Quantifying decadal changes in Arctic lake ice phenology

Tereza Šmejkalová

May, 2014

Quantifying decadal changes in Arctic lake ice phenology

by

Tereza Šmejkalová

Thesis submitted to the University of Southampton, UK, in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialisation: Environmental Modelling and Management

Project Supervisors:

Dr. Jadu Dash
Professor Mary Edwards

UNIVERSITY OF
Southampton

DECLARATION OF AUTHORSHIP

I, Tereza Šmejkalová, declare that the thesis entitled “**Thesis Title**” and the work presented in the thesis are both my own and have been generated by me as the result of my own scholarship. I confirm that:

- This work was done wholly while in candidature for a Masters degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others accreditation has always been given.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.
- It parts of this work have been published, they are listed below.

1. I have read and understood the University’s Academic Integrity Statement for Students, including the information on practice to avoid given in appendix 1 of the Statement and that in this thesis I have worked within the expectations of this Statement.
http://www.calendar.soton.ac.uk/sectionIV/academic-integrity-statement.html

2. I am aware that failure to act in accordance with the Academic Integrity Statement for Students may lead to the imposition of penalties which, for the most serious cases, may include termination of programme.

3. I consent to the University copying and distributing my work and using third parties to verify whether my work contains plagiarised material. If a paper copy is also required for submission it must be identical to this electronic copy. Any discrepancies between these two copies may be considered as an act of cheating.

Signed .... Date 27th May, 2014

Disclaimer

This document describes work undertaken as part of a programme of study at the University of Southampton. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the University.

Abstract

As a major component of the landscape, lakes play an important role in the Arctic. The ecological and thermal processes, including gas emissions (CH_4, CO_2) are strongly dependent on the seasonal freeze-thaw cycle. The phenology of the lake ice (i.e. timing of the freezing cycle) is in large part controlled by the variation in climate. . Air temperature, especially, has been determined to be the major controlling factor of break-up and freeze-up timing. Lake ice phenology is therefore considered a robust indicator of both long term changes and short term variability in regional climate.

The aim of this research is to develop an automated method to derive the timing of phenological variables such as start and end of the freeze and break-up periods from remote sensed data. These dates were derived for lakes larger than $1km^2$ in five study areas distributed evenly over the Arctic to capture the variation in local climatic conditions. Newly developed New Arctic Lake Geodatabase (NALGDB) in combination with the Global Lake and Wetland Database (GLWD) was used to locate the lakes. 13 years of time series of daily surface reflectance data at 250m spatial resolution derived from the Moderate Resolution Imaging spectroradiometer (MODIS) was used to extract the lake ice phenology. The dates for the end of break-up (BUE) and end of freeze-up (FUE) were validated against *in-situ* observations. Results for FUE were not robust enough due to very low sun illumination during the start of the freeze-up period resulting in limited data availability.

For BUE the results are more encouraging and were strongly correlated with the *in-situ* data (R^2 0.65, RMSE 6.16). Trends were analysed for all study areas showing shift towards earlier break-up (BUE) ranging from average -0.89 days/year for Northern Europe to -1.10 days/year for area south of Lake Taymir in northern Russia. Derived ice phenology was also related to various climatic and non-climatic factors such as daily air temperature, precipitation, snow depth, wind speed, and lake size. Mean Temperature over 45 days before the break-up explained up to 60% of observed variability in break-up dates. Deeper understanding of how various factors affect the timing of ice phenological events will help to predict the effect of ongoing climate change on the Arctic.

Key words: Arctic lakes, ice phenology, climate change, remote sensing, MODIS

Acknowledgements

First and foremost I would like to express my gratitude to Dr. Jadu Dash and Professor Mary Edwards, my research supervisors, for their patient guidance, encouragement and useful critiques of this research work.

I am grateful to the GEM Consortium for giving me the opportunity to participate in this programme and the Erasmus Mundus project for financial support that allowed me to take it.

Further thanks belongs to all the professors and teachers at ITC for their support during the first year of the MSc programme.

I would also like to thank my family, without whose support I could not accomplish my goals, not just in recent years but throughout my life.

Last but not least, I great thanks my fellow GEM students for words of encouragement when they were needed. And all the other wonderful people I have met during these two years, both at ITC and in Southampton.

