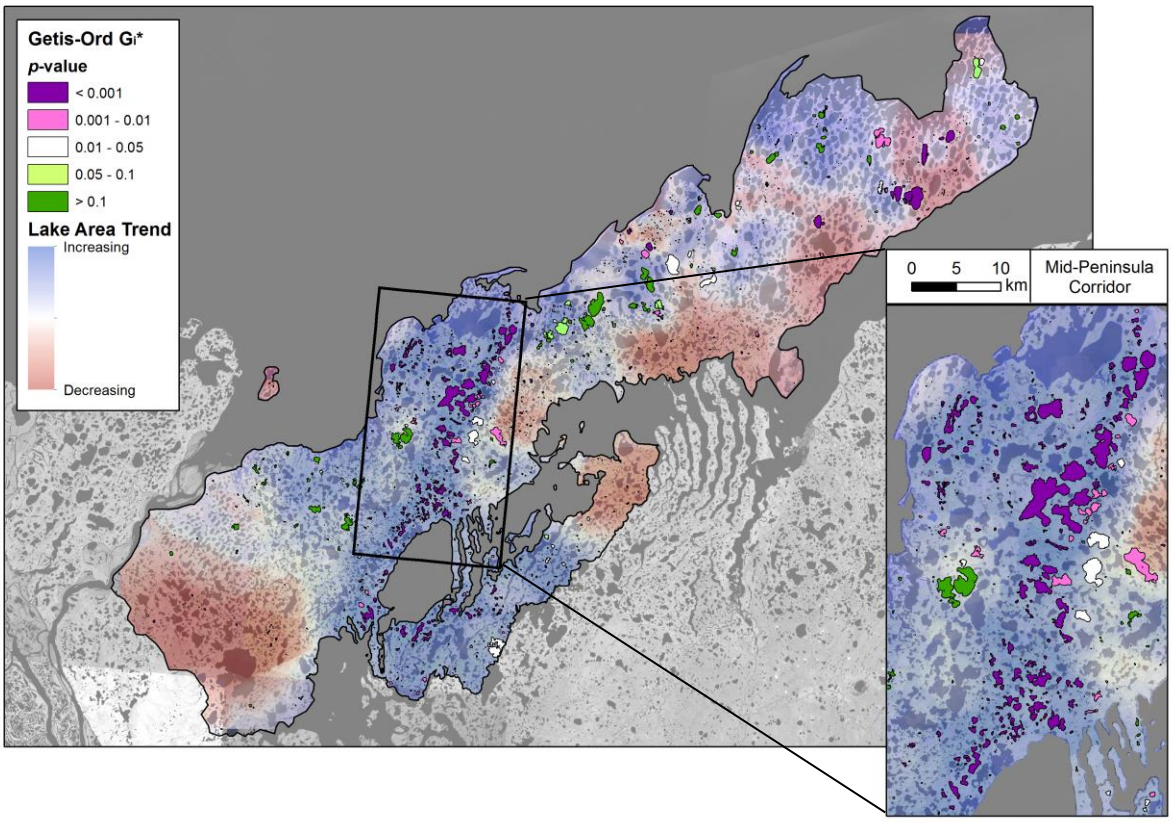


Figure 6.5 (continued).

6.3. Relationship between lake area and weather variables.

6.3.1. Overall rates of change.

Analysis of climate data derived from the ERA-Interim Reanalysis Dataset revealed either weak or no significant trends in climatological time series (Figure 6.6). Both the annual mean and annual growing season (June to September) temperature over the Tuktoyaktuk Peninsula showed a weak linear increase and decrease over time, respectively ($R^2 = 0.093$ and $R^2 = 0.004$). Although not statistically significant, average growing season temperatures were noted to increase almost linearly after 1999.

A slight linear increase in total annual precipitation was also found; however, the correlation was extremely low ($R^2 = -0.04$) and consequently is better described as having no observable trend during the short time period specified. The total precipitation during the growing seasons revealed a statistically significant structural break between the years 1997 and 1998, separating the time series into an initial positive correlation that increased at a moderate rate, and a subsequent correlation that was also positive, but with a lower rate of increase, and with significantly lower average values ($p < 0.05$) than the initial temporal period. The presence of the same structural break in both the total growing season precipitation and the lake surface area time series suggests that the structural break is robust. Consequently, the reductions in total growing season precipitation rates and values during the second temporal period are coincident with the declining trend seen in lake surface areas. For most of the years corresponding to the Landsat images, both the total annual and total growing season precipitation were within 1 standard deviation of the 1979 – 2012 mean (227.608 mm ± 34.452 mm; 127.905 mm ± 27.866 mm).

Similarly, the effect of climatic variables on overall lake surface area over the study region revealed little correlation (Table 6.6). Mean annual temperatures were found to have the largest effect among the climatic variables, decreasing lake areal extents by -13 $km^2/°C$ ($R^2 = 0.058$).

Table 6.6. Statistics of GLMs regressing individual lake surface areal change over individual climatic variables. Chow tests were conducted on each variable time series (variable ~ year) to determine the temporal period used for the GLM. The regression estimate represented the relative influence of each climatic variable in driving increase or decrease in lake area. None of the variables were significant at $p = 0.05$.

Variable	Temporal Period	Estimate	Intercept	Adjusted R^2	GLM *p*-value	Chow *p*-value
Mean Annual Temperature	1985 – 2011	-12.837	13.335	0.058	0.153	0.378
Mean Growing Season Temperature	1985 – 2011	-2.573	14.118	0.373	0.775	0.872
Total Annual Precipitation	1985 – 2011	0.063	13.887	-0.050	0.844	0.101
Total Growing Season Precipitation	1985 – 1997	0.987	2,027.875	0.052	0.256	0.042*
	1998 – 2011	-0.374	2,203.666	0.056	0.209	

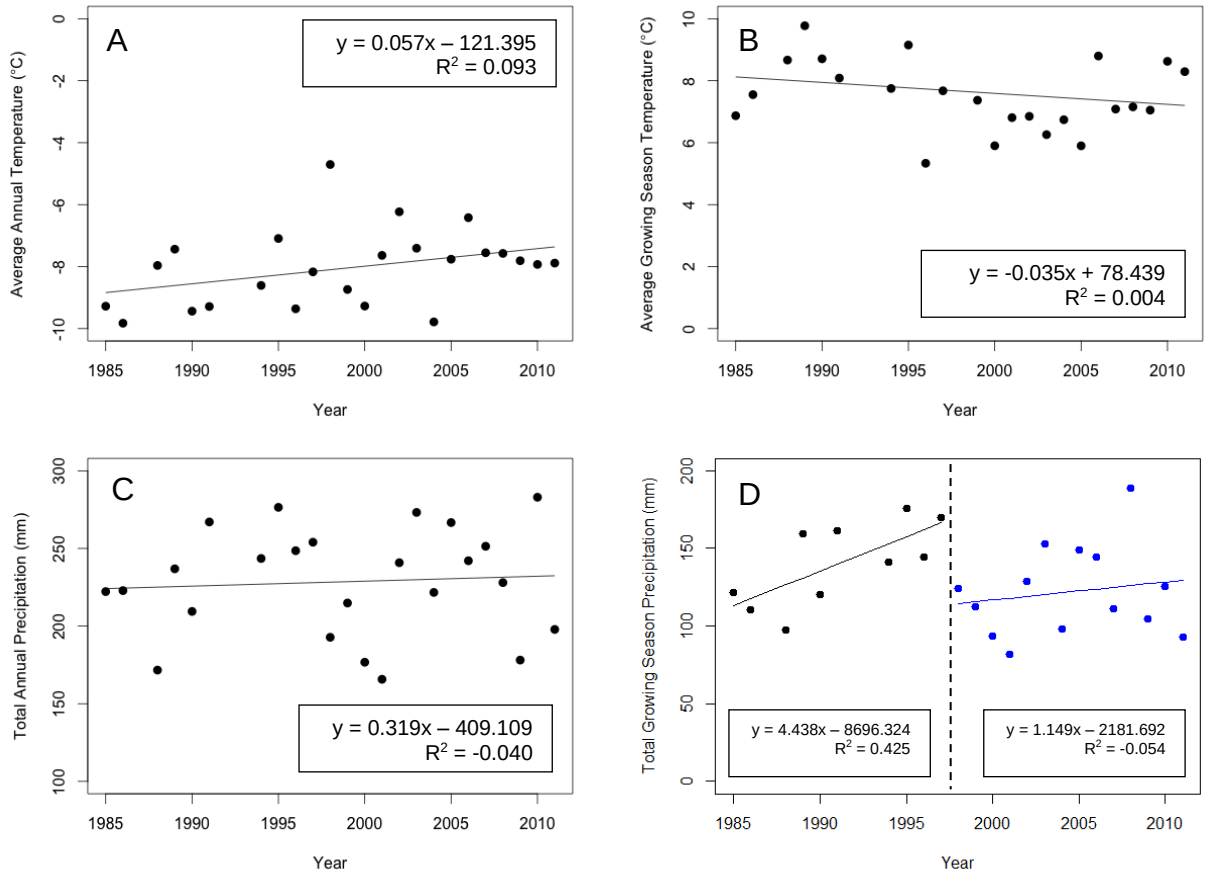


Figure 6.6. Time series of **(A)** mean annual temperature**; (B)** mean temperature during the growing season; **(C)** total annual precipitation; and **(D)** total precipitation during the growing season. The growing season was defined as the period between 01 June and 30 September of the coinciding year. Time series were derived from the ERA-Interim Reanalysis dataset. The significant structural break, as identified by a Chow test, was denoted by the dotted line between years 1997 and 1998, where $p < 0.05$. The structural break separated the time series into an initial strong positive correlation (black) and a subsequent weaker positive correlation (blue).

6.3.2. Spatiotemporal variability in lake area.

Examining the effect of climatic variables on lake surface area at the individual scale revealed varying ranges that included both increases and decreases in lake area (Table 6.7). Lakes were found to decrease both with increasing mean annual temperatures (-2,105 m^2/°C) and with increasing mean growing season temperatures (-404 m^2/°C). The total annual precipitation was found to negatively correlate with lake area at a rate of -47 m^2/mm; however, increasing total growing season precipitation had a positive effect on lake area (+60 m^2/mm). Overall, negative correlations were found to be more extreme than positive correlations (Figure 6.7).

Table 6.7. Summary statistics for GLMs regressing lake area over climatic variables. The Global Moran's *I* statistic was used to evaluate whether lakes exhibiting significant increasing or decreasing areal trends were spatially clustered, dispersed, or randomly distributed across the overall study area. The corresponding scale was the optimized distance band that determined the spatial extent of clustering. The total number of lakes regressed (both statistically significant and insignificant) was 1,872, as determined in Table 6.5.

Variable	# of Lakes	Mean Estimate (m^2/°C or mm)	Estimate Standard Error (m^2/°C or mm)	Moran's *I* Value (Scale; m)	Moran's *I* *p*-value
Mean Annual Temperature	189	-2,104.89	906.001	0.094 (20,000)	<0.001*
Mean Growing Season Precipitation	100	-403.656	603.872	0.053 (20,000)	0.021*
Total Annual Precipitation	47	-46.761	21.816	0.157 (17,000)	0.008*
Total Growing Season Precipitation	105	59.732	27.536	0.049 (22,000)	<0.001*

Moran's *I* values were positive and significant (all $p < 0.05$) for all climatic variables, indicating that each climatic variable had a greater effect on certain regions than others within the entire study area (Table 6.7; Figure 6.7). Lakes were found to have the highest degree of clustering at scales between 17 and 22 km, indicating that spatial autocorrelation was most significant at the broad scale.

The spatial distribution of individual lake areal-climate trends exhibited similar spatial patterns observed with the distribution of individual lake areal trends over the entire temporal period (Figure 6.9). Consequently, the lake areal-climate trends were also spatially segregated, with lakes depicting positive trends located almost exclusively within the upper half of the peninsula, and lakes depicting negative correlations located on the lower half. Overall, annual mean

temperatures were found to exert the largest effect on the lower half of the peninsula, especially around Kugmallik Bay, where an increase in temperatures tended to reduce individual lake surface area. Conversely, annual growing season temperatures may have had a more profound effect on the upper half of the study area, possibly driving expansion in lake areal extents with increasing temperatures. The significant effects of total annual precipitation on lake areal trends were restricted to within the lower half the peninsula, where interestingly, an increase in precipitation resulted in a reduction of lake surface area. This trend was similar with total growing season precipitation. These results could be interpreted as either 1. a delayed response of lakes to precipitation that was not identified in this study; 2. the relative weakness of precipitation as a forcing weather variable on lake surface areal change; or 3. faults in the statistical analysis and/or methodology. These interpretations are discussed further in the Discussion of the study (Section 6.1.2).

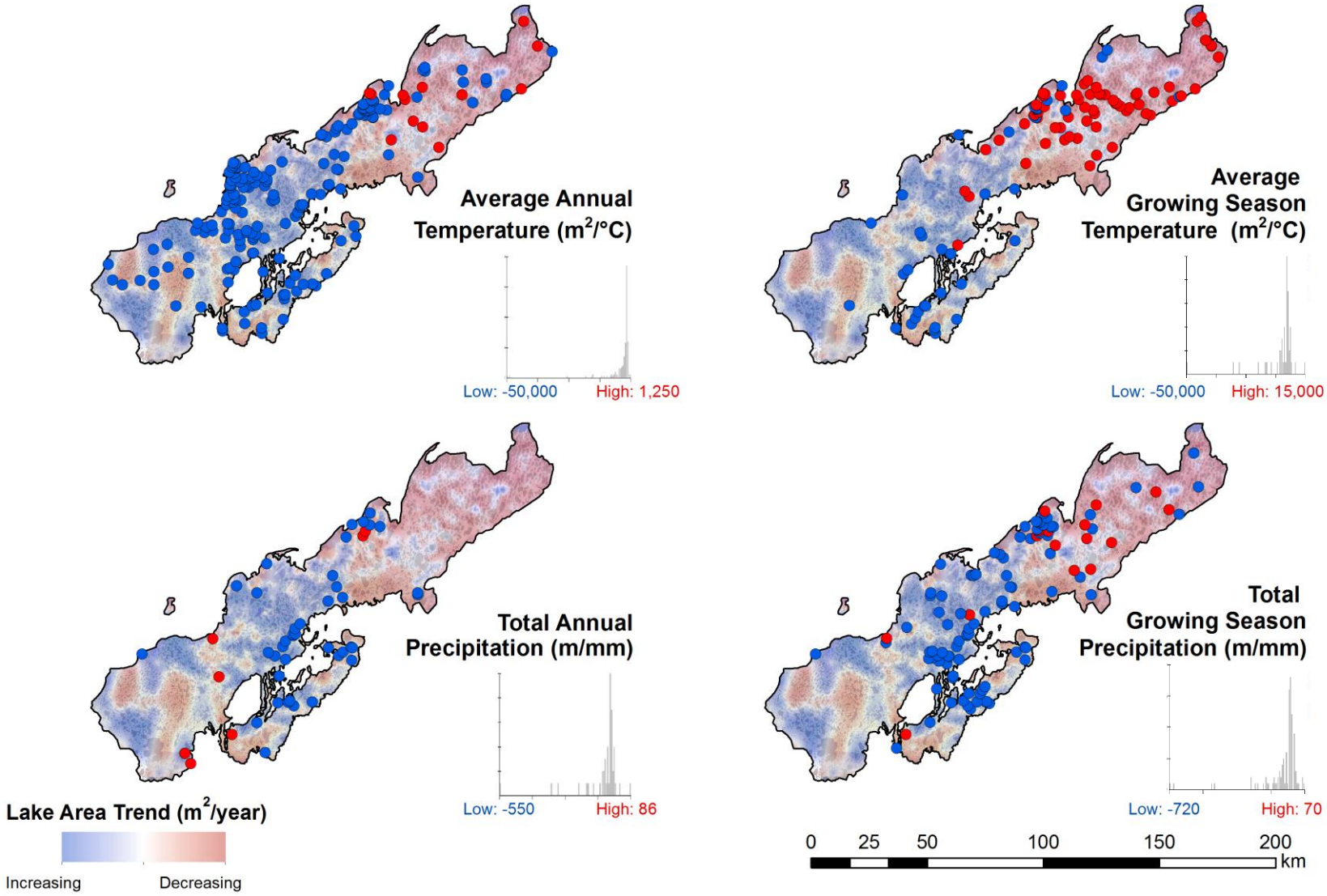


Figure 6.7. Spatial distribution depicting the effect of climatic variables on individual lake surface area over the Tuktoyaktuk Peninsula. The surface area of individual lakes were found to be affected by (A) yearly average temperatures; (B) average growing season temperatures; (C) total annual precipitation; and (D) total growing season precipitation. Trends were determined by GLMs regressing lake area over each climatic variable, where the estimate (slope; represented as red [positive] and blue [negative] points) was represented as the change in areal extent (m^2) per climatic variable unit (°C or mm). Histograms show the distribution of slope values for each climatic variable trend. Trend points were overlaid on lake surface areal trends over the time period 1985 – 2011.

7. DISCUSSION.

This study was among the first to implement a fully automated land-cover classification algorithm to detect and examine changes seen in lake surface area, with relative success. This study was also among the first to evaluate a near-complete yearly time series of homogeneous geospatial imagery in a lacustrine environment. Intra-annual variability was addressed by employing fuzzy membership techniques to represent water level fluctuations throughout the growing season. Inter-annual variability was addressed by implementing general linear models (GLMs) to estimate linear and bilinear trends on individual and overall lake area.

To explain and contextualise the Results of the study (Section 5), the discussion of the study examined and interpreted individual and overall trends in lake area across the Tuktoyaktuk Peninsula (Section 7.1). The section explored the effect of weather forcings found on these areal trends, and investigated other non-climatic effects on the spatiotemporal evolution of lakes. Because this study investigated modern polar lacustrine change in a purely geospatial context, and because of the scope this study was limited to, an additional section detailing priorities for further research was included to contextualise this study with other modern issues in the Tuktoyaktuk Peninsula and the Mackenzie Delta region (Section 7.2).

7.1. Analysis of temporal trends.

The results of this study indicate that, even though lakes showed negligible changes in surface area throughout the entire temporal domain of the modern Landsat archive (1985 – 2011), significant lake areal trends were present at the decadal scale (Figure 6.2). Similarly, areal trends did not appear to be homogeneous over the Tuktoyaktuk Peninsula, but instead bifurcated the Peninsula into broad regions of increasing and decreasing areal extents, with local areas of spatial heterogeneity at the finer scale.

7.1.1. Comparison of longer-term and recent trends.

Previous studies have focused on broad spatiotemporal patterns and trends in lake surface area and count in Arctic and sub-Arctic regions (Table 2.1). Thermokarst lakes located within the

zone of continuous permafrost were generally observed to increase both in area and in count, while lakes located within the zones of discontinuous, sporadic, and isolated permafrost decreased in area and number (Smith et al. 2005, Kirpotin et al. 2008, Kirpotin et al. 2009, Carroll et al. 2011). Although the temporal scale of this study was shorter and more recent in comparison, the results obtained are based on considerably more satellite images within data processing steps, and documented a different pattern for thermokarst lake change in a region of continuous permafrost. This study reported negligible changes in the total area of lakes over 1985 to 2011, while observing a unilateral decline in lake number. The discrepancy between areal and numeral trends can primarily attributed to the presence of a significant structural break between the years 1997 and 1998 in overall lake areal trends, which highlights major differences in trends within areal classes.

The temporal structural break between the years 1997 and 1998 divides the time series into an initial positive growth in lake surface area, and a subsequent negative areal decline overall. This structural break is highly significant, and this pattern in lake areal trends has been observed in past literature. In particular, the presence of the break supports the claims of Plug et al. (2008), the only other study to examine lake areal trends over the entire Tuktoyaktuk Peninsula. Although their study identified general increases in lake surface area primarily between 1978 to 1992 and general decreases between 1992 to 2001, only images corresponding to the years 1978 – 1978, 1991 – 1992, and 2000 – 2001 were examined, due to the previous restrictions concerning Landsat imagery availability (as explained in Section 2.4.2). With a total of 352 Landsat images used, this study was considerably higher in temporal resolution, and arguably more accurate in the analysis of time series trends. The inter-annual timing of the structural break determined in this present study is located between the periods of missing data in Plug et al.'s study, and therefore should be interpreted as a temporal refinement of lake areal trends. In addition, this study extended the temporal series of Plug et al.'s study by ten years, and similarly reported no persistent lake changes in overall lake extent on the Tuktoyaktuk Peninsula.