Table of Contents

List of figures

List of tables

List of abbreviations

AO	Arctic Oscillation
AVHRR	Advanced Very High Resolution Radiometer
BUE	break-up end
BUS	break-up start
CIS	Canadian Ice Service
CLIMo	Canadian Lake Ice Model
CRCM	Canadian Regional Climate Model
DMSP	Defence meteorological satellite program
DMSP OLS	Defence Meteorological Satellite Program Operational Linescan System
ECMWF	European Centre for Medium-range Weather Forecast
ENSO	El Niño-La Niña/Southern Oscillation
ERS2	European Remote-Sensing Satellite 2
FO	freeze onset
FU	freeze-up date
FUE	freeze-up end
FUS	freeze-up start
GCM	Global Climate Model
GMS	Geostationary Meteorological Satellite
GOES	Geostationary Operational Environmental Satellites
GOESS	Geostationary Operational Environmental Satellite
HIGHTSI	High Resolution Thermodynamic Snow and Ice
IMS	Interactive Multisensor Snow and Ice Mapping System
IPCC	International Panel for Climate Change
LIMNOS	Lake Ice Model Numerical Operational Simulator
MAD	median absolute error
MAE	mean absolute error
MODIS	Medium-Resolution Imaging Spectroradiometer
NAO	North Atlantic Oscillation
NIR	near infrared
NOAA	National Oceanic and Atmospheric Administration
NSIDC	National Snow and Ice Data
PDO	Pacific Decadal Oscillation
PNA	Pacific North American pattern
POES	Polar-orbiting Operational Environmental Satellites
SAR	Synthetic Aperture Radar

SHMI	Swedish Meteorological and Hydrological
SYKE	Finish Environmental Institute
TOA	top-of-atmosphere
VISSR	Visible Infrared Spin-Scan Radiometer
WCI	water clean of ice

1 Introduction

Ice sheets, ice shelves, glaciers, ice caps, sea ice, river and lake ice, snow cover and permafrost together form an important part of the Earth system we call cryosphere. The main driver shaping the cryosphere is the climate, making it particularly susceptible to ongoing climate change. Permafrost thawing, loss of the sea ice, shrinking of glaciers, longer ice-off periods on lakes and rivers are only few processes recognized as indicators of the changing climate. (IPCC, 2013) Cryosphere, in return, significantly affects the regional and global climate, either directly or in many cases indirectly, through complex negative and more often positive feedbacks.

It is estimated that around 8 000 000 lakes larger than 1 ha exist on Earth and account for approximately 2% of global land area, this number can increase up to 2.4% when smaller lakes are included. Vast majority of lakes can be found north of 40° latitude. (Lehner & Döll, 2004) It is, therefore, safe to conclude that the majority of Earth's lakes seasonally freeze, at least to a certain degree.(Kirillin et al., 2012) The timing, duration and thickness of seasonal ice cover is major determinant of all the chemical, biological and ecological processes occurring in the lake, especially in northern latitudes, where ice cover can last for better part of the year. (Duguay et al., 2003) In the Arctic lakes occupy 15% to 40%, and at extremes even 90%, of the landscapes and thus the effect they have on regional and even global climate is considerable.

Significant shifts toward later freeze and earlier break-up have been observed around the Northern Hemisphere, resulting in longer open water conditions. Longer ice free season brings higher water temperatures, increased primary productivity and overall change of ecosystems within the lake.(Callaghan et al., 2012) Northern lakes also act both sink and source of atmospheric greenhouse gasses mainly methane. It is expected, that with longer ice free duration and more heat absorbed by the water, the methane production and release to atmosphere will rise, which could lead to overall significant increase in global methane budget and further warming. (Callaghan et al., 2012; Walter et al., 2007)

Monitoring of the changes in the timing of start and end of the ice-on season and deeper understanding of how various factors affect the timing of the freeze-up and break-up is crucial to predict the effect of ongoing climate change on the Arctic ecosystems.

1.1 Arctic lakes

The Arctic can be delineated in several different ways. With the exception of Arctic circle (66°33’44” N), most of the defined boundaries, such the 10° July isotherm, tree line or southern extent of discontinuous permafrost, are strongly influenced by the climatic conditions and are likely to move due to the effect of climate change.(Vincent & Laybourn-Parry, 2008) Regardless of which definition is used, lakes and ponds are prominent feature shaping the Arctic areas of Europe, Asia and North America. They occupy depressions of various geomorphologic and geologic origins. (Woo, 2012) According to Smith et al. (2007) the abundance of lakes is most dependent on the presence of permafrost and glaciation history, with the largest concentration in glaciated permafrost peatlands (14.4 lakes/1000 km^2) and the lowest in unglaciated, permafrost-free areas (1.2 lakes/1000 km^2). During the Pleistocene epoch repeated glaciation periods resulted in postglacial landscape, richly endowed with lakes occupying the depressions carved into bedrock by the moving ice masses.