Furthermore, the presence of the structural break is a robust result, which remains applicable when overall trends in lake surface area were partitioned into areal classes. When examining the magnitude of trends within each areal class, it appears that large lakes (Classes 3 and 4) increased relatively more substantially from 1985 to 1997, and decreased from 1998 to 2011 at rates relatively higher and lower than other classes, respectively (Table 6.2; up to 4.693 km^2/year and down to -1.679 km^2/year, respectively). Increases in lake area were concomitant with a

unidirectional decline in the number of lakes, which were largest in magnitude in the smallest areal class (Table 6.3; Class 1; of down to -34.33 lakes/year). A sole analysis of lake abundance without addressing lake area changes over time may be misleading (Jones et al. 2011), as Class 1 lakes have the lowest area-to-count ratio (Table 6.4), and have little effect on the overall changes observed in surface area. Therefore, because large areal classes saw the greatest increases in absolute lake surface area between the years 1985 to 1997, it is likely that these classes masked the reduction in lake count to an extent during this first temporal period. In addition, as the large areal classes hold a relatively high area-to-count ratio (Table 6.4), it is likely that these classes drove the overall changes in lake area throughout the entire temporal domain.

The apparent rapid decline in overall lake area since 1998 in comparison to the rates of change observed over the entire temporal domain may indicate the presence of larger-scale climatic phenomena. A plausible explanation for this negative areal trend after 1998 is a shift towards a stronger Beaufort Sea High during the summer months beginning in the late 1990s (Moore 2012). This particular shift in the intensity and position of the Beaufort Sea High has been associated with a reduction in cyclogenesis over the Beaufort Sea area (Moore 2012). When the cyclone pattern is poorly developed, net precipitation tends to be anomalously negative over the Arctic Ocean (Serreze and Barrett 2008). In addition, a strong Beaufort Sea High has also been associated with positive lower-tropospheric temperature anomalies over the majority of the Arctic Ocean (Serreze and Barrett 2008). The simultaneous combination of positive temperature and negative net precipitation anomalies often results in a major imbalance of the water balance of the landscape, as given by the equation:

$$\Delta S = P - ET - Q - D, \qquad \text{[Eq. 7.1.]}$$

where ΔS represents the change in water stored within a given area, and is determined by precipitation (P), evapotranspiration (ET), streamflow (Q), and groundwater recharge (D). Increased annual mean temperatures are highly correlated with evapotranspiration rates (Gibson et al. 1996), and have been correlated with regional decreases in lake surface area across Alaska (Riordan et al. 2006, Jones et al. 2011). In this study, the yearly growing season precipitation totals were found to also possess a statistically significant break between 1997 and 1998, and to be lower both in trend and in average after 1998 (Figure 6.6D). Although not statistically significant, average temperatures during the growing season appear to be rising in the late 1990s (Figure 6.6A). Because these climatic time series were found to have an overall good fit with the

trends seen in overall lake area during the second temporal period (1998 – 2011), the overall trend of shrinking areal extent during this may have been particularly responsive to negative shifts in water balance due to the combination of temperature and precipitation trends after 1998.

7.1.2. Spatial heterogeneity in lake morphometrics.

Trends in lake surface area over the Tuktoyaktuk Peninsula were found to be regional, which was consistent with prior studies that identified spatial regional heterogeneity in areal trends in Alaska (Riordan et al. 2006, Roach 2011, Rover et al. 2012, Roach et al. 2013) and in Siberia (Smith et al. 2005). Similar with the findings of (Roach 2011), the lack of an association with several climatic variables (Table 6.7) instead suggests that substrate characteristics, topographic features, or other terrestrial, periglacial, and atmospheric climatic gradients are potentially more influential in determining spatial heterogeneity in areal trends.

A spatial visualisation of individual lake surface areal trends over the Tuktoyaktuk Peninsula indicated the presence of a latitudinal divide in net study area trends that bifurcated the Peninsula into two regions of approximately equal area (Figure 6.4). Within these two regions, several hot and cold spots were identified, some that exhibited trends in concordance with their local surroundings, and others that instead represented local heterogeneity in areal trends. The presence of meso-scale and local spatial heterogeneity within the boundaries of a given study area may indicate complex interactions between climate forcing and fine-scale variability in lake hydrological processes, which include factors such as permafrost and thermokarst degradation, the variability of sub-lacustrine substrate permeability, and/or regional and local topography (Roach et al. 2013).

A closer look at the permafrost and ground ice conditions at and around Tuktoyaktuk Peninsula reveals that the types of periglacial environments covering the Peninsula are spatially heterogeneous (Figure 7.1), and that the locations of different permafrost class types closely match the spatial distribution of individual lake surface areal trends (Figure 6.4). Specifically, areas with noticeable ice content at various depths, and silty and clay diamicton deposits form the same spatial patterns with observed regions of lake area increase, and areas with low or no ice content, and coarser sand and gravel sediment types were present where lakes were seen to decrease in area. Furthermore, these areas are spatially coincident with various hot (e.g. Figure 6.5A Boxes 1 and 2; Figure 6.5B Inset; Figure 6.5C Inset) and cold spots (e.g. Figure 6.5A Boxes

3 and 4) that respectively represent localized zones of significant areal increases and decreases. The spatial associations found between lake areal trends and the corresponding local surficial geology supports the studies of Roach (2011), who determined a positive correlation between trends in lake area, soil grain size and type, and local lacustrine water retention. The occurrence of frozen sublacustrine silt is also significantly associated with expanding lakes (Jepsen et al. 2013a). With the unidirectional increases in annual temperature (Figure 6.6A; Figure 6.7A) and the increases in growing season temperature observed after 1998 (Figure 6.6B; Figure 6.7B), it can be hypothesized that the continued warming of the earth surface in regions of declining lake areal trends may promote sub-lacustrine permafrost degradation, increasing the width of the active layer, and connecting the affected lakes to subsurface groundwater networks or overflowing into river systems. Warmer air temperatures have been linked with similar increases in lake temperature (Marsh et al. 2009), and often results in higher rates of lake outflow downwards due to rapid melt of ground ice during drainage (Marsh and Neumann 2001).

The associations found in this study was also consistent with the hypothesis that continued degradation of permafrost in areas with coarse-grained soils may lead to declining lake areas (Yoshikawa and Hinzman 2003, Smith et al. 2005, Roach et al. 2013), with the sole exception that the Tuktoyaktuk Peninsula is located entirely in the continuous permafrost region, and that other observations supporting this hypothesis have all been made in regions of discontinuous permafrost. Therefore, a better understanding of the surficial geology and permafrost-thermokarst interactions in the Tuktoyaktuk Peninsula will enable empirical testing of this hypothesis whether or not trends and correlations in permafrost thaw, soil characteristics, and lake areal change are transferrable across different classifications of permafrost zones.

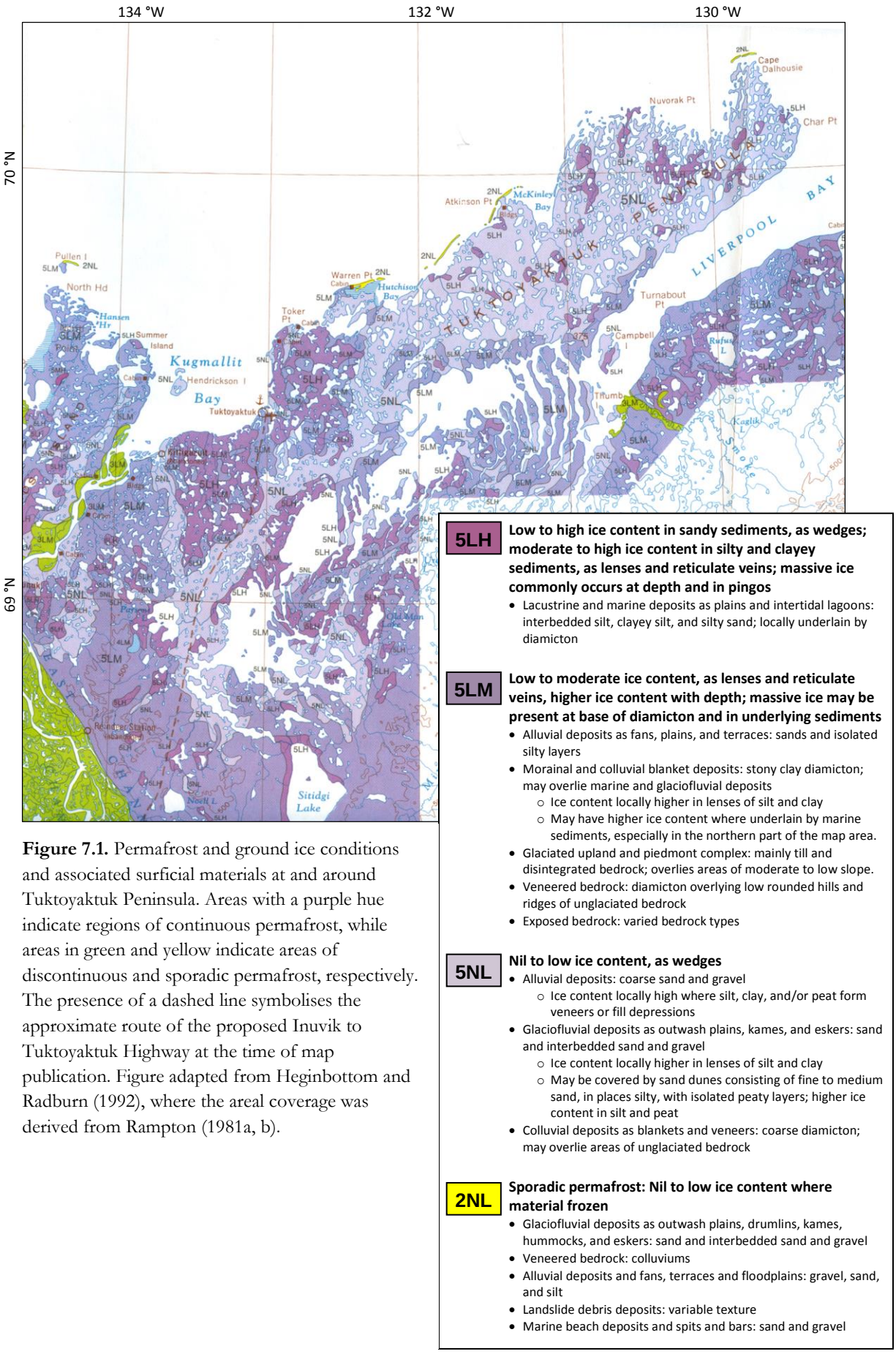


Figure 7.1. Permafrost and ground ice conditions and associated surficial materials at and around Tuktoyaktuk Peninsula. Areas with a purple hue indicate regions of continuous permafrost, while areas in green and yellow indicate areas of discontinuous and sporadic permafrost, respectively. The presence of a dashed line symbolises the approximate route of the proposed Inuvik to Tuktoyaktuk Highway at the time of map publication. Figure adapted from Heginbottom and Radburn (1992), where the areal coverage was derived from Rampton (1981a, b).

The presence of wildfire within the realm of the study area may also influence local and regional lake areal trends. Wildfires result in an immediate impact on the permafrost and ground thermal regimes, increasing soil thermal conductivity and ground heat flux, and decreasing soil moisture retention due to an increase in active layer thickness and the growth of primary and secondary succession (Yoshikawa et al. 2003). The loss of frozen soil, which often serves as an aquiclude for lakes, may promote long-term soil drying (Swanson 1996, Riordan et al. 2006). Very few fires were documented on the Tuktoyaktuk Peninsula due to a combination of the Peninsula being located above the Arctic treeline ecotone and a lack of physical observation at the time of potential fires (EIRB 2011). However, given the available data, all known locations of wildfires and wildfire ignition sites correspond to localised areas of lake surface areal decline during the wildfire occurrence period specified (Figure 7.2), and the majority of these locations coincide with identified cold spots of lake area decrease (Figure 6.5).

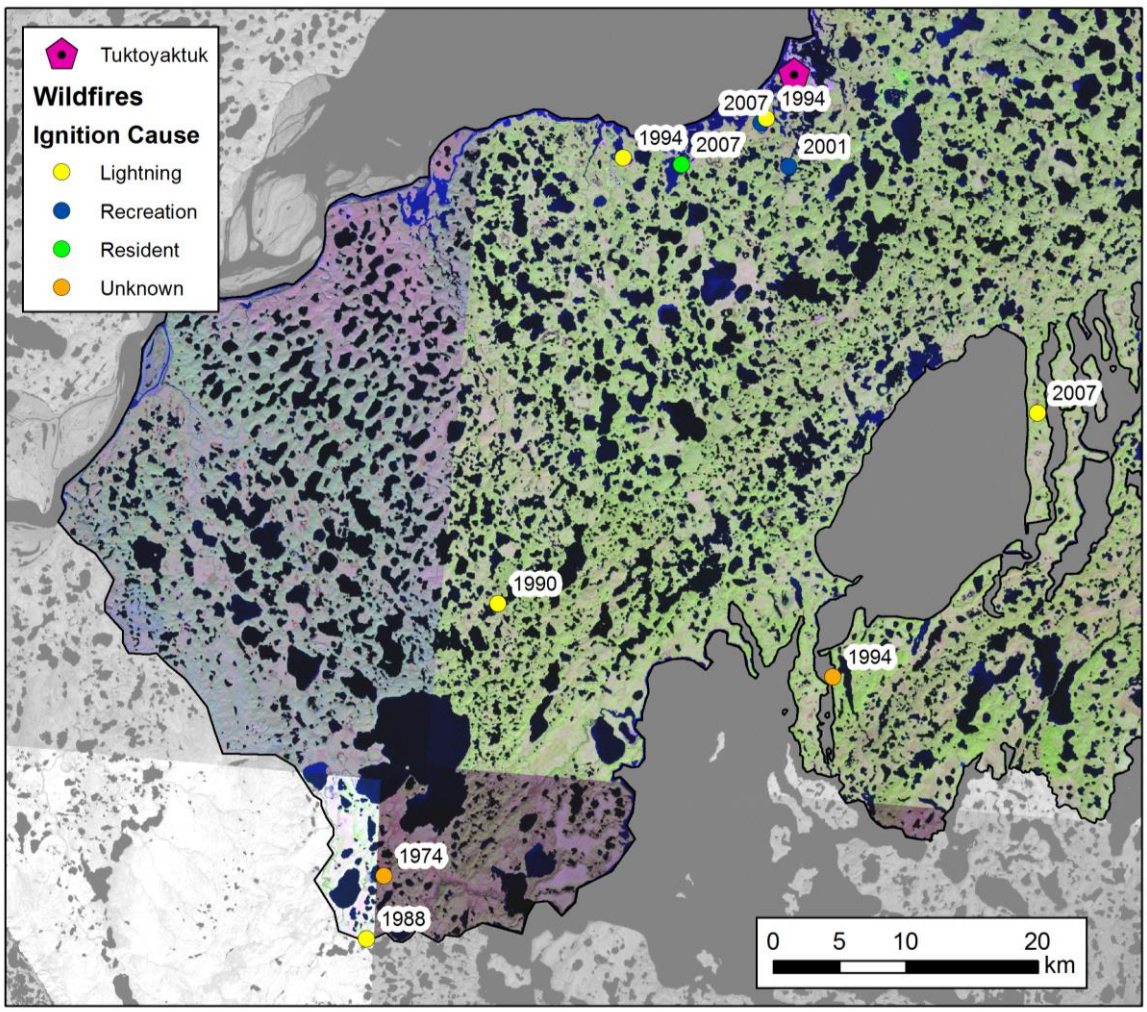


Figure 7.2. Locations of known historical fires and fire ignition causes between Inuvik, N. T., and Tuktoyaktuk, N. T. Note that although there are several points of fire ignition reported, these instances did not all become fires. Fire locations were not known for the region above Tuktoyaktuk, N. T. Adapted from the Environmental Impact Statement for Construction of the Inuvik to Tuktoyaktuk Highway, N. W. T. (EIRB 2011).

The spatial patterns observed examining the effect of precipitation totals on individual lake surface area were less intuitive in comparison. While increased precipitation rates have been correlated with both an increase in lake surface area through the filling of basin catchments (Tarasenko 2013), the same increases have also been associated with a reduction in areal extent by stimulating ground water flow through higher lake levels (Brewer et al. 1993). In this study, an increase in precipitation totals both throughout the year and during the growing season were associated with a reduction in lake surface area in regions of general areal increase (Figure 6.9C and D). It is feasible that increased precipitation could cause diminished lake areas due to higher erosion caused by higher stream level connecting lakes (Hinkel et al. 2007). However, this cannot explain the converse also observed, where reduced amounts of precipitation correlate with increased lake surface area. Several hypotheses can be made to address this discrepancy. First, there may be a significant time lag in the responses of thermokarst lakes to changes in precipitation, due to the delayed response of permafrost to climate forcing (Harris et al. 2009). The coupled thermal and geomorphic processes that operate on thaw lake margins may give rise to a complicated temporal response of lakes to climate change, such as the effect of thaw slump material accumulating along lake margins that buffers permafrost from the effects of increasing waterlevels and temperatures until it is redeposited into the deeper water basin in future years (Plug et al. 2008). Second, precipitation may not be an important factor causing lake areal change. Other studies have made similar claims, due to the absence of lake drainage phenomena and limited new lake formation and lateral lake expansion rates during periods of increased precipitation (Sannel and Kuhry 2011). Third, due to the low number of lakes (47 and 105, respectively) in comparison to the amount of input lakes (1872), there exists the possibility that the results obtained from the GLM are purely coincidental. Last, the ERA-Interim Reanalysis Dataset may not be the best source of precipitation data for localised studies. Because precipitation estimates in the ERA-Interim Dataset were produced by a forecast model based on temperature and humidity (Dee et al. 2011), in addition to known methodological limitations in the ground measurement of precipitation data (explained in Section 5.2), it is difficult to determine true and unbiased totals of precipitation, and future studies should incorporate procedures to correct for precipitation uncertainties.

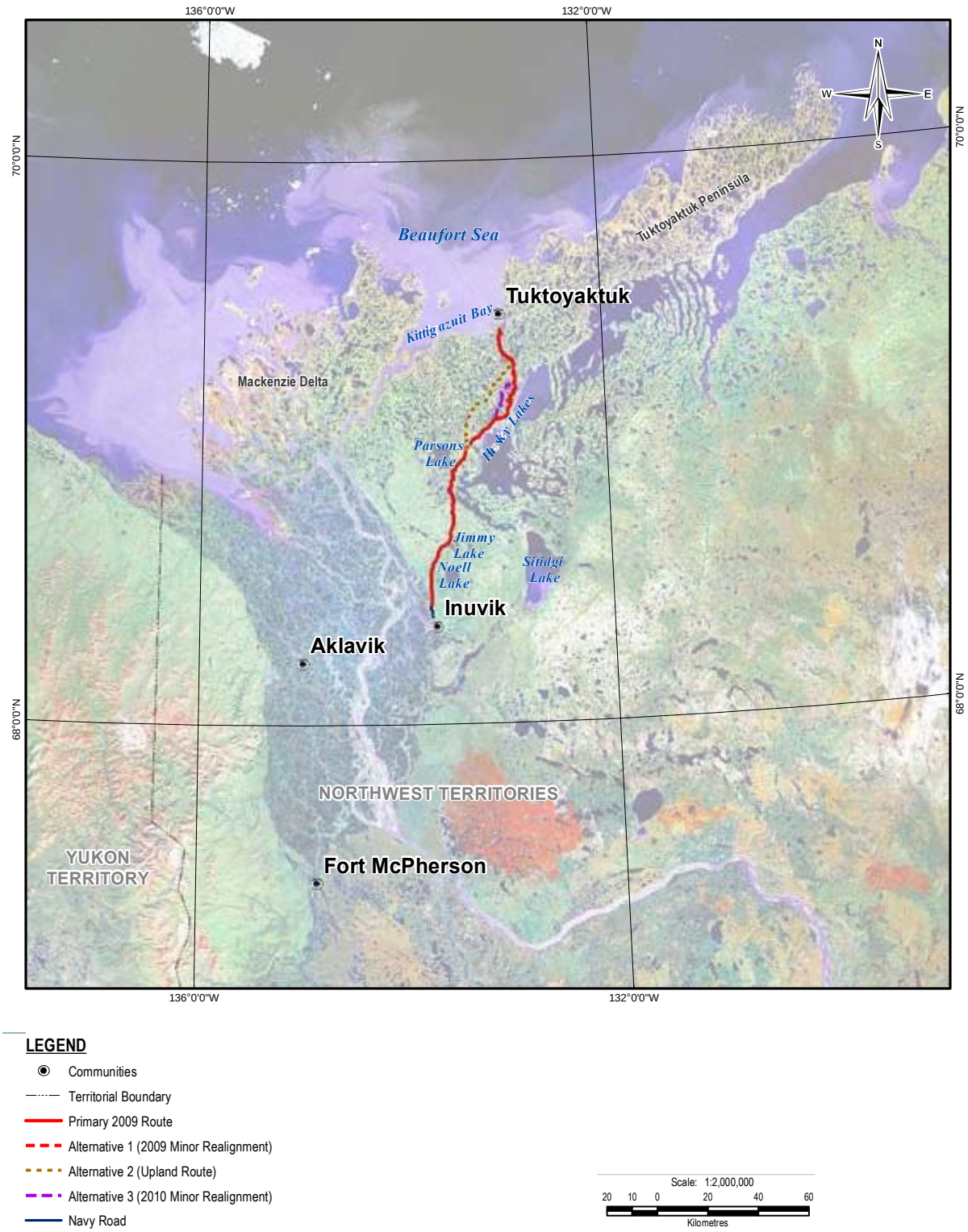


Figure 7.3. Proposed Inuvik-Tuktoyaktuk Highway routes. Adapted from the Environmental Impact Statement for Construction of the Inuvik to Tuktoyaktuk Highway, N. W. T. (EIRB 2011).

7.2. Study implications and priorities for future research.

The results of this study may guide present and future land planning projects in this region. Currently, continued progress is being made at the planning stage for the construction of an all-weather road between the town of Inuvik, N. T., to the hamlet of Tuktoyaktuk, N. T. (EIRB 2011), which is located on the lower half of the northern coastline on the Tuktoyaktuk Peninsula. Although a comprehensive Environmental Impact Statement was recently conducted, the report fails to consider the potential impacts that a changing lacustrine landscape may impact surrounding permafrost and thermokarst conditions. Approximately 80 kilometres of the highway is spatially coincident with identified regions of lake increase, and transcends the Mid-Peninsula Corridor cold spot (Figures 5.4, 5.5, 6.3). Therefore, it would likely be beneficial to understand the local areal dynamics of lakes, particularly in those locations where the highway is to run alongside lakes that have been known to significantly expand and/or shrink over time.

Although historical maps (Figure 7.1; e.g. Rampton 1981a, b, Heginbottom and Radburn 1992) have illustrated the surficial geology and spatial distribution of permafrost conditions of the Tuktoyaktuk Peninsula in the recent past, the instability of ground conditions in the region make consistent and/or complete physical coverage surveys of the Peninsula difficult to achieve. As dramatic changes in the landscape of the Arctic have been observed at an increasing rate, future studies investigating thermokarst lake change in the Tuktoyaktuk Peninsula should involve a field component to measure the present state of permafrost and associated active layer, in the hopes of obtaining more accurate data to successfully correlate lacustrine variability with landscape characteristics.

A better comprehensive knowledge of the ecosystem characteristics, in turn, can potentially quantify explicit relationships between important landscape characteristics and lake areal change. Incorporating these important characteristics in an ecological distribution model, for instance, have been relatively successful at describing the present system by correlative statistics and projecting the system into the future under altered, but monitored conditions (Evans 2012). An improved understanding of the mechanisms underlying lake area trends is imperative to enhance spatiotemporal projections of future lacustrine change by identifying hot and cold spots that are and will be most susceptible to future change (Roach et al. 2013).

Ramifications of the land surface transition due to climate change could potentially trigger changes in the carbon and surface energy budgets of the Arctic (Sturm et al. 2005), as well as alter the habitat and phenology of the system (Jia et al. 2003). Drained thermokarst lake basins have been a significant sink for soil organic carbon over historic and recent geological time scales (Serreze et al. 2000, Schuur et al. 2008). However, lacustrine regions with significant declines in lake area and count, such as the Kuparuk River Basin (Kling et al. 1991) have been identified as recent hotspots of CO_2 and methane export (Zona et al. 2010). Specifically, thermokarst lake basins serve most as a significant CO_2 source within the first 15 years of drainage (Wickland et al. 2009, Jones et al. 2011). Knowing how much carbon will be released from specific permafrost zones during the course of this century is crucial for determining future responses to climate change (Schuur and Abbott 2011). With the recent decreases observed in lake area over the Tuktoyaktuk Peninsula, future research could clarify the net declines in lake area on carbon balance by identifying the size and fraction of the permafrost carbon pool, and the projected increases in methane emissions as a result of carbon release from basin exposure as a result of lake drainage.

8. CONCLUSION.

Changes in thermokarst lakes have been identified across many high-latitude ecosystems (Table 2.1), and these changes are often attributed to consequences from a warming Arctic climate (Roach et al. 2013). This study has examined the dynamics of lake surface area over different spatial and temporal scales in the Tuktoyaktuk Peninsula, N. T., an area located completely within the zone of continuous permafrost. The results of this study supported the findings of past studies in this region (e.g. Plug et al. 2008), while also increasing the temporal resolution to a near-complete yearly time series of homogeneous geospatial imagery, achieving temporal levels that were previously unattainable. The major findings of the study can be separated into five principal components:

1. While no significant net change was found in the total surface area of lakes between 1985 and 2011, the presence of a significant and robust structural temporal break between the years 1997 and 1998 divided the time series into an initial positive growth in lake surface area during the period 1985 to 1997, and a subsequent negative decline in lake area during the period 1998 to 2011.
2. Large lakes saw the greatest changes both in lake areal increase and decrease coincident with the overall observed trends, and so drove the overall change in surface area.
3. The declines seen in lake surface area after 1998 were suggested to be a ramification of a stronger summer Beaufort Sea High that began simultaneously at the end of the old millennium, due to the effects of the pressure system on temperature and precipitation trends around the Tuktoyaktuk Peninsula.
4. Regional differences were observed in the spatial distribution of individual lake area trends. On the broad scale, a latitudinal divide bifurcated the peninsula into two regions of approximately equal area, where the northern region exhibited general trends of lake areal decline, while the southern region exhibited general trends of areal increase. Within these regions, meso- and local hot and cold spots were identified, some that exhibited trends in concordance with local surroundings, while others represented local spatial heterogeneity in areal trends.
5. The spatial trends in lake area seen at multiple scales were suggested to be influenced to a greater extent by characteristics in coincident permafrost and surficial geology, and to a

lesser extent by atmospheric and climatic variables. Most of these features were linked to a warming climate.

In addition, the study demonstrated reasonable success in implementing a largely automated technique on land cover classification using large amounts of Landsat satellite imagery in order to achieve these component goals.

Aside from low-permeability environments and/or beneficial net water balance adjustments, the ultimate effect of continued climate warming on high-latitude permafrost-controlled lakes has continually and consistently been observed (Smith et al. 2005). Annual rates of change of individual lakes within a given region may compound to yield substantial cumulative reductions in overall lake surface area (Roach et al. 2013). As the Arctic continues to warm, a continued observation of thermokarst lake evolution over both broad landscapes and localised regions will be increasingly valuable to future studies that investigate the resulting transformation of the Arctic tundra ecosystem.

9. BIBLIOGRAPHY.

Arctic Climate Impact Assessment [ACIA]. 2005. Pages 183-352 *in* C. Symon, L. Arris, and B. Heal, editors. Arctic Climate Impact Assessment - Scientific report. Cambridge University Press, Cambridge, UK.

Anema, C., R. E. Hecky, E. J. Fee, D. Nernberg, and S. J. Guildford. 1990. Water chemistry of some lakes and channels in the Mackenzie Delta and on the Tuktoyaktuk Peninsula, NWT, 1985.*in* Canadian Data Report of Fisheries and Aquatic Sciences, Department of Fisheries and Oceans. No. 726, Winnipeg, MB.

Anisimov, O. A. and S. Reneva. 2006. Permafrost and changing climate: the Russian perspective. Ambio **35**:169-175.

Anisimov, O. A., N. I. Shiklomanov, and F. E. Nelson. 1997. Global warming and active-layer thickness: results from transient general circulation models. Global and Planetary Change **15**:61-77.

Arp, C. D., B. M. Jones, F. E. Urban, and G. Grosse. 2011. Hydrogeomorphic processes of thermokarst lakes with grounded-ice and floating-ice regimes on the Arctic Coastal Plain, Alaska. Hydrological Processes **25**:2422-2438.

Barber, V. A., G. P. Juday, and B. P. Finney. 2000. Reduced growth of Alaskan white spruce in the twentieth century from temperature-induced drought stress. Nature **405**:668-673.

Bigras, S. C. 1990. Hydrological regime of lakes in the Mackenzie delta, Northwest Territories, Canada. Arctic and Alpine Research **22**:163-174.

Bijker, W., N. A. S. Hamm, J. Ijumulana, and M. K. Wole. 2011. Monitoring a fuzzy object: The case of Lake Naivasha. 6th International Workshop on the Analysis of Multi-temporal Remote Sensing Images (Multi-Temp):153-156.

Boike, J., K. Roth, and P. P. Overdoin. 1998. Thermal and hydrologica dynamics of the active layer at a continuous permafrost site (Taymyr Peninsula, Siberia). Water Resources Research **34**:355-363.

Bolstad, P. V., P. Gessler, and T. M. Lillesand. 1990. Uncertainty in manually digitized map data. International Journal of Geographic Information Systems **4**.

Bowling, L. C. and D. P. Lettenmaier. 2010. Modeling the effects of lakes and wetlands on the water balance of the Arctic Environments. Journal of Hydrometeorology **11**:276-295.

Brewer, M. C., L. D. Carter, and R. Glenn. 1993. Sudden drainage of a thaw lake on the Alaskan Arctic Coastal Plain. Pages 48-53 Proceedings of the Sixth International Conference on Permafrost, Beijing, China, 5-9 July 1993. South China University of Technology Press, Beijing, China.

Brown, J., O. J. Ferrians Jr., J. A. Heginbottom, and E. S. Melnikov. 1997. Circum-Arctic Map of Permafrost and Ground-Ice Conditions. U. S. Geological Survey, Washington, D. C. .

Brown, J., K. M. Hinkel, and F. E. Nelson. 2000. The Circumpolar Active Layer Monitoring (CALM) program: Research designs and initial results. Polar Geography **24**:165-258.

Buckeridge, K. M. and P. Grogan. 2008. Deepened snow alters soil microbial nutrient limitations in arctic birch hummock tundra. Applied Soil Ecology **39**:210-222.

Burn, C. R. 2002. Tundra lakes and permafrost, Richards Island, western Arctic coast, Canada. Canadian Journal of Earth Sciences **39**:1281-1298.

Burn, C. R. 2003. Lake-bottom thermal regime in thermokarst terrain near Mayo, Yukon Territory, Canada.*in* M. Phillips, S. M. Springman, and L. U. Arenson, editors. Proceedings of the Eigth International Conference on Permafrost, Zürich, Switzerland, 21-25 July 2003. Swets & Zeitlinger, Lisse, the Netherlands.

Burn, C. R. 2005. Lake-bottom thermal regimes, western Arctic coast, Canada. Permafrost and Periglacial Processes **16**:355-367.

Carroll, M. L., J. R. G. Townshend, C. M. DiMiceli, T. Loboda, and R. A. Sohlberg. 2011. Shrinking lakes of the Arctic: Spatial relationships and trajectory of change. Geophysical Research Letters **38**:L20406.

CAVM. 2003. Circumpolar Arctic Vegetation Map (1:7,500,000 scale), Conservation of Arctic Flora and Fauna (CAFF) Map No. 1.*in* U. S. F. a. W. Service, editor., Anchorage, AK.

Chander, G., B. L. Markham, and D. L. Helder. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Remote Sensing of Environment **113**:893-903.

Chapman, W. L. and J. E. Walsh. 1993. Recent variations of sea ice and air temperature in high latitudes. Bulletin of the American Meteorological Society **74**:33-47.

Chow, G. C. 1960. Tests of equality between sets of coefficients in 2 linear regressions. Econometrica **28**:591-605.

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

Côté, M. M. and C. R. Burn. 2002. The oriented lakes of Tuktoyaktuk Peninsula, Western Arctic Coast, Canada: a GIS-based analysis. Permafrost and Periglacial Processes **13**:61-70.

Davis, N. 2003. Permafrost: A Guide to Frozen Ground in Transition. University of Alaska Press, Fairbanks, AK.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Galsamo, and P. Pauer. 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society **137**:553-597.

Desyatkin, A. R., F. Takakai, P. P. Fedorov, M. C. Nikolaeva, R. V. Desyatkin, and R. Hatano. 2009. CH_4 emission from different stages of thermokarst formation in Central Yakutia, East Siberia. Soil Science and Plant Nutrition **55**:558-570.

Dobinski, W. 2011. Permafrost. Earth-Science Reviews **108**:158-169.

Duguay, C. R., Y. Ernou, and J. Hawkings. 1999. SAR and optical satellite observations of ice covered thermokarst lakes, Old Crow Flats, Yukon Territory. 56th Eastern Snow Conference, University of Laval, Fredericton, N. B., Canada.

Environmental Impact Research Board [EIRB]. 2011. Environmental impact statement for construction of the Inuvik to Tuktoyaktuk Highway, NWT. Kiggiak - EBA Consulting Ltd., Inuvik, NT.

Environment Canada. 2013. Canadian Climate Normals or Averages, 1971 - 2000, http://www.climate.weatheroffice.ec.gc.ca/. Ottawa, ON, Canada.

Evans, M. R. 2012. Modelling ecological systems in a changing world. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences **367**:181-190.

Fisher, P. F. 2010. Remote sensing of land cover classes as type 2 fuzzy sets. Remote Sensing of Environment **114**:309-321.

Foody, G. M. 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment **80**:185-2001.

Foody, G. M. and P. M. Atkinson. 2002. Uncertainty in Remote Sensing and GIS. John Wiley and Sons, Ltd., Chichester, UK.

French, H. M. 2007. The Periglacial Environment. 3rd edition. John Wiley and Sons Ltd., Chichester, UK.

Getis, A. and J. K. Ord. 1992. The analysis of spatial association by use of distance statistics. Geographical Analysis **24**:189-206.

Gibson, J. J. and T. W. D. Edwards. 2002. Regional surface water balance and evaporation-transpiration partitioning from a stable isotope survey of lakes in northern Canada. Global Biogeochemical Cycles **16**:10.

Gibson, J. J., T. D. Prowse, and T. W. D. Edwards. 1996. Evaporation of a small lake in the continental Arctic using multiple methods. Nordic Hydrology **27**:1-24.

Gómez-Chova, L., G. Camps-Valls, J. Calpe-Maravilla, L. Guanter, and J. Moreno. 2007. Cloud-screening algorithm for ENVISAT/MERIS multispectral images. IEEE Transactions on Geoscience and Remote Sensing **45**:4105-4118.

Goodwin, N. R., L. J. Collett, R. J. Denham, and N. Flood. 2013. Cloud and cloud shadow screening across Queensland, Australia: An automated method for Landsat TM/ETM+ time series. Remote Sensing of Environment **134**:50-65.

Gouttevin, I., M. Menegoz, F. Dominé, G. Krinner, C. Koven, P. Ciais, C. Tarnocai, and J. Boike. 2012. How the insulating properties of snow affect soil carbon distribution in the continental pan-Arctic area. Journal of Geophysical Research **117**:G02020.

Grosse, G., B. M. Jones, and C. D. Arp. 2013. Thermokarst lakes, drainage, and drained basins. Treatise on Geomorphology **8**:325-353.

Grosse, G., L. Schirrmeister, and T. J. Malthus. 2006. Application of Landsat-7 satellite data and a DEM for the quantification of thermokarst-affected terrain types in the periglacial Lena-Anabar coastal lowland. Polar Research **25**:51-67.

Guerschman, J. P., J. M. Paruelo, C. Di Bella, M. C. Giallorenzi, and F. Pacín. 2003. Cloud and cloud shadow screening across Queensland, Australia: An automated method for Landsat TM/ETM+ time series. International Journal of Remote Sensing **24**:3381-3402.

Hall, D. K., G. A. Riggs, and V. V. Salomonson. 2001. Algorithm theoretical basis document (ATBD) for the MODIS snow and sea ice-mapping algorithms. Hydrological Sciences Branch, NASA/Goddard Space Flight Center, Greenbelt, MD.

Harris, C., L. U. Arenson, H. H. Christiansen, B. Etzelmüller, R. Frauenfelder, S. Gruber, W. Haeberli, C. Hauck, M. Hölzle, and O. Humlum. 2009. Permafrost and climate in Europe: Monitoring and modelling thermal, geomorphological and geotechnical responses. Earth-Science Reviews **92**:117-171.

Heginbottom, J. A. and L. K. Radburn. 1992. Permafrost and ground conditions of northwestern Canada. Geological Survey oof Canada. Map 1691A (2 sheets, scale 1:1,000,000).

Hinkel, K. M., B. M. Jones, W. R. Eisner, C. J. Cuomo, R. A. Beck, and R. Frohn. 2007. Methods to assess natural and anthropogenic thaw lake drainage on the western Arctic coastal plain of northern Alaska. Journal of Geophysical Research **112**:F02S16.

Hinkel, K. M., Y. Sheng, J. D. Lenters, E. A. Lyons, R. A. Beck, W. R. Eisner, and J. Wang. 2012. Thermokarst lakes on the Arctic Coastal Plain of Alaska: geomorphic controls on bathymetry. Permafrost and Periglacial Processes **23**:218-230.

Hudak, A. T. and B. H. Brockett. 2004. Mapping fire scars in a southern African savannah using Landsat imagery. International Journal of Remote Sensing **25**:3231-3243.

Irish, R. R. 2000. Landsat 7 automatic cloud cover assessment algorithms for multispectral, hyperspectral, and ultraspectral imagery. The International Society for Optical Engineering **4049**:348-355.

Jepsen, S. M., C. I. Voss, M. A. Walvoord, B. J. Minsley, and J. Rover. 2013a. 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., C. I. Voss, M. A. Walvoord, J. R. Rose, B. J. Minsley, and B. D. Smith. 2013b. Sensitivity analysis of lake mass balance in discontinuous permafrost: the example of disappearing Twelvemile Lake, Yukon Flats, Alaska (USA). Hydrologeology Journal **21**:185-200.

Jia, G. J., H. E. Epstein, and D. A. Walker. 2003. Greening of arctic Alaska, 1981-2001. Geophysical Research Letters **30**:2067.

Jones, B. M., G. Grosse, C. D. Arp, M. C. Jones, K. M. Walter Anthony, and V. E. Romanovsky. 2011. Modern thermokarst lake dynamics in the continuous permafrost zone, northern Seward Peninsula, Alaska. Journal of Geophysical Research **2011**:G00M03.

Jorgenson, M. T. and T. Osterkamp. 2005. Response of boreal ecosystems to varying modes of permafrost degradation. Canadian Journal of Forestry Resources **35**:2100-2111.