The most common lake type in the Arctic are the thermokarst or thaw lakes. According to Walter et al. (2006) 90% of lakes in the Russian Yedoma zone (Pleistocene-age organics rich permafrost) are of this type. Their evolution occurs in cycles, Jorgenson & Shur (2007) identified six development stages: 1) initial flooding of depressions in melting permafrost in sandy soils of degradation of ice wedges in thick silt soils, 2) lateral erosion and expansion, and sediment redistribution and sorting, 3) lake drainage (full or partial), 4) ice aggradation in silty centres and sandy margins, 5) formation of secondary lakes and ponds, 6) basin stabilization. Appart from Yedoma, they are widespread on Northern Slope of Alaska, and in the lowlands of Northern Canada and Siberia.

Another lake type frequently present in in the Arctic especially in Greenland are ice-dammed lakes. They develop close to glacier fronts in mountainous terrain, or beside or on ice sheets. (Pienitz et al., 2008) Many lakes among the coast are product of glaciostatic uplift processes that follow the retreat of the glacier mass. In some areas such as the Canadian Hudson bay this uplift is still ongoing today at rate approximately 1 meter/century. During this process the offshore bars rose transforming lagoons to coastal lakes. (Woo, 2012) Flood plain lakes and former meanders occur in deltas of most Siberian and Canadian rivers (Lena, Mackenzie). Volcanic, tectonic and meteorite impact crater lakes are also found in the Arctic, however, they are much less frequent.

In areas with continuous permafrost, lakes of all origins are often subject to thermokarsting processes. By permafrost erosion and thaw on lake margins, organic matter is introduced to the lake bottom and broken down by anaerobic bacteria. The bacteria release methane which is consequently introduced to atmosphere through the process of ebullition. (Walter et al., 2007) The length of ice free season determines the heat accumulated in the lake and therefore, the magnitude of permafrost degradation at its margins, related methane production and other biochemical processes. (Callaghan et al., 2011)

1.2 Lake ice phenology

The word phenology is derived from the Greek *phaino* meaning "to appear". It is defined as the study of periodically recurring natural events influenced by the environmental variables, especially temperature and its changes driven by weather and climate conditions. (Schwartz, 2003) The Oxford Dictionaries define phenology similarly as "the study of cyclic and seasonal natural phenomena, especially in relation to climate and plant and animal life." (Simpson, 2002)

Lake ice phenology studies the timing of annual ice formation and decay and its changes. Kropáček et al. (2013) and Latifovic & Pouliot (2007) define four main phenological events. Although the terms are different in both studies, they refer to identical events. 1) The date when ice is first detected is referred to in the studies as *freeze onset* (FO) or *freeze-up start* (FUS) respectively. 2) The *freeze-up date* (FU) or the *freeze-up end* (FUE) is defined as the date when there is no longer any detectable open water on the lake. 3) The date when ice-free water appears is referred to as *break-up start* (BUS). 4) When the water is in the spring first completely free of ice, it is denoted as *water clean of ice* (WCI) or *break-up end* (BUE) or just *break-up date* (BU). In this study the second set of terms (FUS, FUE, BUS, BUE) by Latifovic & Pouliot (2007) will be used. Most studies use only the terms *freeze-up* and *break-up,* however, they are not united in their definition. *Freeze-up* mostly refers to FUE but *break-up* is used in relation to both, BUS and BUE, almost equally often. Here, if event is referred to as *break-up* it is meant BUE, except when reference to another work is made then original definition is adopted. The time span between FUE and BUS defines the ice-on duration (the time when the lake is completely covered by ice) and therefore would be the most sensible combination of events to study. However, mots available *in-situ* data sources such as The Global Lake and River Ice Phenology Database, the Swedish Meteorological and Hydrological Institute (SHMI) or the Finish Environmental Institute (SYKE) only record timing of FUE and BUE events though with yet another terminology (ice-on and ice-off, freeze and thaw). (NSIDC, 2012)

1.3 The role of lake ice

Vast majority of processes occurring in and around Arctic lakes is governed by seasonal ice cycle. The duration and thickness of ice cover affects, among others, the water temperature and thermal stratification, light penetration, nutrient supply and the overall dynamics of the phytoplankton within a lake. (Blenckner et al., 2010) Shifts toward shorter ice cover duration are expected to cause increase in primary productivity and, in consequence, changes in trophic relationships within a lake. (Prowse et al., 2011a; Vincent et al., 2011) The effect of lake ice cover duration and timing, however, extends far beyond the borders of the lake. As suggested by the word itself, the interaction between lake ice and climate is not one way process.