Jorgenson, M. T., V. Romanovsky, J. Harden, Y. L. Shur, J. O'Donnell, E. A. G. Schuur, M. Z. Kanevskiy, and S. Marchenko. 2010. Resilience and vulnerability of permafrost to climate change. Canadian Journal of Forest Research **40**:1219-1236.

Ju, J. and D. P. Roy. 2008. The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally. Remote Sensing of Environment **112**:1196-1211.

Katamura, F., M. Fukuda, N. P. Bosikov, R. V. Desyatkin, T. Nakamura, and J. Moriizumi. 2007. Thermokarst formation and vegetation dynamics inferred from a palynological study in Central Yakutia, Eastern Siberia, Russia. Arctic, Antarctic, and Alpine Research **38**:561-570.

Kaufman, Y. J. 1987. The effect of subpixel clouds on remote sensing. International Journal of Remote Sensing **8**:839-857.

Kirpotin, S., A. Berezin, V. Bazanov, Y. Polishchuk, S. Vorobiov, N. Mironycheva-Tokoreva, N. Kosykh, I. Volkova, B. Dupre, O. Pokrovsky, A. V. Kouraev, E. Zakharova, L. Shirokova, N. Mognard, S. Biancamaria, J. Viers, and M. Kolmakova. 2009. Western Siberia wetlands as indicator and regulator of climate change on the global scale. International Journal of Environmental Studies **66**:409-421.

Kirpotin, S., Y. Polishchuck, E. Zakharova, L. Shirokova, O. Pokrovsky, M. Kolmakova, and B. Dupre. 2008. One of possible mechanisms of thermokarst lake drainage in West-Siberian North. International Journal of Environmental Studies **65**:631-635.

Kling, G. W., G. W. Kipphut, and M. C. Miller. 1991. Arctic lakes and streams as gas conduits to the atmosphere: implications for tundra carbon budgets. Science **251**:298-301.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel. 2006. World map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift **15**:259-263.

Kozlenko, N. and M. O. Jeffries. 2000. Bathymetric mapping of shallow water in thaw lakes on the North Slope of Alaska with spaceborne imaging radar. Arctic **53**:306-316.

Kravtsova, V. I. and A. G. Bystrova. 2009. Changes in thermokarst lake sizes in different regions of Russia for the last 30 yr. Kriosfera Zemli (Earth and Cryosphere) **13**:16-26.

Kravtsova, V. I. and T. V. Tarasenko. 2010. Studying and mapping of the dynamics of thermokarst lakes in Western Siberia by space images taken in different times. Dinamika okruzhayushchei sredy i global'nye izmeneniya klimata (Dynamics of the Environment and Global Climate Changes) **1**:88-93.

Labrecque, S., D. Lacelle, C. R. Duguay, B. Lauriol, and J. Hawkings. 2009. Contemporary (1951-2001) evolution of lakes in the Old Crow Basin, Northern Yukon, Canada: remote sensing, numerical modeling, and stable isotope analysis. Arctic **62**:225-238.

Lachenbruch, A. H., M. C. Brewer, G. W. Greene, and B. V. Marshall. 1962. Temperatures in permafrost. Pages 791-803 *in* C. M. Herzfeld, editor. Temperature: Its Measurement and Control in Science and Industry. Reinhold, New York, NY.

Lachenbruch, A. H. and B. V. Marshall. 1986. Changing climate - geothermal evidence from permafrost in the Alaskan Arctic. Science **234**:689-696.

Legates, D. R. and C. J. Willmott. 1990. Mean seasonal and spatial variability in global surface air temperature. Theoretical and Applied Climatology **41**:11-21.

Luo, Y., A. P. Trishchenko, and K. V. Khlopenkov. 2008. Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land band over Canada and North America. Remote Sensing of Environment **112**:4167-4185.

Lyapustin, A., Y. Wang, and R. Frey. 2008. An automatic cloud mask algorithm based on time series of MODIS measurements. Journal of Geophysical Research **113**:D16207.

Mackay, J. R. 1971. The origin of massive icy beds in permafrost, Western Arctic Coast, Canada. Journal of Earth Sciences **8**:397-422.

Mackay, J. R. 1979. Pingos of the Tuktoyaktuk Peninsula area, Northwest Territories. Géographie physique et Quaternaire **33**:3-61.

Mackay, J. R. 1981. An experiment in lake drainage, Richards Island, Northwest Territories: a progress report. Current Research, Part A. Geological Survey of Canada, Paper 81-1A.

Mackay, J. R. 1988. Catastrophic lake drainage, Tuktoyaktuk Peninsula area, District of Mackenzie. Pages 83-90 *in* Current Research, Part D, Geological Survey of Canada, Paper 88-1D.

Mackay, J. R. 1997. A full-scale field experiment (1978-1995) on the growth of permafrost by means of lake drainage, western Arctic coast: a discussion of the method and some results. Canadian Journal of Earth Sciences **34**:17-33.

Maling, D. H. 1989. Measurements from Maps: Principles and Methods of Cartometry. Pergamon Press, Oxford, UK.

Marsh, P. and S. C. Bigras. 1988. Evaporation from Mackenzie delta lakes, NWT, Canada. Arctic and Alpine Research **20**:220-229.

Marsh, P. and N. N. Neumann. 2001. Processes controlling the rapid drainage of two ice-rich permafrost-dammed lakes in NW Canada. Hydrological Processes **15**:3433-3446.

Marsh, P., M. Russell, C. Onclin, and H. Haywood. 2008. Modeling discharge during the rapid drainage of thaw lakes in the western Canadian Arctic.*in* D. L. Kane and K. M. Hinkel, editors. Ninth International Conference on Permafrost. Institute of Northern Engineering, University of Alaska Fairbanks, Fairbanks, AK.

Marsh, P., M. Russell, S. Pohl, H. Haywood, and C. Onclin. 2009. Changes in thaw lake drainage in the Western Canadian Arctic from 1950 to 2000. Hydrological Processes **23**:145-158.

Martinuzzi, S., W. A. Gould, and O. M. Ramos Gonzáles. 2007. Creating cloud-free Landsat ETM+ data sets in tropical landscapes: cloud and cloud-shadow removal. Page 12, U.S. Department of Agriculture, International Institute of Tropical Forestry. , Rio Piedras, PR.

Masek, J. G., E. F. Vermote, N. E. Saleous, R. Wolfe, F. G. Hall, K. F. Huemmrich, F. Gao, J. Kutler, and T.-K. Lim. 2006. A Landsat surface reflectance dataset for North America, 1990 - 2000. IEEE Geoscience and Remote Sensing Letters **3**:68-72.

Melaas, E. K., M. A. Friedl, and Z. Zhu. 2013. Detecting interannual variation in deciduous broadleaf forest phenology using Landsat TM/ETM+ data. Remote Sensing of Environment **132**.

Moore, G. W. K. 2012. Decadal variability and a recent amplification of the summer Beaufort Sea High. Geophysical Research Letters **39**:L10807.

Moran, P. A. P. 1950. Notes on continuous stochastic phenomena. Biometrika **37**:17-23.

Mugford, R. I., P. Christoffersen, J. A. Dowdeswell, and A. Consonni. 2012. Evaluation of the ERA-Interim Reanalysis for modelling permafrost on the North Slope of Alaska. North Press, Tyumen, Russia.

Murton, J. B. 1996. Thermokarst-lake-basin sediments, Tuktoyaktuk Coastlands, Western Arctic Canada. Sedimentology **43**:737-760.

North American Cartographical Information Society [NACIS]. 2013. Physical vectors representing single-line drainages including optional lake centerlines and supplementary data for North America and Europe at the 1:10m scale, http://www.naturalearthdata.com/. Milwaukee, WI.

Nicol, S., J. K. Roach, and B. Griffith. 2013. Spatial heterogeneity in statistical power to detect changes in lake area in Alaskan National Wildlife Refuges. Landscape Ecology **28**:507-517.

Osterkamp, T. E. 2007. Characteristics of the recent warming of permafrost in Alaska. Journal of Geophysical Research **112**:F02S02.

Osterkamp, T. E. and J. C. Jorgenson. 2006. Warming of permafrost in the Arctic National Wildlife Refuge, Alaska. Permafrost and Periglacial Processes **17**:65-69.

Osterkamp, T. E., M. T. Jorgenson, E. A. G. Schuur, Y. L. Shur, M. Z. Kanevskiy, J. G. Vogel, and V. E. Tumskoy. 2009. Physical and ecological changes associated with warming permafrost and thermokarst in Interior Alaska. Permafrost and Periglacial Processes **20**:235-256.

Osterkamp, T. E. and V. E. Romanovsky. 1999. Evidence for warming and thawing of discontinuous permafrost in Alaska. Permafrost and Periglacial Processes **10**:17-37.

Osterkamp, T. E., L. Viereck, Y. L. Shur, M. T. Jorgenson, C. Racine, A. Doyle, and R. D. Boone. 2000. Observations of thermokarst and its impact on boreal forests in Alaska. Arctic, Antarctic, and Alpine Research **32**:303-315.

Payette, S. 2004. Accelerated thawing of subarctic peatland permafrost over the last 50 years. Geophysical Research Letters **31**:L18208.

Pienitz, R., J. P. Smol, and D. R. S. Lean. 1997. Physical and chemical limnology of 59 lakes located between the southern Yukon and the Tuktoyaktuk Peninsula, Northwest Territories (Canada). Canadian Journal of Fisheries and Aquatic Sciences **54**:330-346.

Pinard, J.-P. 2011. Wind Monitoring Update for Tuktoyaktuk, Winter 2009. Aurora Research Institute, Aurora College, Inuvik, NT.

Plug, L. J., C. Walls, and B. M. Scott. 2008. Tundra lake changes from 1978 to 2001 on the Tuktoyaktuk Peninsula, western Canadian Arctic. Geophysical Research Letters **35**:L03502.

Pollard, W. H. and H. M. French. 1980. A first approximation of the volume of ground ice, Richards Island, Pleistocene Mackenzie Delta, N.W.T., Canada. Canadian Geotechnology Journal **17**:509-516.

Rampton, V. N. 1981a. Surficial geology of Mackenzie Delta, District of Mackenzie, Northwest Territories. Map 30-1979, scale 1:250,000. Geological Survey of Canada.

Rampton, V. N. 1981b. Surficial geology of Stanton, District of Mackenzie, Northwest Territories. Map 33-1979, scale 1:250,000. Geological Survey of Canada.

Rampton, V. N. and M. Bouchard. 1975. Surficial geology of Tuktoyaktuk, District of Mackenzie. Geological Survey of Canada **Paper 74-53**:18.

Riordan, B., D. Verbyla, and A. D. McGuire. 2006. Shrinking ponds in subarctic Alaska based on 1950-2002 remotely sensed images. Journal of Geophysical Research **111**:G04002.

Riseborough, D. 2006. Discussion of C.R. Burn's 'Lake-bottom thermal regimes, western Arctic coast, Canada'. Permafrost and Periglacial Processes **17**:87-89.

Roach, J. K. 2011. Lake area change in Alaskan National Wildlife Refuges: magnitude, mechanisms, and heterogeneity. PhD Thesis, University of Alaska Fairbanks, Fairbanks, AK.

Roach, J. K., B. Griffith, and D. Verbyla. 2012. Comparison of three methods for estimating number and area of boreal lakes from TM/ETM+ Landsat imagery. Canadian Journal of Remote Sensing **38**:427-440.

Roach, J. K., B. Griffith, and D. Verbyla. 2013. Landscape influences on climate-related lake shrinkage at high latitudes. Global Change Biology **19**:2276-2284.

Roach, J. K., B. Griffith, D. Verbyla, and J. Jones. 2011. Mechanisms influencing changes in lake area in Alaskan boreal forest. Global Change Biology **17**:2567-2583.

Rover, J., L. Ji, B. K. Wylie, and L. L. Tieszen. 2012. Establishing water body areal extent trends in interior Alaska from multi-temporal Landsat data. Remote Sensing Letters **3**:595-604.

Rozenstein, O. and A. Karnieli. 2011. Comparison of methods for land-use classification incorporating remote sensing and GIS inputs. Applied Geography **31**:533-544.

Runyan, C. W. and P. O'Dorico. 2012. Ecohydroogical feedbacks between permafrost and vegetation dynamics. Advances in Water Resources **49**:1-12.

Sannel, A. B. K. and P. Kuhry. 2011. Warming-induced destabilization of peat plateau/thermokarst lake complexes. Journal of Geophysical Research **116**:G03035.

Sano, E. E., L. G. Ferreira, G. P. Asner, and E. T. Steinke. 2007. Spatial and temporal probabilities of obtaining cloud-free Landsat images over the Brazilian tropical savanna. International Journal of Remote Sensing **28**:2739-2752.

Schuur, E. A. G. and B. Abbott. 2011. Climate change: High risk of permafrost thaw. Nature **480**:32-33.

Schuur, E. A. G., J. Bockheim, J. G. Canadell, E. S. Euskirchen, C. B. Field, S. V. Goryachkin, S. Hagemann, P. Kuhry, H. L. Lafleur, G. Mazhitova, F. E. Nelson, A. Rinke, V. E. Romanovsky, N. I. Shiklomanov, C. Tarnocai, S. Venevsky, J. G. Vogel, and S. A. Zimov. 2008. Vulnerability of permafrost carbon to climate change: implications for the global carbon cycle. Bioscience **58**:701-714.

Schuur, E. A. G., K. G. Crummer, J. G. Vogel, and M. C. Mack. 2007. Plant species composition and productivity following permafrost thaw and thermokarst in Alaskan tundra. Ecosystems **10**:280-292.

Sellmann, P. V., J. Brown, R. I. Lewellen, H. L. McKim, and C. J. Merry. 1975. The classification and geomorphic implications of thaw lakes on the Arctic Coastal Plain, Alaska. Hanover, NH, USA.

Serreze, M. C. and A. P. Barrett. 2008. The summer cyclone maximum over the Central Arctic Ocean. Journal of Climate **21**:1048-1065.

Serreze, M. C., J. E. Walsh, F. S. Chapin, T. E. Osterkamp, M. Dyurgerov, V. Romanovsky, W. C. Oechel, J. Morison, T. Zhang, and R. G. Barry. 2000. Observational evidence of recent change in the northern high-latitude environment. Climactic Change **46**:159-207.

Sharma, C. S., M. D. Behera, A. Mishra, and S. N. Panda. 2011. Assessing flood induced land-cover changes using remote sensing and fuzzy approach in Eastern Gujarat (India). Water Resources Management **25**:3219-3246.

Shur, Y. L. and M. T. Jorgenson. 2007. Patterns of permafrost formation and degradation in relation to cliamte and ecosystems. Permafrost and Periglacial Processes **18**:7-19.

Smith, G. I. 1976. Permafrost in the Mackenzie delta, Northwest Territories. Pages Paper 75-28, Geological Survey of Canada.

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

Smith, M. W. and D. Riseborough. 1996. Ground temperature monitoring and detection of climate change. Permafrost and Periglacial Processes **7**:301-310.

Smits, P. C., S. G. Dellepiane, and R. A. Schowengerdt. 1999. Quality assessment of image classification algorithms for land-cover mapping: a review and proposal for a cost-based approach. International Journal of Remote Sensing **20**:1461-1486.

Smol, J. P. and M. S. V. Douglas. 2007. Crossing the final ecological threshold in high Arctic ponds. Proceedings of the National Academy of Sciences of the United States of America **104**:12395-12397.

Soille, P. 1999. Morphological image analysis: Principles and applications. Springer-Verlag Telos, Berlin, Germany.

Soille, P., J. Vogt, and R. Columbo. 2003. Carving and adapting drainage enforcement of grid digital elevation models. Water Resources Research **12**:1-13.

Storey, J., P. Scaramuzza, and G. A. Schmidt. 2005. Landsat 7 Scan Line Corrector-Off Gap-Filled Product Development. Pecora 16 "Global Priorities in Land Remote Sensing". American Society for Photogrammetry and Remote Sensing, Sioux Falls, SD.

Sturm, M., J. Holmgren, J. P. McFadden, G. E. Liston, F. S. Chapin III, and C. H. Racine. 2001. Snow-shrub interactions in Arctic tundra: a hypothesis with climatic implications. Journal of Climate **14**:336-344.

Sturm, M., J. P. Schimel, G. Michaelson, J. M. Welker, S. F. Oberbauer, G. E. Liston, J. Fahnestock, and V. E. Romanovsky. 2005. Winter biological processes could help convert Arctic tundra to shrubland. Bioscience **55**:17-26.

Swanson, D. K. 1996. Susceptibility of permafrost soils to deep thaw after forest fires in interior Alaska, USA, and some ecologic implications. Arctic and Alpine Research **28**:217-227.

Tarasenko, T. V. 2013. Interannual variations in the areas of thermokarst lakes in Central Yakutia. Water Resources **40**:111-119.

Toniolo, H., P. Kodial, L. D. Hinzman, and K. Yoshikawa. 2009. Spatio-temporal evolution of a thermokarst in Interior Alaska. Cold Regions Science and Technology **56**:39-49.

Townshend, J. R. G., C. Huang, S. N. V. Kalluri, R. S. Defries, S. Liang, and K. Yang. 2000. Beware of per-pixel characterization of land cover. International Journal of Remote Sensing **21**:839-843.

Tsoukalas, L. H. and R. E. Uhrig. 1997. Fuzzy and Neural Approaches in Engineering. John Wiley and Sons, Inc., New York, NY.

Uppala, S. M., P. W. Kållberg, A. J. Simmons, U. Andrae, V. Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, and G. A. Kelly. 2005. The ERA-40 re-analysis. Quarterly Journal of the Royal Meteorological Society **131**:2961-3012.

Vermote, E. F. and N. E. Saleous. 2007. LEDAPS surface reflectance product description, Version 2.0.

Vitt, D. H., L. A. Halsey, and S. C. Zoltai. 2000. The changing landscape of Canada's western boreal forest: the current dynamics of permafrost. Canadian Journal of Forest Research **30**:283-287.

Walsh, J. E., V. Kattsov, D. Portis, and V. Meleshko. 1998. Arctic precipitation and evaporation: Model results and observational estimates. Journal of Climate **11**:72-87.

Wang, F. 1990. Fuzzy supervised classification of remote sensing images. IEEE Transactions on Geoscience and Remote Sensing **28**:194-201.

Wickland, K., J. A. O'Donnell, J. C. Koch, M. T. Jorgenson, J. W. Harden, R. G. Striegl, and L. Anderson. 2009. Fate of carbon in sediments of a drying high latitude lake, interior Alaska. EOS, Transactions American Geophysical Union **90**:U41C-0057.

Woodcock, C. E., R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. Cohen, F. Gao, S. N. Goward, D. Helder, E. Helmer, R. R. Nemani, L. Oreopoulos, J. Schott, P. S. Thenkabail, E. F. Vermote, J. Vogelmann, M. A. Wulder, and R. Wynne. 2008. Free access to Landsat imagery. Science **320**:1011.

Wulder, M. A., J. G. Masek, W. B. Cohen, T. R. Loveland, and C. E. Woodcock. 2012. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. Remote Sensing of Environment **122**:2-10.

Yang, D., D. L. Kane, Z. Zhang, D. R. Legates, and B. Goodison. 2005. Bias correction of long-term (1973-2004) daily precipitation data over the northern regions. Geophysical Research Letters **32**:L19501.

Yoshikawa, K., W. R. Bolton, V. E. Romanovsky, M. Fukuda, and L. D. Hinzman. 2003. Impacts of wildfire on the permafrost in the boreal forests of Interior Alaska. Journal of Geophysical Research **108**:8148.

Yoshikawa, K. and L. D. Hinzman. 2003. Shrinking thermokarst ponds and groundwater dynamics in discontinuous permafrost. Permafrost and Periglacial Processes **14**:151-160.

Zhang, T., R. G. Barry, K. Knowles, J. A. Heginbottom, and J. Brown. 1999. Statistics and characteristics of permafrost and ground-ice distribution in the Northern Hemisphere. Polar Geography **23**:132-154.

Zhang, Y., B. Guindon, and J. Cihlar. 2002. An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images. Remote Sensing of Environment **82**:173-187.

Zhu, Z. and C. E. Woodcock. 2012. Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sensing of Environment **118**:83-94.

Zhu, Z., C. E. Woodcock, and P. Olofsson. 2012. Continuous monitoring of forest disturbance using all available Landsat imagery. Remote Sensing of Environment **122**.

Zona, D., W. C. Oechel, K. M. Peterson, R. J. Clements, U. K. T. Paw, and S. L. Ustin. 2010. Characterization of the carbon fluxes of a drained lake basin chronosequence on the Alaskan Arctic Coastal Plain. Global Change Biology **18**:1870-1882.

APPENDIX.

Appendix 1. Complete tabular list of Landsat imagery used in methodology.

Year	Month	Day	Path	Row	Landsat	Sensor
1985	7	15	61	11	5	TM
1985	7	31	61	11	5	TM
1985	7	15	61	12	5	TM
1985	7	31	61	12	5	TM
1985	9	1	61	12	5	TM
1985	8	7	62	11	5	TM
1985	8	7	62	12	5	TM
1985	6	27	63	12	5	TM
1985	8	5	64	11	5	TM
1985	6	25	65	11	5	TM
1986	7	18	61	11	5	TM
1986	7	2	61	12	5	TM
1986	7	18	61	12	5	TM
1986	8	10	62	11	5	TM
1986	9	11	62	11	5	TM
1986	8	10	62	12	5	TM
1986	9	11	62	12	5	TM
1986	8	17	63	10	5	TM
1986	8	17	63	11	5	TM
1986	9	2	63	11	5	TM
1986	7	23	64	11	5	TM
1986	6	21	64	12	5	TM
1986	7	7	64	12	5	TM
1986	8	8	64	12	5	TM
1988	6	28	62	11	5	TM
1988	6	28	62	12	5	TM
1988	7	5	63	10	5	TM
1988	8	6	63	10	5	TM
1988	7	5	63	11	5	TM
1988	7	3	65	11	5	TM
1989	6	24	61	11	5	TM
1989	8	18	62	11	5	TM
1989	8	18	62	12	5	TM

Appendix 1 (continued).

Year	Month	Day	Path	Row	Landsat	Sensor
1989	7	8	63	10	5	TM
1989	7	8	63	11	5	TM
1989	6	22	63	12	5	TM
1989	7	15	64	11	5	TM
1989	7	22	65	11	5	TM
1989	8	7	65	11	5	TM
1989	8	23	65	11	5	TM
1990	6	27	61	11	5	TM
1990	8	14	61	11	5	TM
1990	8	14	61	12	5	TM
1990	7	4	62	11	5	TM
1990	9	22	62	11	5	TM
1990	7	4	62	12	5	TM
1990	9	22	62	12	5	TM
1990	8	12	63	10	5	TM
1990	9	13	63	10	5	TM
1990	9	13	63	11	5	TM
1990	6	25	63	12	5	TM
1990	7	2	64	11	5	TM
1990	7	2	64	12	5	TM
1991	7	16	61	11	5	TM
1991	6	21	62	12	5	TM
1991	7	30	63	10	5	TM
1991	9	16	63	10	5	TM
1991	9	16	63	11	5	TM
1991	6	28	63	12	5	TM
1991	7	14	63	12	5	TM
1991	7	28	65	11	5	TM
1994	7	24	61	11	5	TM
1994	6	13	62	11	5	TM
1994	6	29	62	11	5	TM
1994	6	13	62	12	5	TM
1994	7	6	63	10	5	TM
1994	7	22	63	10	5	TM
1994	8	7	63	10	5	TM
1994	7	22	63	11	5	TM
1994	8	7	63	11	5	TM
1994	7	22	63	12	5	TM
1994	8	7	63	12	5	TM

Appendix 1 (continued).

Year	Month	Day	Path	Row	Landsat	Sensor
1994	7	13	64	11	5	TM
1994	7	29	64	11	5	TM
1994	7	4	65	11	5	TM
1995	7	11	61	11	5	TM
1995	7	11	61	12	5	TM
1995	7	2	62	11	5	TM
1995	9	20	62	11	5	TM
1995	7	2	62	12	5	TM
1995	7	9	63	10	5	TM
1995	8	10	63	10	5	TM
1995	8	10	63	11	5	TM
1995	6	30	64	11	5	TM
1995	6	30	64	12	5	TM
1995	9	18	64	12	5	TM
1996	7	13	61	11	5	TM
1996	6	18	62	12	5	TM
1996	9	6	62	12	5	TM
1996	8	12	63	10	5	TM
1996	8	28	63	10	5	TM
1996	8	12	63	11	5	TM
1996	8	28	63	11	5	TM
1996	7	11	63	12	5	TM
1996	8	28	63	12	5	TM
1997	6	30	61	11	5	TM
1997	6	30	61	12	5	TM
1997	6	5	62	12	5	TM
1997	7	30	63	11	5	TM
1997	7	30	63	12	5	TM
1998	6	17	61	11	5	TM
1998	7	19	61	12	5	TM
1998	8	20	61	12	5	TM
1998	7	1	63	10	5	TM
1998	7	17	63	10	5	TM
1998	6	15	63	11	5	TM
1998	7	1	63	11	5	TM
1998	6	15	63	12	5	TM
1998	7	1	63	12	5	TM
1998	8	2	63	12	5	TM
1998	6	22	64	11	5	TM

Appendix 1 (continued).

Year	Month	Day	Path	Row	Landsat	Sensor
1998	7	24	64	11	5	TM
1998	6	29	65	11	5	TM
1999	7	14	61	11	7	ETM+ SLC-on
1999	7	30	61	11	7	ETM+ SLC-on
1999	9	16	61	11	7	ETM+ SLC-on
1999	7	14	61	12	7	ETM+ SLC-on
1999	7	30	61	12	7	ETM+ SLC-on
1999	7	22	61	11	5	TM
1999	8	7	61	11	5	TM
1999	8	7	61	12	5	TM
1999	6	11	62	12	5	TM
1999	8	21	63	10	5	TM
1999	6	25	64	11	5	TM
2000	7	16	61	11	7	ETM+ SLC-on
2000	8	1	61	11	7	ETM+ SLC-on
2000	9	2	61	11	7	ETM+ SLC-on
2000	7	16	61	12	7	ETM+ SLC-on
2000	8	17	61	12	7	ETM+ SLC-on
2000	9	2	61	12	7	ETM+ SLC-on
2000	7	23	62	11	7	ETM+ SLC-on
2000	7	23	62	12	7	ETM+ SLC-on
2000	8	31	63	11	7	ETM+ SLC-on
2000	8	31	63	12	7	ETM+ SLC-on
2000	6	19	64	12	7	ETM+ SLC-on
2001	7	3	61	11	7	ETM+ SLC-on
2001	7	19	61	11	7	ETM+ SLC-on
2001	8	20	61	11	7	ETM+ SLC-on
2001	9	5	61	11	7	ETM+ SLC-on
2001	9	21	61	11	7	ETM+ SLC-on
2001	7	3	61	12	7	ETM+ SLC-on
2001	7	19	61	12	7	ETM+ SLC-on
2001	8	4	61	12	7	ETM+ SLC-on
2001	8	20	61	12	7	ETM+ SLC-on
2001	8	27	62	11	7	ETM+ SLC-on
2001	9	12	62	11	7	ETM+ SLC-on
2001	6	24	62	12	7	ETM+ SLC-on
2001	8	27	62	12	7	ETM+ SLC-on
2001	9	12	62	12	7	ETM+ SLC-on
2001	9	19	63	10	7	ETM+ SLC-on

Appendix 1 (continued).

Year	Month	Day	Path	Row	Landsat	Sensor
2001	7	17	63	11	7	ETM+ SLC-on
2001	9	19	63	11	7	ETM+ SLC-on
2001	7	17	63	12	7	ETM+ SLC-on
2001	8	18	63	12	7	ETM+ SLC-on
2001	9	19	63	12	7	ETM+ SLC-on
2001	9	10	64	11	7	ETM+ SLC-on
2001	10	3	65	11	7	ETM+ SLC-on
2002	7	22	61	11	7	ETM+ SLC-on
2002	9	8	61	11	7	ETM+ SLC-on
2002	7	22	61	12	7	ETM+ SLC-on
2002	9	8	61	12	7	ETM+ SLC-on
2002	7	13	62	11	7	ETM+ SLC-on
2002	7	20	63	10	7	ETM+ SLC-on
2002	9	6	63	10	7	ETM+ SLC-on
2002	9	6	63	11	7	ETM+ SLC-on
2002	7	20	63	12	7	ETM+ SLC-on
2002	9	6	63	12	7	ETM+ SLC-on
2002	8	28	64	11	7	ETM+ SLC-on
2002	9	29	64	11	7	ETM+ SLC-on
2002	8	28	64	12	7	ETM+ SLC-on
2002	7	18	65	11	7	ETM+ SLC-on
2003	9	2	62	11	7	ETM+ SLC-off
2003	7	21	65	11	7	ETM+ SLC-off
2003	8	2	61	11	5	TM
2003	8	2	61	12	5	TM
2003	8	9	62	12	5	TM
2003	9	1	63	10	5	TM
2004	8	12	61	12	7	ETM+ SLC-off
2004	6	16	62	12	7	ETM+ SLC-off
2004	6	23	63	12	7	ETM+ SLC-off
2004	7	25	63	12	7	ETM+ SLC-off
2004	8	17	64	11	7	ETM+ SLC-off
2004	7	23	65	11	7	ETM+ SLC-off
2004	9	25	65	11	7	ETM+ SLC-off
2004	7	19	61	11	5	TM
2004	7	26	62	11	5	TM
2004	7	26	62	12	5	TM
2004	7	1	63	10	5	TM
2004	7	17	63	10	5	TM

Appendix 1 (continued).

Year	Month	Day	Path	Row	Landsat	Sensor
2004	7	1	63	11	5	TM
2004	8	18	63	11	5	TM
2004	7	1	63	12	5	TM
2004	8	18	63	12	5	TM
2004	9	19	63	12	5	TM
2004	8	9	64	11	5	TM
2004	9	26	64	11	5	TM
2004	8	9	64	12	5	TM
2005	7	14	61	11	7	ETM+ SLC-off
2005	8	15	61	11	7	ETM+ SLC-off
2005	7	14	61	12	7	ETM+ SLC-off
2005	8	15	61	12	7	ETM+ SLC-off
2005	9	16	61	12	7	ETM+ SLC-off
2005	6	19	62	12	7	ETM+ SLC-off
2005	8	29	63	11	7	ETM+ SLC-off
2005	6	26	63	12	7	ETM+ SLC-off
2005	8	29	63	12	7	ETM+ SLC-off
2005	8	4	64	11	7	ETM+ SLC-off
2005	8	4	64	12	7	ETM+ SLC-off
2005	7	26	65	11	7	ETM+ SLC-off
2005	6	20	61	12	5	TM
2005	8	14	62	11	5	TM
2005	6	27	62	12	5	TM
2005	8	14	62	12	5	TM
2005	6	18	63	12	5	TM
2005	7	27	64	11	5	TM
2006	8	2	61	11	7	ETM+ SLC-off
2006	9	3	61	11	7	ETM+ SLC-off
2006	6	15	61	12	7	ETM+ SLC-off
2006	8	2	61	12	7	ETM+ SLC-off
2006	7	8	62	11	7	ETM+ SLC-off
2006	7	24	62	11	7	ETM+ SLC-off
2006	7	8	62	12	7	ETM+ SLC-off
2006	7	31	63	11	7	ETM+ SLC-off
2006	6	20	64	11	7	ETM+ SLC-off
2006	9	8	64	11	7	ETM+ SLC-off
2006	9	15	65	11	7	ETM+ SLC-off
2006	7	9	61	11	5	TM
2006	7	25	61	11	5	TM

Appendix 1 (continued).

Year	Month	Day	Path	Row	Landsat	Sensor
2006	8	10	61	11	5	TM
2006	8	26	61	11	5	TM
2006	7	9	61	12	5	TM
2006	7	25	61	12	5	TM
2006	8	10	61	12	5	TM
2006	6	30	62	11	5	TM
2006	8	1	62	11	5	TM
2006	9	18	62	11	5	TM
2006	6	14	62	12	5	TM
2006	6	30	62	12	5	TM
2006	9	18	62	12	5	TM
2006	7	14	64	11	5	TM
2006	9	16	64	11	5	TM
2006	9	7	65	11	5	TM
2007	7	4	61	11	7	ETM+ SLC-off
2007	6	18	61	12	7	ETM+ SLC-off
2007	7	11	62	11	7	ETM+ SLC-off
2007	8	12	62	11	7	ETM+ SLC-off
2007	8	28	62	11	7	ETM+ SLC-off
2007	6	9	62	12	7	ETM+ SLC-off
2007	6	25	62	12	7	ETM+ SLC-off
2007	7	11	62	12	7	ETM+ SLC-off
2007	8	12	62	12	7	ETM+ SLC-off
2007	8	28	62	12	7	ETM+ SLC-off
2007	9	4	63	11	7	ETM+ SLC-off
2007	9	4	63	12	7	ETM+ SLC-off
2007	7	16	65	11	7	ETM+ SLC-off
2007	9	18	65	11	7	ETM+ SLC-off
2007	7	28	61	11	5	TM
2007	8	29	61	11	5	TM
2007	6	10	61	12	5	TM
2007	8	29	61	12	5	TM
2007	7	3	62	11	5	TM
2007	8	4	62	11	5	TM
2007	9	5	62	11	5	TM
2007	6	17	62	12	5	TM
2007	7	3	62	12	5	TM
2007	8	4	62	12	5	TM
2007	7	10	63	10	5	TM

Appendix 1 (continued).

Year	Month	Day	Path	Row	Landsat	Sensor
2007	7	10	63	11	5	TM
2007	7	26	63	11	5	TM
2007	7	10	63	12	5	TM
2007	8	27	63	12	5	TM
2007	9	12	63	12	5	TM
2007	9	3	64	11	5	TM
2007	8	25	65	11	5	TM
2008	7	6	61	11	7	ETM+ SLC-off
2008	6	20	61	12	7	ETM+ SLC-off
2008	7	4	63	11	7	ETM+ SLC-off
2008	6	25	64	11	7	ETM+ SLC-off
2008	7	27	64	11	7	ETM+ SLC-off
2008	6	25	64	12	7	ETM+ SLC-off
2008	8	12	64	12	7	ETM+ SLC-off
2008	7	14	61	12	5	TM
2008	7	5	62	11	5	TM
2008	7	5	62	12	5	TM
2008	6	26	63	10	5	TM
2008	7	28	63	10	5	TM
2008	8	13	63	10	5	TM
2008	6	26	63	11	5	TM
2008	6	10	63	12	5	TM
2008	6	26	63	12	5	TM
2008	7	3	64	11	5	TM
2008	6	24	65	11	5	TM
2008	7	10	65	11	5	TM
2008	7	26	65	11	5	TM
2009	6	23	61	11	7	ETM+ SLC-off
2009	7	9	61	11	7	ETM+ SLC-off
2009	6	23	61	12	7	ETM+ SLC-off
2009	7	9	61	12	7	ETM+ SLC-off
2009	6	30	62	11	7	ETM+ SLC-off
2009	6	30	62	12	7	ETM+ SLC-off
2009	9	2	62	12	7	ETM+ SLC-off
2009	8	24	63	11	7	ETM+ SLC-off
2009	8	24	63	12	7	ETM+ SLC-off
2009	9	16	64	11	7	ETM+ SLC-off
2009	8	31	64	12	7	ETM+ SLC-off
2009	9	16	64	12	7	ETM+ SLC-off

Appendix 1 (continued).

Year	Month	Day	Path	Row	Landsat	Sensor
2009	8	18	61	11	5	TM
2009	7	1	61	12	5	TM
2009	8	18	61	12	5	TM
2009	9	3	61	12	5	TM
2009	7	8	62	11	5	TM
2009	8	25	62	11	5	TM
2009	7	8	62	12	5	TM
2009	7	6	64	11	5	TM
2009	8	23	64	11	5	TM
2010	7	28	61	11	7	ETM+ SLC-off
2010	6	10	61	12	7	ETM+ SLC-off
2010	8	4	62	11	7	ETM+ SLC-off
2010	8	4	62	12	7	ETM+ SLC-off
2010	9	5	62	12	7	ETM+ SLC-off
2010	6	24	63	11	7	ETM+ SLC-off
2010	7	10	63	11	7	ETM+ SLC-off
2010	6	8	63	12	7	ETM+ SLC-off
2010	7	10	63	12	7	ETM+ SLC-off
2010	9	12	63	12	7	ETM+ SLC-off
2010	8	18	64	11	7	ETM+ SLC-off
2010	9	19	64	11	7	ETM+ SLC-off
2010	8	5	61	11	5	TM
2010	7	4	61	12	5	TM
2010	8	5	61	12	5	TM
2010	7	27	62	11	5	TM
2010	6	9	62	12	5	TM
2010	7	27	62	12	5	TM
2010	8	19	63	10	5	TM
2010	8	19	63	11	5	TM
2010	8	19	63	12	5	TM
2010	8	17	65	11	5	TM
2011	7	15	61	11	7	ETM+ SLC-off
2011	7	15	61	12	7	ETM+ SLC-off
2011	10	3	61	12	7	ETM+ SLC-off
2011	6	20	62	11	7	ETM+ SLC-off
2011	6	27	63	11	7	ETM+ SLC-off
2011	6	27	63	12	7	ETM+ SLC-off
2011	8	5	64	11	7	ETM+ SLC-off
2011	7	14	62	11	5	TM

Appendix 1 (continued).

Year	Month	Day	Path	Row	Landsat	Sensor
2011	6	28	62	12	5	TM
2011	7	14	62	12	5	TM
2011	8	31	62	12	5	TM
2011	7	5	63	11	5	TM
2011	6	19	63	12	5	TM
2011	7	5	63	12	5	TM
2011	6	26	64	11	5	TM

Appendix 2. Extended methods section detailing the Fmask algorithm.

In order to spatially determine lake coverage over the scope of a Landsat image, a systematic classification of various land and water classes must first be applied for each image pixel. From these cover types, only those that are not classified as cloud and cloud shadow should be applied for further data processing. Clouds and their associated shadows inherently obstruct and complicate the view of the earth surface, and are an impediment to the assembly and analysis of dense time series of geospatial analysis (Martinuzzi et al. 2007, Sano et al. 2007, Ju and Roy 2008, Goodwin et al. 2013). However, as Goodwin et al. (2013) explain, a complete reliance on cloud-free imagery greatly restricts sampling opportunities, which could be alleviated by incorporating cloud-contaminated imagery in analyses, many of which contain large numbers of usable pixels. Therefore, to maximize the potential number of valid observations for any one pixel throughout the entirety of the time series, it is advantageous to include cloud and cloud shadow contaminated imagery and mask out the affected pixels.

In this light, the published Function of mask (Fmask) approach (Version 2.2, Windows Executable) was applied to all obtained Landsat imagery. Fmask is a newly-developed cloud and cloud shadow detection algorithm for the Landsat TM and ETM+ instruments that builds on the results of previous screening approaches. The Fmask algorithm is described in this section; a complete approach and evaluation can be found in Zhu and Woodcock (2012), and the code script downloadable at the project website URL (https://code.google.com/p/fmask/). Digital Number (DN) values were converted to top-of-atmosphere (TOA) reflectances and brightness temperatures (BT; °C) using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) atmospheric correction procedure (Masek et al. 2006, Vermote and Saleous 2007). Then, various spectral tests were conducted to identify and extract the land cover classes of terrain, water, cloud, cloud shadow, and snow/ice. The algorithm can be summarized into four distinct "passes":

Pass 1: Identification of potential cloud pixels;
Pass 2: Extraction of a potential cloud layer;
Pass 3: Identification and extraction of a potential cloud shadow layer; and
Pass 4: Identification and extraction of a potential snow and ice layer.

In this method, clouds are identified using a probability mask and a scene-based threshold, and cloud shadows are calculated using a combination of object matching, lapse rates, and a flood-filled transformation. In addition, land and water cover types are identified in Pass 2, and are designated for each pixel after identification of all other cover types. Because more than one class may exist within any one pixel, the algorithm assigns a class hierarchy in the following order from highest to lowest: clouds, cloud shadows, snow and ice, land and water.