1.3.1 Feedbacks to climate

Due to the enormous areas occupied by lakes, they play important role in shaping the local and regional climate. Some feedbacks to climate are direct, others are result of complex processes. Especially, large lakes have been shown to extend substantial moderation effect on local climate. During the summer open water phase the lake acts as heat sink absorbing the solar radiation and effectively cooling its surroundings in the process. Conversely in during the winter season the lake can become a heat source. (Brown & Duguay, 2010)

During winter period the solid ice acts as a lid on the lake, stopping evapotranspiration and direct heat exchange with the atmosphere, leaving energy exchange only through radiative and conductive processes. (Brown & Duguay, 2010) Shift toward longer open water phase will cause rise in water temperatures and consequently evaporation, resulting in higher water vapour concentrations in the atmosphere. While water vapour can have cooling effects on local scale it can induce warming over larger areas due to its mixing and transport in the atmosphere. (Callaghan et al., 2011) One of more direct feedbacks to climate is the contribution of Arctic lakes to regional and global greenhouse gas budget. As mentioned above, Arctic lakes have been observed to be both sink a source of carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O). During winter ice cover these gasses, especially CH_4, accumulate under the ice and are released during the spring melt in a large pulse. Higher water temperature will promote permafrost thawing on the lake margins and therefore cause increase in the amount of CH_4 released to the atmosphere. (Brown & Duguay, 2010)

1.4 Trends

Evaluation of historical spatial and temporal trends in lake ice phenology around the Northern Hemisphere has proven difficult to quantify. This is due to various reasons such as the use of different phenological events or rather different terminology or absence of metadata for older records, different periods for which records are available (many not continuous or years missing) and changes in the observation methods from *in-situ* over visual interpretation of remote sensing imagery to development of automated extraction methods. (Brown & Duguay, 2010; Prowse et al., 2011b) Moreover, *in-situ* records are generally determined by point observations, which can represent yet another issue, as entire lake is rarely visible form single point. However, most studies, short term or long term, regional or global, agree on overall trend leading to later freeze-up and earlier break-up dates and general shortening of ice-on period. Many of recent studies evaluating trend in lake ice phenology on the Northern Hemisphere are summarized in Prowse et al. (2011b), Walsh (2005) and Brown & Duguay (2010). (Figure 1)