Pass 1 – Potential clouds, cloud shadows, and snow.

The first pass compiles a series of spectral tests to flag pixels that may represent either clouds, termed "Potential Cloud Pixels (PCPs)". Pixels not included in the first pass are considered as confident in representing clear-sky pixels. Pixels will be labelled as PCPs if they satisfy all four of the following spectral tests.

Filter 1: Brightness and temperature threshold.

As a first step, the Basic Test was performed to fundamentally identify clouds by their temperature and reflectivity. The test is formulated as follows:

$$Basic\ Test: \begin{cases} B_7 > 0.03 \\ BT < 27 \\ NDSI < 0.8 \\ NDVI < 0.8 \end{cases}, \qquad \text{[Eq. A2.1.]}$$

where the Normalised Difference Snow Index (NDSI) and the Normalised Difference Vegetation Index (NDVI) was defined for Landsat images as:

$$NDSI = \frac{B_2 - B_5}{B_2 + B_5} \text{ and } NDVI = \frac{B_4 - B_3}{B_4 + B_3}. \qquad \text{[Eq. A2.2.]}$$

The top-of-atmosphere (TOA) reflectance for clouds usually decreases with increasing wavelength (Lyapustin et al. 2008) extending into the mid-infrared (Band 7) spectrum; additionally, the relatively low temperature of clouds was also encapsulated by thresholding the

brightness temperature at 27 °C. To distinguish PCPs from vegetated or snow-covered areas, additional thresholds were applied to the NDSI and NDVI.

Filter 2: Whiteness threshold.

As clouds always appear white due to their uniform spectral signature across the visible spectrum, the Whiteness Test (Gómez-Chova et al. 2007) was applied as follows:

$$Whiteness\ Test: \sum_{i=1}^{3} \left| \frac{B_i - \overline{l_v}}{\overline{l_v}} \right| < 0.7 , \qquad \text{[Eq. A2.3.]}$$

where the mean pixel value of the Landsat bands within the visible spectrum are defined as:

$$\overline{l_v} = \frac{B_1 + B_2 + B_3}{3} . \qquad \text{[Eq. A2.4.]}$$

The Whiteness Test was used to include pixels that are considered white enough to be classified as clouds; this test may also include pixels representing soil, sand, snow, and ice due to their similar flat reflectance in the visible bands.

Filter 3: Haze optimised transformation.

The Haze Optimised Transformation (HOT) for Landsat data was originally developed to target and characterize haze and thin cloud spatial distributions (Zhang et al. 2002). The image transform uses a simple spectral space consisting of blue and red bands based on the idea that, although the spectral responses of visible bands for all land cover types are highly correlated, the effect of haze on apparent radiance is significantly different between blue and red wavelengths. The HOT Test was applied as follows:

$$HOT\ Test: B_1 - 0.5 \times B_3 - 0.08 > 0 , \qquad \text{[Eq. A2.5]}$$

which separated thin and thick clouds, as well as thick aerosols from clear-sky pixels. Like the Whiteness Test, the HOT Test may also include pixels representing rocks, turbid water, snow, and ice surfaces due to their bright TOA reflectance values in the specified bands.

Filter 4: Band 4/band 5 ratio.

In the Automated Cloud Cover Assessment (ACCA; Irish 2000), a Band 4/Band 5 ratio was employed to distinguish highly reflective rocks and sands in arid landscapes from clouds, targeting those pixels not identified as PCPs using the HOT Test. The Band 4/5 ratio was applied as a test as follows:

$$B_4/B_5 \ Test: \ \frac{B_4}{B_5} > 0.75 \quad . \qquad \text{[Eq. A2.6.]}$$

2.2.2. Pass 2: Potential cloud layer.

The second pass used absolutely cloud-free pixels (pixels not identified as PCPs in Pass One) to compute potential cloud probability for all pixels in each image. Because temperature and reflectance distributions may vary between each input image, cloud probability was calculated separately for water and land pixels, described below. Discrimination between water and land pixels were conducted using a Water Test as follows:

$$Water\ Test: \begin{cases} NDVI < 0.01 \\ B_4 < 0.11 \end{cases} or \begin{cases} NDVI < 0.1 \\ B_4 < 0.05 \end{cases} , \qquad \text{[Eq. A2.7.]}$$

in which land pixels were identified by their low TOA reflectance and NDVI values, and water pixels identified by the converse. From this, cloud-free pixels over water and land bodies were respectively identified as follows:

$$Clear\text{-}Sky_w: \begin{cases} Water\ Test\ (True) \\ B_7 < 0.03 \end{cases} \quad \text{and} \qquad \text{[Eq. A2.8.]}$$

$$Clear\text{-}Sky_l: \begin{cases} Water\ Test\ (False) \\ PCP\ (False) \end{cases} . \quad \text{[Eq. A2.9.]}$$

By combining the results from Passes 1 and 2 (PCPs and potential cloud probability), a potential cloud layer was generated. In addition, missed cloud pixels are identified if the cloud probability over land is extremely large ($P > 0.99$) or if the BT is extremely cold ($BT < T_{low} - 35$). Finally, to err on cloud pixel commission rather than omission, the cloud mask was enlarged over a 3-by-3 pixel neighbourhood if 5 or more pixels represented clouds.

Filter 1: Cloud probability over water.

The overall cloud probability over water represents the product of its respective temperature and brightness probabilities, outlined below. A water pixel was designated as a cloud pixel if its overall cloud probability exceeds the 50th percentile.

The temperature probability over water was calculated as follows, normalizing the pixels' BT by 4 °C (Vermote and Saleous 2007):

$$P_{w,T} = \frac{T_w - BT}{4}, \quad \text{[Eq. A2.10.]}$$

where the clear-sky water temperature was classified using the upper level of clear-sky brightness temperatures over water:

$$T_w: P\left(BT_{w,C}\right) > 0.825. \quad \text{[Eq. A2.11.]}$$

The brightness probability over water was calculated as follows:

$$P_{w,B} = \frac{B_5}{0.11},\ B_5 \leq 0.11, \quad \text{[Eq. A2.12.]}$$

which takes advantage of water's generally dark TOA reflectance values, especially at the near-IR domain. The pixels' TOA reflectance values were normalized by 0.11, which was the highest observed value in Band 5.

Filter 2: Cloud probability over land.

The overall cloud probability over land represents the product of its respective temperature and variability probabilities, outlined below. A land pixel was designated as a cloud pixel if its overall cloud probability exceeds a threshold, given by the addition of the 82.5 percentile of clear-sky land pixels' probability and a constant of 0.2, based on sensitivity analyses.

The temperature probability over land was calculated as follows:

$$P_{l,T} = \frac{T_{high} + 4 - BT}{T_{high} + 4 - \left(T_{low} - 4\right)}, \qquad \text{[Eq. A2.13.]}$$

where the clear-sky land temperature was classified respectively using the lower and upper level of clear-sky brightness temperatures over water:

$$T_{low}: P\left(BT_{l,C}\right) = 0.175 \text{ and } T_{high}: P\left(BT_{l,C}\right) = 0.825. \qquad \text{[Eq. A2.14.]}$$

Pass 3: Cloud shadow layer

Filter 1: Flood-filling algorithm.

Because direct solar radiation towards the earth surface is blocked by clouds, their respective shadows are illuminated mostly by scattered light, and as the atmospheric scattering is stronger at shorter wavelengths, the intensity of the diffusive radiation within cloud shadows will be larger at these wavelengths (Luo et al. 2008). First, a flood-fill transformation was applied to the NIR band reflectance values, reducing the darkening effect of cloud shadows by matching the brightness values of its surrounding pixels (Soille 1999, Soille et al. 2003). From the output, the

difference between the flood-filled and the original reflectance values therefore assigns potential cloud shadow pixels with higher positive values, as follows:

$$\textit{Shadow Test: } B_{4,Ff} - B_4 > 0.02 \quad \text{[Eq. A2.15.]}$$

Filter 2: Cloud and cloud shadow attribution.

Cloud shadows result from the projection of the cloud onto the earth surface, as determined by the direction of incoming solar radiation (Simpson and Stitt 1998). Potential cloud and shadow layers can be matched assuming that a cloud and its assigned shadow are geometrically related. The location and shape of cloud shadows with respect to its attributed cloud was therefore matched by a series of geometric and trigonometric calculations relating the spatial location of the cloud, cloud top and base heights, and the sun and satellite positions. This was achieved by a series of iterations to determine these locations and positions.

Lastly, similar to the potential cloud mask, because the matched cloud shadow may contain spatial gaps, the shadow pixels were dilated by 3 pixels in 8-connected directions to fill these gaps.

Pass 4: Snow/ice layer.

The spectral tests used to detect snow and ice-contaminated pixels were derived from the MODIS snow mapping algorithm (Hall et al. 2001), with the exception of the NDSI threshold, which was lowered from 0.4 to 0.15. The MODIS algorithm uses the former threshold to only identify pixels containing at or over 50% coverage of snow and/or ice; for the current algorithm, the threshold was lowered to include pixels with snow/ice coverage less than 50% and snow in vegetated areas. The Snow/Ice Test was applied as follows:

$$\textit{Snow/Ice Test: } \begin{cases} NDSI > 0.15 \\ BT < 3.8 \\ B_4 > 0.11 \\ B_2 > 0.1 \end{cases} \quad \text{[Eq. A2.16.]}$$

Appendix 3. Python scripts to streamline batch processing in methodology.

3.1. Python ArcPy script to assemble Fmask images into yearly mosaics.

```
# ---------------------------------------------------------------------------
# fmask2poly.py
# Created on: 2013-04-12 15:40:06.00000
#   (generated by ArcGIS/ModelBuilder)
# Created by: TJ Young (SPRI)
# Description: This script was written to convert TIF images [generated from
#           the Fmask program version 2.2 (Windows Executable format) via
#           ENVI 4.6] to polygon files of only lakes. Output is the sum of
#           all lakes within a given year.
# ---------------------------------------------------------------------------
print "Running script 'fmask2poly.py'..."

# Import arcpy module
import os, sys, shutil
import arcpy

from arcpy import env
from arcpy.sa import *

# Check out any necessary licenses
arcpy.CheckOutExtension("spatial")

# Overwrite pre-existing files
arcpy.env.overwriteOutput = True

# Create input variables
year_input = raw_input('Enter year to process data: ')
data_path = "H:\\dissertation\\data\\"
scratch_path = "H:\\dissertation\\scratch\\"

print "\nProcessing data: " + str(year_input)

# Creates a list of contents in the working directory
workdir0 = "H:\\dissertation\\data\\fmask\\"
workdir1 = workdir0 + year_input + "\\"

dircontents = os.listdir(workdir1)
dirlist0 = list(dircontents)   # list contents in specified data folder
dirlist1 = []
for fname in dirlist0:
    fname_front, fname_end = os.path.splitext(fname)
    if fname_end == ".tif":
        dirlist1.append(fname) # only adds TIF files to dirlist1

mlakelist = []
mlandlist = []
micelist = []

# ArcPy Process 1a: Extract by Attributes (water)
print "\nRunning ArcPy Process 1a: Extract by Attributes (water)"
for fname in dirlist1:
    fname_path0 = workdir1 + fname
    fname_lakeras = fname[4:16] + "_lakeras.tif"
    fname_path1 = scratch_path + fname_lakeras # lists full path and file name for output water files

    arcpy.gp.ExtractByAttributes_sa(fname_path0, "\"Value\" = 1", fname_path1)
    mlakelist.append(fname_path1) # updates mlakelist
    print "Successfully completed tool 'Extract by  Attributes (water)' for raster file '" + fname +
"'."
    print "Saved output as file and path '" + fname_path1 + "'."

# ArcPy Process 1b: Raster Calculator (water)
print "\nRunning ArcPy Process 1b: Raster Calculator (water)"

mosaic_clip_back = "H:\\dissertation\\data\\mosaic_erdas_clip_back.tif"
mlakelist1 = []
for fname in mlakelist:
```

```
    fname1 =os.path.basename(fname[:-4]) + "1.tif"
    arcpy.MosaicToNewRaster_management([fname,mosaic_clip_back], scratch_path, fname1,
"PROJCS['WGS_84_UTM_zone_9N',GEOGCS['GCS_WGS_1984',DATUM['D_WGS_1984',SPHEROID['WGS_1984',6378137.0,29
8.257223563]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]],PROJECTION['Transverse_Mercat
or'],PARAMETER['false_easting',500000.0],PARAMETER['false_northing',0.0],PARAMETER['central_meridian',
-129.0],PARAMETER['scale_factor',0.9996],PARAMETER['latitude_of_origin',0.0],UNIT['Meter',1.0]]",
"8_BIT_UNSIGNED", "", "1", "MAXIMUM", "FIRST")
    # Converts null values of water rasters to 0 values; expands extent to full study area

    fname_path1 = os.path.dirname(fname) + "\\" + fname1
    mlakelist1.append(fname_path1) # updates mlakelist1

    print "Successfully completed tool 'Mosaic to New Raster (water)' for raster file '" + fname1 +
"'."
    print "Saved output as file and path '" + fname_path1 + "'."

print "Summing lake rasters..."

i = 0
for fname in mlakelist1:
    out2 = Raster(fname)
    if i == 0:
        out1 = out2
        i += 1
    else:
        out1 = out1 + out2
        i += 1
fname_lakeras_sum = scratch_path + str(year_input) + "_lakeras_sum.tif"
out1.save(fname_lakeras_sum)
# Sums all lake rasters together to create composite gradient value lake raster
print "Successfully summed all lake rasters for year '" + str(year_input) + "'."

# ArcPy Process 2a: Reclassify using Con (land)
print "\nRunning ArcPy Process 2a: Extract by Attributes (land)"
for fname in dirlist1:
    fname_path0 = workdir1 + fname
    fname_landras = fname[4:16] + "_landras.tif"
    fname_path1 = scratch_path + fname_landras # lists full path and file name for output land files

    inraster = Raster(fname_path0)
    outraster = Con(inraster == 0, 1, 0) # converts 0 values (land) to 1, converts all other values to
0.
    outraster.save(fname_path1)

    mlandlist.append(fname_path1)
    print "Successfully completed tool 'Raster Calculator (land)' for raster file '" + fname + "'."
    print "Saved output as file and path '" + fname_path1 + "'."

# ArcPy Process 2b: Raster Calculator (land)
print "\nRunning ArcPy Process 2b: Raster Calculator (land)"

mosaic_clip_back = "H:\\dissertation\\data\\mosaic_erdas_clip_back.tif"
mlandlist1 = []
for fname in mlandlist:
    fname1 =os.path.basename(fname[:-4]) + "1.tif"
    arcpy.MosaicToNewRaster_management([fname,mosaic_clip_back], scratch_path, fname1,
"PROJCS['WGS_84_UTM_zone_9N',GEOGCS['GCS_WGS_1984',DATUM['D_WGS_1984',SPHEROID['WGS_1984',6378137.0,29
8.257223563]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]],PROJECTION['Transverse_Mercat
or'],PARAMETER['false_easting',500000.0],PARAMETER['false_northing',0.0],PARAMETER['central_meridian',
-129.0],PARAMETER['scale_factor',0.9996],PARAMETER['latitude_of_origin',0.0],UNIT['Meter',1.0]]",
"8_BIT_UNSIGNED", "", "1", "MAXIMUM", "FIRST")
    # Converts null values of land rasters to 0 values; expands extent to full study area

    fname_path1 = os.path.dirname(fname) + "\\" + fname1
    mlandlist1.append(fname_path1)

    print "Successfully completed tool 'Mosaic to New Raster (land)' for raster file '" + fname1 +
"'."
    print "Saved output as file and path '" + fname_path1 + "'."

print "Summing land rasters..."

i = 0
for fname in mlandlist1:
```

```
    out2 = Raster(fname)
    if i == 0:
        out1 = out2
        i += 1
    else:
        out1 = out1 + out2
        i += 1
fname_landras_sum = scratch_path + str(year_input) + "_landras_sum.tif"
out1.save(fname_landras_sum)
# Sums all lake rasters together to create composite gradient value land raster
print "Successfully summed all land rasters for year '" + str(year_input) + "'."

# ArcPy Process 3a: Extract by Attributes (snow/ice)
print "\nRunning ArcPy Process 3a: Extract by Attributes (snow/ice)"
for fname in dirlist1:
    fname_path0 = workdir1 + fname
    fname_iceras = fname[4:16] + "_iceras.tif"
    fname_path1 = scratch_path + fname_iceras # lists full path and file name for output water files

    arcpy.gp.ExtractByAttributes_sa(fname_path0, "\"Value\" = 3", fname_path1)
    micelist.append(fname_path1) # updates micelist
    print "Successfully completed tool 'Extract by  Attributes (snow/ice)' for raster file '" + fname
+ "'."
    print "Saved output as file and path '" + fname_path1 + "'."

# ArcPy Process 3b: Raster Calculator (snow/ice)
print "\nRunning ArcPy Process 3b: Raster Calculator (snow/ice)"

mosaic_clip_back = "H:\\dissertation\\data\\mosaic_erdas_clip_back.tif"
micelist1 = []
for fname in micelist:
    fname1 =os.path.basename(fname[:-4]) + "1.tif"
    arcpy.MosaicToNewRaster_management([fname,mosaic_clip_back], scratch_path, fname1,
"PROJCS['WGS_84_UTM_zone_9N',GEOGCS['GCS_WGS_1984',DATUM['D_WGS_1984',SPHEROID['WGS_1984',6378137.0,29
8.257223563]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]],PROJECTION['Transverse_Mercat
or'],PARAMETER['false_easting',500000.0],PARAMETER['false_northing',0.0],PARAMETER['central_meridian',
-129.0],PARAMETER['scale_factor',0.9996],PARAMETER['latitude_of_origin',0.0],UNIT['Meter',1.0]]",
"8_BIT_UNSIGNED", "", "1", "MAXIMUM", "FIRST")
    # Converts null values of water rasters to 0 values; expands extent to full study area

    fname_path1 = os.path.dirname(fname) + "\\" + fname1
    micelist1.append(fname_path1) # updates mlakelist1

    print "Successfully completed tool 'Mosaic to New Raster (snow/ice)' for raster file '" + fname1 +
"'."
    print "Saved output as file and path '" + fname_path1 + "'."

print "Summing lake rasters..."

i = 0
for fname in micelist1:
    out2 = Raster(fname)
    if i == 0:
        out1 = out2
        i += 1
    else:
        out1 = out1 + out2
        i += 1
fname_iceras_sum = scratch_path + str(year_input) + "_iceras_sum.tif"
out1.save(fname_iceras_sum)
# Sums all lake rasters together to create composite gradient value ice raster
print "Successfully summed all ice rasters for year '" + str(year_input) + "'."

# ArcPy Process 4: Mosaic To New Raster (combined)
print "\nRunning ArcPy Process 4: Mosaic to New Raster (combined)"

fname_path1 = fname_lakeras_sum[:-4] + "10.tif"
fname_path2 = fname_landras_sum[:-4] + "10.tif"
fname_path3 = fname_iceras_sum[:-4] + "10.tif"
fname_path4 = data_path + "fmask_raster\\blend\\" + str(year_input) + "_ras_blend.tif"

inraster1 = Raster(fname_lakeras_sum)
outraster1 = 10*inraster1 # rescales water rasters from 0 (land) to 110 (highest water)
outraster1.save(fname_path1)
```

```
inraster2 = Raster(fname_landras_sum)
outraster2 = -10*inraster2 # rescales land rasters from -110 (water) to 0 (highest land)
outraster2.save(fname_path2)

inraster3 = Raster (fname_iceras_sum)
outraster3 = 5*inraster3
outraster3.save(fname_path3)

inraster4 = Raster(fname_path1)
inraster5 = Raster(fname_path2)
inraster6 = Raster(fname_path3)
outraster4 = inraster4 + inraster5 + inraster6
outraster4.save(fname_path4)

inras_lake = scratch_path + str(year_input) + "_lakeras_sum10.tif"
inras_land = scratch_path + str(year_input) + "_landras_sum10.tif"
inras_ice = scratch_path + str(year_input) + "_iceras_sum10.tif"
inras_comp = scratch_path + str(year_input) + "_compras_sum.tif"

data_path_fmaskras_blend = data_path + "fmask_raster\\blend\\"
ras_blend = str(year_input) + "_ras_blend.tif"

print "Successfully completed tool 'Mosaic to New Raster (combined)' for rasters in year: " +
str(year_input) + "'."

# ArcPy Process 4a: Extract by Attributes (water)
print "\nRunning ArcPy Process 4a: Extract by Attributes (water)"

print "Converting combined raster to binary values..."
year_ras_blend = data_path_fmaskras_blend + ras_blend
fname_path_bin = data_path + "fmask_raster\\binary\\median_bin\\" + ras_blend[:-10] + "_bin.tif"
inraster = Raster(year_ras_blend)
outraster = Con(inraster <0, 0, 1)
outraster.save(fname_path_bin)

print "Removing land pixels..."
fname_lakeras_bin = scratch_path + str(year_input) + "_ras_bin_lake.tif"
arcpy.gp.ExtractByAttributes_sa(fname_path_bin, "\"Value\" = 1", fname_lakeras_bin)

# ArcPy Process 4b: Raster to Polygon (combined)
print "\nRunning ArcPy Process 4b: Raster to Polygon (combined)"

print "Extracting pixels within study mask..."
fname_path_ras_mask = scratch_path + str(year_input) + "_ras_bin_lake_mask.tif"
study_mask = data_path + "\\study_site_image\\fig1_mask_upd.shp"
outExtractByMask = ExtractByMask(fname_lakeras_bin, study_mask)
outExtractByMask.save(fname_path_ras_mask)

print "Converting binary raster to polygon (combined)..."
fname_path_poly = data_path + "fmask_polygon\\median\\" + str(year_input) + "_poly_upd.shp"
arcpy.RasterToPolygon_conversion (fname_path_ras_mask, fname_path_poly, "NO_SIMPLIFY", "VALUE")

print "Successfully complete tool 'Raster to Polygon (combined) to create polygon shapefile '" +
fname_path_poly+ "'."

print "\nScript complete."
```

3.2. Python ArcPy script to extract Fmask polygons by specified fuzzy membership thresholds.

```
# ---------------------------------------------------------------------------
# fmask2poly_conf.py
# Created on: 2013-05-07 20:30:09.00000
#   (generated by ArcGIS/ModelBuilder)
# Created by: TJ Young (SPRI)
# Description: This script was written to assess the confidence of the lake
#           output given as the median of all pixels (on a gradient from land
#           to water).
# ---------------------------------------------------------------------------
print "Running script 'fmask2poly_conf.py'..."

# Import arcpy module
import os, sys, shutil
import arcpy

from arcpy import env
from arcpy.sa import *

# Check out any necessary licenses
arcpy.CheckOutExtension("spatial")

# Overwrite pre-existing files
arcpy.env.overwriteOutput = True

# Create input variables
year_input = raw_input('Enter year to process data: ')
del_input = raw_input('Delete intermediary files? (y/n): ')
data_path = "H:\\dissertation\\data\\"
scratch_path = "H:\\dissertation\\scratch\\"

print "\nProcessing data: " + str(year_input)

# Creates a list of contents in the working directory
workdir0 = "H:\\dissertation\\data\\fmask\\"
workdir1 = workdir0 + year_input + "\\"

dircontents = os.listdir(workdir1)
dirlist0 = list(dircontents)   # list contents in specified data folder
dirlist1 = []
for fname in dirlist0:
    fname_front, fname_end = os.path.splitext(fname)
    if fname_end == ".tif":
        dirlist1.append(fname) # only adds TIF files to dirlist1

print "Defining number of layers for each pixel..."

# ArcPy Process 1a: Con (Spatial Analyst)
print "\nRunning ArcPy Process 1a: Con (Spatial Analyst)"

env.workspace = workdir1 # Set environment workspace
dirlist_back = []

# Runs through conditional functions for each raster in list
for fname in dirlist1:
    inRaster = Raster(fname)
    outRaster = Con((inRaster <= 3) & (inRaster <> 2), 1, 0) # defines measured area (excludes NoData)
    outRaster_path = scratch_path + fname[4:16] + "_back.tif"
    outRaster.save(outRaster_path)
    dirlist_back.append(outRaster_path)
    print "Successfully completed tool 'Con (Spatial Analyst)' for raster file '" + fname + "'."

print "Successfully completed tool 'Con (Spatial Analyst)' for all files."

# ArcPyProcess 1b: Mosaic to New Raster (Spatial Analyst)
print "\nRunning ArcPy Process 1b: Mosaic to New Raster (Spatial Analyst)"

env.workspace = scratch_path # Change environment workspace

dirlist_back2 = []
mosaic_clip_back = "H:\\dissertation\\data\\mosaic_erdas_clip_back.tif"
outLocation = scratch_path
```

```
coordSystem =
"PROJCS['WGS_84_UTM_zone_9N',GEOGCS['GCS_WGS_1984',DATUM['D_WGS_1984',SPHEROID['WGS_1984',6378137.0,29
8.257223563]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]],PROJECTION['Transverse_Mercat
or'],PARAMETER['false_easting',500000.0],PARAMETER['false_northing',0.0],PARAMETER['central_meridian',
-129.0],PARAMETER['scale_factor',0.9996],PARAMETER['latitude_of_origin',0.0],UNIT['Meter',1.0]]"
pixelType = "8_BIT_UNSIGNED"
cellSize = ""
bands = "1"
mosaicMethod = "MAXIMUM"
mosaicMode = "FIRST"

for fname in dirlist_back:
    outRaster = os.path.basename(fname[:-4]) + "1.tif"
    arcpy.MosaicToNewRaster_management([fname,mosaic_clip_back], outLocation, outRaster, coordSystem,
                                       pixelType, cellSize, bands, mosaicMethod, mosaicMode)
    outRaster_path = os.path.dirname(fname) + "\\" + outRaster
    dirlist_back2.append(outRaster_path)
    print "Successfully completed tool 'Mosaic to New Raster' for raster file '" +
os.path.basename(fname) + "'."

print "Successfully completed tool 'Mosaic to New Raster (Spatial Analyst)' for all files.."

print "\nSumming background rasters..."
i = 0
for fname in dirlist_back2:
    out2 = Raster(fname)
    if i == 0:
        out1 = out2
        i += 1
    else:
        out1 = out1 + out2
        i += 1
backras_sum = scratch_path + str(year_input) + "_backsum.tif"
out1.save(backras_sum)
# Sums all lake rasters together to create composite gradient value lake raster
print "Successfully summed all background rasters for year '" + str(year_input) + "'."

print "Successfully obtained pixel-based layered background raster for year " + str(year_input) + "."

# ArcPy Process 2: Raster Calculator (Spatial Analyst)
print "\nRunning ArcPy Process 2: Raster Calculator (Spatial Analyst)"

fmaskras_blend = data_path + "fmask_raster\\blend\\" + str(year_input) + "_ras_blend.tif"
fmaskras_rescale = scratch_path + str(year_input) + "_ras_rescale.tif"

inRaster1 = Raster(backras_sum)
inRaster2 = Raster(fmaskras_blend)
outRaster = (10 * inRaster2)/inRaster1
outRaster.save(fmaskras_rescale)

print "Statistics for background raster in year: " + str(year_input)
inRaster = fmaskras_rescale
propertyType = "MAXIMUM"

rescale_result = arcpy.GetRasterProperties_management(inRaster, propertyType)
rescale_value = rescale_result.getOutput(0)
rescale_max = float(rescale_value)
print propertyType + " for rescaled raster: " + str(rescale_value)

propertyType = "MINIMUM"

rescale_result = arcpy.GetRasterProperties_management(inRaster, propertyType)
rescale_value = rescale_result.getOutput(0)
rescale_min = float(rescale_value)
print propertyType + " for rescaled raster: " + str(rescale_value)

dif_factor = (rescale_max - rescale_min)/2
dif_shift = ((5 * rescale_min)/dif_factor) * -1

fmaskras_10 = data_path + "fmask_raster\\10\\" + str(year_input) + "_ras_10.tif"

inRaster = Raster(fmaskras_rescale)
outRaster = (5 * inRaster)/dif_factor + dif_shift
outRaster.save(fmaskras_10)
```

```
print "Successfully completed tool 'Raster Calculator (Spatial Analyst)"

# ArcPy Process 3: Fuzzy Membership (Spatial Analyst)
print "\nRunning ArcPy Process 3: Fuzzy Membership (Spatial Analyst)"

env.workspace = data_path + "fmask_raster\\10\\"  # Set environment workspace

# Set local variables
inRaster = str(year_input) + "_ras_10.tif"
outRaster_large = scratch_path + str(year_input) + "_fuzzy.tif"

# Create the Fuzzy algorithm object and raster (Large)
propertyType = "MEAN"

rescale_result = arcpy.GetRasterProperties_management(inRaster, propertyType)
rescale_value = rescale_result.getOutput(0)
rescale_mean = float(rescale_value)
print propertyType + " for rescaled raster: " + str(rescale_value)

# For Fuzzy Large
midpoint = 5
spread = 5
fuzzyAlgorithm = FuzzyLarge(midpoint, spread)

outRaster = FuzzyMembership(inRaster, fuzzyAlgorithm)
outRaster.save(outRaster_large)
fuzzy_inc = outRaster_large

# For Fuzzy MSLarge

print "Successfully completed tool 'Fuzzy Membership (Spatial Analyst)' for file '" + str(year_input)
+ "_fuzzy_large.tif'."

# ArcPy Process 4: Reclassify (Spatial Analyst)
print "\nRunning ArcPy Process 4: Reclassify (Spatial Analyst)"

print "Reclassifying IsNull values..."
inRaster = Raster(fuzzy_inc)
outRaster = scratch_path + str(year_input) + "_fuzzy_isnull.tif"
fuzzy_isnull = IsNull(inRaster)
fuzzy_isnull.save(outRaster)
fuzzy_null = outRaster

inConRaster = Raster(fuzzy_null)
inTrueConstant = 0.00000001
inFalseRaster = Raster(fuzzy_inc)
outRaster = data_path + "fmask_raster\\fuzzy\\" + str(year_input) + "_fuzzy_large.tif"

outCon = Con(inConRaster == 1, inTrueConstant, inFalseRaster)
outCon.save(outRaster)
fuzzy = outRaster

print "Successfully reclassified IsNull values. Successfully ran ArcPy Process 4."

# ArcPy Process 5: Reclassify (Spatial Analyst)
print "\nRunning ArcPy Process 5: Reclassify (Spatial Analyst)"

thresholdlist = [0.3, 0.5, 0.7]

for threshold in thresholdlist:
    print "\nConverting fuzzy raster to binary values using set threshold: " + str(threshold)
    outFolder = data_path + "fmask_raster\\binary\\"
    outFolderYear = outFolder + "fuzzy_" + str(int(threshold * 100)) + "\\"
    if not os.path.exists(outFolderYear):
        os.makedirs(outFolderYear)
    inRaster = Raster(fuzzy)
    outRaster = outFolderYear + str(year_input) + "_fuzzy" + str(int(threshold * 100)) + "_bin.tif"
    fuzzy_threshold = Con(inRaster > float(threshold), 1, 0)
    fuzzy_threshold.save(outRaster)
    fuzzy_threshold_bin = outRaster

    print "Removing land pixels..."
    outRaster = scratch_path + str(year_input) + "_fuzzy" + str(int(threshold * 100)) + "_bin_inc.tif"
```

```
    arcpy.gp.ExtractByAttributes_sa(fuzzy_threshold_bin, "\"Value\" = 1", outRaster)
    fuzzy_threshold_bin = outRaster

    print "Extracting pixels within study mask..."
    inRaster = fuzzy_threshold_bin
    outRaster = scratch_path + str(year_input) + "_fuzzy" + str(int(threshold * 100)) +
"_bin_lake.tif"
    study_mask = data_path + "\\study_site_image\\fig1_mask_upd.shp"
    outExtractByMask = ExtractByMask(inRaster, study_mask)
    outExtractByMask.save(outRaster)
    fuzzy_threshold_bin_extract = outRaster

    print "Converting binary raster to polygon (combined)..."
    inRaster = fuzzy_threshold_bin_extract
    simplify = "NO_SIMPLIFY"
    rasterField = "VALUE"
    outFolder = data_path + "fmask_polygon\\"
    outFolderYear = outFolder + "fuzzy_" + str(int(threshold * 100)) + "\\"
    if not os.path.exists(outFolderYear):
        os.makedirs(outFolderYear)
    outRaster = outFolderYear + str(year_input) + "_fuzzy" + str(int(threshold * 100)) + "_poly.shp"
    arcpy.RasterToPolygon_conversion(inRaster, outRaster, simplify, rasterField)

    print "Successfully completed tool 'Raster to Polygon' to create polygon shapefile '" +
os.path.basename(outRaster) + "'."

# Delete scratch files
if del_input == "y":
    print "\nDeleting intermediary files in scratch..."
    for fname in dirlist_back: # Deletes output in Process 1a
        inData = fname
        dataType = ""
        arcpy.Delete_management(inData, dataType)
    for fname in dirlist_back2: # Deletes output in Process 1b
        inData = fname
        dataType = ""
        arcpy.Delete_management(inData, dataType)
    arcpy.Delete_management(backras_sum, "")
    arcpy.Delete_management(fmaskras_rescale, "") # Deletes output in Process 2
    arcpy.Delete_management(outRaster_large, "") # Deletes output in Process 3
    arcpy.Delete_management(fuzzy_null, "") # Deletes output in Process 4
    for threshold in thresholdlist:
        inData = scratch_path + str(year_input) + "_fuzzy" + str(int(threshold * 100)) +
"_bin_inc.tif"
        dataType = ""
        arcpy.Delete_management(inData, dataType)
        inData = scratch_path + str(year_input) + "_fuzzy" + str(int(threshold * 100)) +
"_bin_lake.tif"
        dataType = ""
        arcpy.Delete_management(inData, dataType)

elif del_input == "n":
    pass

else:
    pass

print "\nScript complete."
```

3.3. Python ArcPy script to assign unique ID numbers to individual lakes for all years.

```
# ---------------------------------------------------------------------------
# fmask_sj.py
# Created on: 2013-05-30 10:26:53.00000
#   (generated by ArcGIS/ModelBuilder)
# Created by: TJ Young (SPRI)
# Description: This script was written to assign each lake polygon a unique
#     ID that corresponds to the same ID for a given master lake mask. This
#     ensures that all polygons with the same X and Y coordinate in any year
#     has the same ID number. In general, polygons increase in unique ID
#     number with latitude. This script also calculates area and perimeter
#     for each polygon.
# ---------------------------------------------------------------------------
print "Running script 'fmask_sj.py'..."

# Import arcpy module
import os, sys, shutil
import arcpy

from arcpy import env
from arcpy.sa import *

# Check out any necessary licenses
arcpy.CheckOutExtension("spatial")

# Overwrite pre-existing files
arcpy.env.overwriteOutput = True

# Create input variables
fmask_path = "H:\\dissertation\\data\\fmask_polygon\\"
data_path = "H:\\dissertation\\data\\"
scratch_path = "H:\\dissertation\\scratch\\"

fuzzylist_int = [30, 50, 70]
yearlist_int = [1994]

#yearlist_int = [1985, 1986, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999,
                #2000, 2001, 2002, 2003, 2004 ,2005, 2006, 2007, 2008, 2009, 2010, 2011]
fuzzylist_str = map(str, fuzzylist_int)
yearlist_str = map(str, yearlist_int) # Converts years from integer to string

for acut in fuzzylist_str:
    fuzzy_path = fmask_path + "fuzzy_" + acut + "\\"
    print "\nRunning polygons for specified years having alpha-threshold: " + acut + "."

    # Copy Features (Data Management)
    inFeature = fmask_path + "join\\fuzzy_union.shp"
    outFeature = fmask_path + "join\\fuzzy_union" + acut + ".shp"

    arcpy.CopyFeatures_management(inFeature, outFeature)

    # Add Field (Data Management)
    inFeature = fmask_path + "join\\fuzzy_union" + acut + ".shp"
    fieldName = "FID_fuzzy_"
    fieldType = "LONG"
    fieldPrecision = ""
    fieldScale = ""
    fieldLength = ""
    fieldAlias = ""
    fieldIsNullable = "NON_NULLABLE"
    fieldIsRequired = "NON_REQUIRED"
    fieldDomain = ""

    arcpy.AddField_management(inFeature, fieldName, fieldType, fieldPrecision, fieldScale,
                              fieldLength, fieldAlias,fieldIsNullable, fieldIsRequired,
                              fieldDomain)

    # Calculate Field (Data Management)
    expression = "!FID!"
    expressionType = "PYTHON_9.3"
    codeBlock = ""
```

```
    arcpy.CalculateField_management(inFeature, fieldName, expression, expressionType, codeBlock)

    # Add Index (Data Management)
    fields = "FID_fuzzy_"
    indexName = ""
    unique = "UNIQUE"
    ascending = "ASCENDING"

    arcpy.AddIndex_management(inFeature, fields, indexName, unique, ascending)

    for year in yearlist_str:
        print "\nProcessing polygon for year: " + year + "."

        # Identity (Analysis)
        print "Assigning Unique ID..."

        inFeature = fuzzy_path + year + "_fuzzy" + acut + "_poly.shp"
        identityFeature = fmask_path + "join\\fuzzy_union.shp"
        outFeature = scratch_path + "fuzzy" + acut + "_" + year + "_id.shp"
        joinAttributes = "ALL"
        clusterTolerance = ""
        relationship = "NO_RELATIONSHIPS"

        arcpy.Identity_analysis(inFeature, identityFeature, outFeature, joinAttributes,
                                clusterTolerance, relationship)

        # Dissolve (Data Management)
        print "Dissolving fields with same Unique IDs..."

        inFeature = scratch_path + "fuzzy" + acut + "_" + year + "_id.shp"
        outFeature = fmask_path + "join\\fuzzy" + acut + "_" + year + "_id.shp"
        dissolveField = "FID_fuzzy_"
        statisticsField = ""
        multiPart = "MULTI_PART"
        unsplitLines = "DISSOLVE_LINES"

        arcpy.Dissolve_management(inFeature, outFeature, dissolveField, statisticsField, multiPart,
unsplitLines)

        # Add Field (Data Management)
        print "Calculating area..."

        inFeature = fmask_path + "join\\fuzzy" + acut + "_" + year + "_id.shp"
        fieldName = "AREA" + year
        fieldType = "LONG"
        fieldPrecision = ""
        fieldScale = ""
        fieldLength = ""
        fieldAlias = ""
        fieldIsNullable = "NON_NULLABLE"
        fieldIsRequired = "NON_REQUIRED"
        fieldDomain = ""

        arcpy.AddField_management(inFeature, fieldName, fieldType, fieldPrecision, fieldScale,
                                  fieldLength, fieldAlias,fieldIsNullable, fieldIsRequired,
                                  fieldDomain)

        # Calculate Field (Data Management)
        print "Copying relevant fields to master polygon shapefile..."

        expression = "!SHAPE.AREA!"
        expressionType = "PYTHON_9.3"
        codeBlock = ""

        arcpy.CalculateField_management(inFeature, fieldName, expression, expressionType, codeBlock)
        """
        # Delete (Data Management)
        print "Deleting intermediary files in scratch..."
        inData = scratch_path + "fuzzy" + acut + "_" + year + "_id.shp"
        dataType = ""
        arcpy.Delete_management(inData, dataType)

        print "\nProcess complete for year: " + year + " having alpha-threshold: " + acut + "."
```

```
"""
    # Delete (Data Management)
    print "Deleting intermediary files in scratch..."
    inData = scratch_path + "fuzzy_union_" + acut + ".shp"
    dataType = ""
    arcpy.Delete_management(inData, dataType)

    print "\nProcess complete for alpha-threshold: " + acut + "."

print "\nScript complete."
```