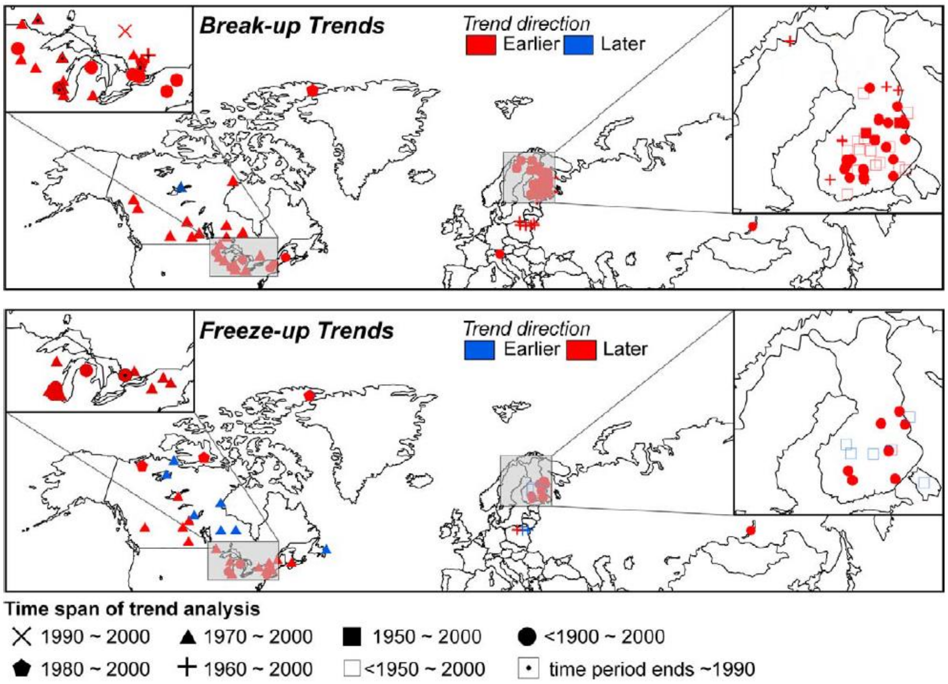

Figure 1: Trends for Northern Hemisphere identified as statistically significant based on studies conducted before 2010 (Brown & Duguay, 2010)

Very small number of long-term records exists for either freeze-up or break-up dates and they are mostly limited to lower latitudes.

Magnuson et al. (2000) analysed thirty-nine lake and river ice phenology series (freeze-up or break-up) available for the period from 1845 to 1995 (150 years) for 26 lakes around the Northern Hemisphere. Some of the analysed series contain records beginning before the year 1800 (Japan, Russia and Finland). 38 out of 39 series show trends toward later freeze-up or earlier break-up, with the exception of 550-year series of lake Suwa (Japan) where the trend did not prove significant. For the period from 1846 to 1995 the overall observed rate of change of the freeze-up date (FUE) was +5.8 days/100 years (± 1.9 days) and -6.5 days/100 years (±1.4 days) for break-up date (BUS) corresponding to increase of 1.2 °C/100 years in air temperature. The study did not identify any statistically significant spatial trends or differences between the slope of change for break-up and freeze-up dates.

Sharma et al. (2013) investigated temporal patterns in ice break-up dates for two lakes from the same dataset (Mendota and Monona in Wisconsin) over an updated 100-year period from 1905 to 2004. Statistically significant linear trends were observed with rate of change slightly lower (-6.7 and -12 days/100 years, respectively) than observed in Magnuson et al. (2000)(-7.5 and -12.2 days/100 years), which can be attributed to the difference in studied period. Benson et al. (2012) have extended on the work of Magnuson et al., (2000) in several ways. Greater number of lakes (75) was analysed, some records were corrected and the database was updated until 2004. Unlike in the previously mentioned studies, the trends were analysed separately for three periods (150-, 100- and 30-year) as well as for four regions. In all regions and on all time scales the trend was, as expected, toward shorter ice duration. The steepest slopes were observed for the 30-year (1975-2004) period break-up dates for European lakes (Scandinavia and Switzerland) with the value of 0.29 days a year and in Northeastern North America with the value of 0.25 days a year. The lowest rates of change were interestingly also observed for the last 30 years in North central North America (0.03 days/year). The slopes for the 150-year and 100-year periods range from 0.05 to 0.06 days /year. None of the mentioned studies, however, explores the spatial trend in south to north direction and although the last analysis included Arctic lakes in Scandinavia, the results were not discussed.

Several studies that concentrated partly of fully on Arctic lakes show contradicting results. Although they agree on the general trend of shortening of the ice duration, they show differences in rates of change in relation to southern latitudes. Analysis of 30-year long records (1960 - 1991) for Swedish lakes identified trends toward earlier break-up more pronounced in southern parts (-0.92 days/year) than in northern

Sweden (-0.25 days/year). (Weyhenmeyer et al., 2005). Similarly, analysis of long term records of Finish lakes (some beginning in early 19th century) revealed significant shortening of ice-on period for the southern and central areas, however, no significant trends were found in the far north for the period of 1960 to 2002. (Korhonen, 2006) These findings are furthermore supported by the assumption that smoothed air temperature is well described by an arccosine function. The form of arccosine function suggests that response to equal change in air temperature will be stronger at latitudes with shorter duration of ice cover (0-125) opposed to higher latitudes with long duration of ice cover (200-250). In Sweden the sensitivity of ice break-up was estimated to ≈ 14 days per 1°C in South as opposed to ≈ 4 days per °C in the North. (Weyhenmeyer et al., 2004) On contrary the study based on AVHRR imagery made by Latifovic & Pouliot (2007) shows considerably higher average change in FUE (+0.76 days/year) and BUE (- 0.99 days/year) and for the lakes in the far north than for lower latitude lakes (0.23 days/year and 0.16 days/year respectively) for the period of 1985 to 2004. Several studies that concentrated purely on one or more lakes in far North have been conducted in Northern Europe. Lei et al. (2012) have examined 44 year record for Lake Kilpisjärvi (69.05°N, 20.83°E) in Northern Finland and their relation to air temperature. Significant trend toward later freeze-up (0.23 days/year) and shorter ice duration (0.32 days/year) was observed. The break up occurred 1 day/decade earlier over the studied period, however, this trend has not proven significant. Recently, Surdu et al. (2014) analysed trends for shallow tundra lakes on the North Slope of Alaska over last six decades (1950 - 2011). The results show shift toward shorter ice-on duration by 24 days as result of later freeze-up by 5.6 days (0.09 days/year) and earlier break-up by 17.7 – 18.6 days (0.3 days/year). This study is the first to analyse the response of small shallow Arctic lakes to changing climate.