3.4. Python ArcPy script to generate random points for each input Landsat image for validation.

```
# ---------------------------------------------------------------------------
# fmask2poly.py
# Created on: 2013-05-26 17:24:30.00000
#   (generated by ArcGIS/ModelBuilder)
# Created by: TJ Young (SPRI)
# Description: This script was written to sample weather rasters using inputs
#    as defined from statistical significance in linear regression. Code was
#    blocked out as necessary.
# ---------------------------------------------------------------------------

print "Running script 'weather_extract.py'..."

# Import arcpy module
import os, sys, shutil
import csv
import arcpy

from arcpy import env
from arcpy.sa import *

# Check out any necessary licenses
arcpy.CheckOutExtension("spatial")

# Overwrite pre-existing files
arcpy.env.overwriteOutput = True

# Create input paths
data_path = "H:\\dissertation\\data\\"
scratch_path = "H:\\dissertation\\scratch\\"
weather_path = "H:\\dissertation\\docs\\weather\\sample\\"

varlist = ["t2m", "tp0", "tp12"]

# Get yearly and summer temperatures and precipitation
t2m_path = "H:\\dissertation\\data\\weather\\temperature_2m\\utm_zone9\\"

year_range_int = list(xrange(1984, 2012))
year_range = map(str, year_range_int)
leapyear_list = ["1984", "1988", "1992", "1996", "2000", "2004", "2008", "2012"]
"""
# Temperature

print "Getting averages for temperature..."
for year in year_range:
    month1 = Raster(t2m_path + year + "01.tif")
    month2 = Raster(t2m_path + year + "02.tif")
    month3 = Raster(t2m_path + year + "03.tif")
    month4 = Raster(t2m_path + year + "04.tif")
    month5 = Raster(t2m_path + year + "05.tif")
    month6 = Raster(t2m_path + year + "06.tif")
    month7 = Raster(t2m_path + year + "07.tif")
    month8 = Raster(t2m_path + year + "08.tif")
    month9 = Raster(t2m_path + year + "09.tif")
    month10 = Raster(t2m_path + year + "11.tif")
    month11 = Raster(t2m_path + year + "11.tif")
    month12 = Raster(t2m_path + year + "12.tif")

    # For temperature yearly averages:
```

```
    print "year" + year
    outRaster = (month1 + month2 + month3 + month4 + month5 + month6 + month7 + month8 +
                 month9 + month10 + month11 + month12) / 12
    outRaster.save("H:\\dissertation\\data\\weather\\temperature_2m\\year\\" + year + ".tif")

    # For temperature growing season averages:
    print "growing season" + year
    outRaster = (month5 + month6 + month7 + month8 + month9) / 5
    outRaster.save("H:\\dissertation\\data\\weather\\temperature_2m\\growing\\" + year + ".tif")
"""
# NOTE: I labelled precipitation files wrong. E.g. tp12201201 is 201201 time step 00,
#       tp00201202 is 201201 time step 12, tp12201202 is 201202 time step 00,
#       tp00201203 is 201202 time step 12, etc. The output is daily averages of monthly means,
#       says ECMWF...
tp0_path = "H:\\dissertation\\data\\weather\\total_precipitation0\\utm_zone9\\"
tp12_path = "H:\\dissertation\\data\\weather\\total_precipitation12\\utm_zone9\\"

# Precipitation

print "Getting totals for precipitation..."
for year in year_range:
    year_int = int(year)
    month1 = (Raster(tp12_path + year + "01.tif") + Raster(tp0_path + year + "02.tif")) * 31
    if year in leapyear_list:
        month2 = (Raster(tp12_path + year + "02.tif") + Raster(tp0_path + year + "03.tif")) * 29
    else:
        month2 = (Raster(tp12_path + year + "02.tif") + Raster(tp0_path + year + "03.tif")) * 28
    month3 = (Raster(tp12_path + year + "03.tif") + Raster(tp0_path + year + "04.tif")) * 31
    month4 = (Raster(tp12_path + year + "04.tif") + Raster(tp0_path + year + "05.tif")) * 30
    month5 = (Raster(tp12_path + year + "05.tif") + Raster(tp0_path + year + "06.tif")) * 31
    month6 = (Raster(tp12_path + year + "06.tif") + Raster(tp0_path + year + "07.tif")) * 30
    month7 = (Raster(tp12_path + year + "07.tif") + Raster(tp0_path + year + "08.tif")) * 31
    month8 = (Raster(tp12_path + year + "08.tif") + Raster(tp0_path + year + "09.tif")) * 31
    month9 = (Raster(tp12_path + year + "09.tif") + Raster(tp0_path + year + "10.tif")) * 30
    month10 = (Raster(tp12_path + year + "11.tif") + Raster(tp0_path + year + "11.tif")) * 31
    month11 = (Raster(tp12_path + year + "11.tif") + Raster(tp0_path + year + "12.tif")) * 30
    month12 = (Raster(tp12_path + year + "12.tif") + Raster(tp0_path + str(year_int + 1) + "01.tif"))
* 31

    # For precipitation yearly averages:
    print "year" + year
    outRaster = (month1 + month2 + month3 + month4 + month5 + month6 + month7 + month8 +
                 month9 + month10 + month11 + month12) * 1000
    outRaster.save("H:\\dissertation\\data\\weather\\total_precipitation\\year\\" + year + ".tif")

    # For temperature growing season averages:
    print "growing season" + year
    outRaster = (month5 + month6 + month7 + month8 + month9) * 1000
    outRaster.save("H:\\dissertation\\data\\weather\\total_precipitation\\growing\\" + year + ".tif")

    # For previous year (June to May):
    print "previous year" + year
    year_int = int(year)
    month1 = (Raster(tp12_path + year + "01.tif") + Raster(tp0_path + year + "02.tif")) * 31
    if year in leapyear_list:
        month2 = (Raster(tp12_path + year + "02.tif") + Raster(tp0_path + year + "03.tif")) * 29
    else:
        month2 = (Raster(tp12_path + year + "02.tif") + Raster(tp0_path + year + "03.tif")) * 28
    month3 = (Raster(tp12_path + year + "03.tif") + Raster(tp0_path + year + "04.tif")) * 31
    month4 = (Raster(tp12_path + year + "04.tif") + Raster(tp0_path + year + "05.tif")) * 30
    month5 = (Raster(tp12_path + year + "05.tif") + Raster(tp0_path + year + "06.tif")) * 31
    month6 = (Raster(tp12_path + str(year_int - 1) + "06.tif") + Raster(tp0_path + str(year_int - 1) +
"07.tif")) * 30
    month7 = (Raster(tp12_path + str(year_int - 1) + "07.tif") + Raster(tp0_path + str(year_int - 1) +
"08.tif")) * 31
    month8 = (Raster(tp12_path + str(year_int - 1) + "08.tif") + Raster(tp0_path + str(year_int - 1) +
"09.tif")) * 31
    month9 = (Raster(tp12_path + str(year_int - 1) + "09.tif") + Raster(tp0_path + str(year_int - 1) +
"10.tif")) * 30
    month10 = (Raster(tp12_path + str(year_int - 1) + "11.tif") + Raster(tp0_path + str(year_int - 1)
+ "11.tif")) * 31
    month11 = (Raster(tp12_path + str(year_int - 1) + "11.tif") + Raster(tp0_path + str(year_int - 1)
+ "12.tif")) * 30
```

```
    month12 = (Raster(tp12_path + str(year_int - 1) + "12.tif") + Raster(tp0_path + year + "01.tif"))
* 31

    print "year" + year
    outRaster = (month1 + month2 + month3 + month4 + month5 + month6 + month7 + month8 +
                 month9 + month10 + month11 + month12) * 1000
    outRaster.save("H:\\dissertation\\data\\weather\\total_precipitation\\previous\\" + year + ".tif")


# ArcPy Process 1: Sample (Spatial Analyst)
print "Running ArcPy Process 1: Sample (Spatial Analyst)"

# Create input paths
period_path = "H:\\dissertation\\data\\spatial_stats\\"
chow1_path = "H:\\dissertation\\data\\spatial_stats\\chow1\\"
chow2_path = "H:\\dissertation\\data\\spatial_stats\\chow2\\"

# Read in sample points
period_pt = period_path + "lakes_reg_pt.shp"
chow1_pt = chow1_path + "lakes_reg_pt.shp"
chow2_pt = chow2_path + "lakes_reg_pt.shp"

tsptlist = [period_pt, chow1_pt, chow2_pt]

# 1a. Yearly average of temperature

print "\n1a. Yearly average of temperature"
year_range_int = list(xrange(1985, 2012))
year_range = map(str, year_range_int)
year_range_chow1 = year_range[:12]
year_range_chow2 = year_range[12:]

for ts in tsptlist:
    if ts == period_pt:
        print "Sampling for entire time domain..."
        yearlist_period = []
        for year in year_range:
            yearfile = "H:\\dissertation\\data\\weather\\temperature_2m\\year\\" + year + ".tif"
            yearlist_period.append(yearfile)
        inRasters = yearlist_period
        outTable = weather_path + "period_tyear.dbf"
    if ts == chow1_pt:
        print "Sampling for chow1 domain..."
        yearlist_chow1 = []
        for year in year_range_chow1:
            yearfile = "H:\\dissertation\\data\\weather\\temperature_2m\\year\\" + year + ".tif"
            yearlist_chow1.append(yearfile)
        inRasters = yearlist_chow1
        outTable = weather_path + "chow1_tyear.dbf"
    if ts == chow2_pt:
        print "Sampling for chow2 domain..."
        yearlist_chow2 = []
        for year in year_range_chow2:
            yearfile = "H:\\dissertation\\data\\weather\\temperature_2m\\year\\" + year + ".tif"
            yearlist_chow2.append(yearfile)
        inRasters = yearlist_chow2
        outTable = weather_path + "chow2_tyear.dbf"
    inPoint = ts
    resamplingType = "NEAREST"

    Sample(inRasters, inPoint, outTable, resamplingType)

# 1b. Growing season average of temperature
print "\n1b. Growing season average of temperature"

year_range_int = list(xrange(1985, 2012))
year_range = map(str, year_range_int)
year_range_chow1 = year_range[:12]
year_range_chow2 = year_range[12:]

for ts in tsptlist:
    if ts == period_pt:
        print "Sampling for entire time domain..."
        yearlist_period = []
```

```
        for year in year_range:
            yearfile = "H:\\dissertation\\data\\weather\\temperature_2m\\growing\\" + year + ".tif"
            yearlist_period.append(yearfile)
        inRasters = yearlist_period
        outTable = weather_path + "period_tgs.dbf"
    if ts == chow1_pt:
        print "Sampling for chow1 domain..."
        yearlist_chow1 = []
        for year in year_range_chow1:
            yearfile = "H:\\dissertation\\data\\weather\\temperature_2m\\growing\\" + year + ".tif"
            yearlist_chow1.append(yearfile)
        inRasters = yearlist_chow1
        outTable = weather_path + "chow1_tgs.dbf"
    if ts == chow2_pt:
        print "Sampling for chow2 domain..."
        yearlist_chow2 = []
        for year in year_range_chow2:
            yearfile = "H:\\dissertation\\data\\weather\\temperature_2m\\growing\\" + year + ".tif"
            yearlist_chow2.append(yearfile)
        inRasters = yearlist_chow2
        outTable = weather_path + "chow2_tgs.dbf"
    inPoint = ts
    resamplingType = "NEAREST"

    Sample(inRasters, inPoint, outTable, resamplingType)

# 1c. Yearly average of precipitation
print "\n1c. Yearly average of precipitation"

year_range_int = list(xrange(1985, 2012))
year_range = map(str, year_range_int)
year_range_chow1 = year_range[:12]
year_range_chow2 = year_range[12:]

for ts in tsptlist:
    if ts == period_pt:
        print "Sampling for entire time domain..."
        yearlist_period = []
        for year in year_range:
            yearfile = "H:\\dissertation\\data\\weather\\total_precipitation\\year\\" + year + ".tif"
            yearlist_period.append(yearfile)
        inRasters = yearlist_period
        outTable = weather_path + "period_tpyear.dbf"
    if ts == chow1_pt:
        print "Sampling for chow1 domain..."
        yearlist_chow1 = []
        for year in year_range_chow1:
            yearfile = "H:\\dissertation\\data\\weather\\total_precipitation\\year\\" + year + ".tif"
            yearlist_chow1.append(yearfile)
        inRasters = yearlist_chow1
        outTable = weather_path + "chow1_tpyear.dbf"
    if ts == chow2_pt:
        yearlist_chow2 = []
        print "Sampling for chow2 domain..."
        for year in year_range_chow2:
            yearfile = "H:\\dissertation\\data\\weather\\total_precipitation\\year\\" + year + ".tif"
            yearlist_chow2.append(yearfile)
        inRasters = yearlist_chow2
        outTable = weather_path + "chow2_tpyear.dbf"
    inPoint = ts
    resamplingType = "NEAREST"

    Sample(inRasters, inPoint, outTable, resamplingType)

# 1d. Growing season average of precipitation
print "\n1d. Growing season average of precipitation"

year_range_int = list(xrange(1985, 2012))
year_range = map(str, year_range_int)
year_range_chow1 = year_range[:12]
year_range_chow2 = year_range[12:]

for ts in tsptlist:
    if ts == period_pt:
```

```
        print "Sampling for entire time domain..."
        yearlist_period = []
        for year in year_range:
            yearfile = "H:\\dissertation\\data\\weather\\total_precipitation\\growing\\" + year +
".tif"
            yearlist_period.append(yearfile)
        inRasters = yearlist_period
        outTable = weather_path + "period_tpgs.dbf"
    if ts == chow1_pt:
        print "Sampling for chow1 domain..."
        yearlist_chow1 = []
        for year in year_range_chow1:
            yearfile = "H:\\dissertation\\data\\weather\\total_precipitation\\growing\\" + year +
".tif"
            yearlist_chow1.append(yearfile)
        inRasters = yearlist_chow1
        outTable = weather_path + "chow1_tpgs.dbf"
    if ts == chow2_pt:
        yearlist_chow2 = []
        print "Sampling for chow2 domain..."
        for year in year_range_chow2:
            yearfile = "H:\\dissertation\\data\\weather\\total_precipitation\\growing\\" + year +
".tif"
            yearlist_chow2.append(yearfile)
        inRasters = yearlist_chow2
        outTable = weather_path + "chow2_tpgs.dbf"
    inPoint = ts
    resamplingType = "NEAREST"

    Sample(inRasters, inPoint, outTable, resamplingType)

# 1e. Previous yearly average of precipitation
print "\n1e. Previous yearly average of precipitation"

year_range_int = list(xrange(1985, 2012))
year_range = map(str, year_range_int)
year_range_chow1 = year_range[:12]
year_range_chow2 = year_range[12:]

for ts in tsptlist:
    if ts == period_pt:
        print "Sampling for entire time domain..."
        yearlist_period = []
        for year in year_range:
            yearfile = "H:\\dissertation\\data\\weather\\total_precipitation\\previous\\" + year +
".tif"
            yearlist_period.append(yearfile)
        inRasters = yearlist_period
        outTable = weather_path + "period_tppyear.dbf"
    if ts == chow1_pt:
        print "Sampling for chow1 domain..."
        yearlist_chow1 = []
        for year in year_range_chow1:
            yearfile = "H:\\dissertation\\data\\weather\\total_precipitation\\previous\\" + year +
".tif"
            yearlist_chow1.append(yearfile)
        inRasters = yearlist_chow1
        outTable = weather_path + "chow1_tppyear.dbf"
    if ts == chow2_pt:
        print "Sampling for chow2 domain..."
        yearlist_chow2 = []
        for year in year_range_chow2:
            yearfile = "H:\\dissertation\\data\\weather\\total_precipitation\\previous\\" + year +
".tif"
            yearlist_chow2.append(yearfile)
        inRasters = yearlist_chow2
        outTable = weather_path + "chow2_tppyear.dbf"
    inPoint = ts
    resamplingType = "NEAREST"

    Sample(inRasters, inPoint, outTable, resamplingType)

print "\nScript complete."
```

3.5. Python ArcPy script to assess the classification accuracy of the Fmask software.

```
# ---------------------------------------------------------------------------
# randompt_check.py
# Created on: 2013-05-09 11:44:25.00000
#   (generated by ArcGIS/ModelBuilder)
# Created by: TJ Young (SPRI)
# Description: This script was written to create a set of randomly-generated
#           points to assess the classification accuracy of the fmask
#           software.
# ---------------------------------------------------------------------------
print "Running script 'randompt_check.py'..."

# Import arcpy module
import os, sys, shutil
import arcpy

from arcpy import env
from arcpy.sa import *

# Check out any necessary licenses
arcpy.CheckOutExtension("spatial")

# Overwrite pre-existing files
arcpy.env.overwriteOutput = True

yearlist_int = [1990, 1991, 1992, 1993, 1995, 1996, 1997, 1998, 1999, 2000,
             2001, 2002, 2004 ,2005, 2006, 2007, 2008, 2009, 2010, 2011]
yearlist_str = map(str, yearlist_int) # Converts years from integer to string

env.workspace = "H:\\dissertation\\data\\randpts\\"  # Set environment workspace

for year in yearlist_str:
    print "\nCreating random points for year: " + year

    Path_fmask = "H:\\dissertation\\data\\fmask\\" + year + "\\"
    Path_data = "H:\\dissertation\\data\\randpts\\"   + year + "\\"
    Path_scratch = "H:\\dissertation\\scratch\\randpts\\"   + year + "\\"

    # Creates output folder by year (if it does not already exist)
    if not os.path.exists(Path_data):
        os.makedirs(Path_data)
    if not os.path.exists(Path_scratch):
        os.makedirs(Path_scratch)

    # Creates a list of contents in the input directory
    dircontents = os.listdir(Path_fmask)
    inPath_list = list(dircontents) # list contents in specified data folder
    inPath_listTIF = []
    for fname in inPath_list:
        fname_front, fname_end = os.path.splitext(fname)
        if fname_end == ".tif":
            inPath_listTIF.append(fname) # only adds TIF files to inPath_listTIF

    # Runs tool referencing every TIF file in specified year
    pt_genlist = []
    for fname in inPath_listTIF:
        outPts = os.path.basename(fname[4:16]) + "_randpts_clean.shp"
        conFC = ""
        fcExtent = Path_fmask + fname
        numPoints = 150

        arcpy.CreateRandomPoints_management(Path_scratch, outPts, conFC, fcExtent, numPoints)

        # Adds X and Y coordinates to point files
        inPts = Path_scratch + outPts
        arcpy.AddXY_management(inPts)

        print "Successfully created " + str(numPoints) + " geo-labelled random points to be used in
year " + year + " for file: '" + fname + "'."

    # Extract values of fmask rasters to set points
```

```
    print "\nExtracting Fmask values to set points..."
    for fname in inPath_listTIF:
        inPts = Path_scratch + os.path.basename(fname[4:16]) + "_randpts_clean.shp"
        inRaster = Path_fmask + fname
        outPts = Path_scratch + os.path.basename(fname[4:16]) + "_fmaskpts.shp"
        interpolate = "NONE"
        attributes = "ALL"

        ExtractValuesToPoints(inPts, inRaster, outPts, interpolate, attributes)

    # Remove NoData points
    print "\nRemoving points with NoData..."
    for fname in inPath_listTIF:
        inPts = Path_scratch + os.path.basename(fname[4:16]) + "_fmaskpts.shp"
        outPts = Path_data + os.path.basename(fname[4:16]) + "_randpts.shp"
        where = "RASTERVALU < 5"
        arcpy.Select_analysis(inPts, outPts, where)

    # Delete scratch files
    print "\nDeleting intermediary files in scratch..."
    inData = Path_scratch
    dataType = ""
    arcpy.Delete_management(inData, dataType) # Deletes folder for year

    print "\nSuccessfully processed all random point shapefiles in year: " + year + "."

# Delete scratch randpts folder
inData = "H:\\dissertation\\scratch\\randpts\\"
dataType = ""
arcpy.Delete_management(inData, dataType)

print "\nScript complete."
```