Phenology series for the lakes in far north are mostly limited to last 50 years and earlier observations are usually not available due to remoteness of the areas. This makes judging the long terms in the Arctic difficult and complicates comparison with the development in southern latitudes. The different rates of change can be attributed to unequal changes in air temperature between the south and north regions as well as on continental scale.

Generally break-up dates show higher rate of long- long term change then freeze-up dates due to higher dependency of break-up on the air temperature. (Jensen et al., 2007) Benson et al. (2012) observed slight shift toward earlier freeze-up date for some lakes especially in Northern Europe. The departure from the overall observed trends was assigned to the complex interactions of variables including regional climate oscillations, lake size and morphometry.

The long-term trends explain only approximately one third of the observed variability. Strong multi-decadal, inter-decadal, and inter-annual variability is associated with regional and global climate oscillations. (Benson et al., 2012; Duguay et al., 2006; Blenckner et al., 2007) Against expectations no significant increase in inter-annual variability was observed. There is, however, substantial shift in the direction of extreme events caused by the long-term changes. Benson et al. (2012) observed increase in frequency of extreme event associated with warmer conditions such as no total freeze-up, extremely late freeze-up or extremely early break-up. It was also observed that the inter-annual variability tends to decrease with increasing latitude. The lowest values ±9 days were observed for lakes north of 61°N. (Weyhenmeyer et al., 2011)

1.4.1 Future development

Several studies have attempted to model future developments in lake ice phenology. (Brown & Duguay, 2010, 2011; Dibike et al., 2011, 2012) It was determined that in Northern hemisphere 1°C change in mean air temperature results in 5 days change to the date of phenological event. (Magnuson et al., 2000) When considering the 4°-7°C increase in mean air temperature predicted for the Arctic (Hassol, 2005) over next 100 years, the decrease in mean ice duration is estimated to 40-70 days. However, relationships observed over past periods may not apply for future climatic conditions, thus they may not reliably portray the future changes and more complex predictive methods are needed. (Bonsal & Prowse, 2003; Prowse et al., 2011b)

There have only been few analyses based on predictive ice phenology models integrated with Global Climate Model (GCM) outputs. Dibike et al. (2011) used the one-dimensional MyLake model to simulate the development of lake ice freeze-up and break-up around northern Hemisphere (40° to 75° latitudinal band) under changing climate conditions. The changes were simulated on hypothetical lakes placed in the grid 2.5° latitude and longitude for future climatic conditions (2040-2079) based on ERA-40 global reanalysis dataset modified according to Canadian Global Climate Model. Results indicated continuing trend of overall decrease in ice duration. The delay in freeze-up is estimated to 5-20 days and break-up is predicted to advance by 10-30 days, resulting in decrease of ice cover duration by 15-50 days compared to current situation (1960-1999). Brown & Duguay (2011) investigated the fate of North American Arctic lake ice through application of the Canadian Lake Ice Model (CLIMo). Lake ice phenology was simulated using two scenarios based on Canadian

Regional Climate Model (CRCM) for 1960 to 2100. Results for the period 1960 to 1990 were validated against *in-situ* observations of 15 Canadian Arctic lakes. The projected differences between the mean values of current 30-year periods of 1961-1990 and 2041-2070 show 10-25 days advance in break-up and 0-15 days delay for freeze-up in most areas. The overall reduction of ice cover is estimated to be 10-25 days for shallow lakes and 10-30 for deeper lakes. The results from both models are comparable in magnitude. However, slightly higher change predicted by MyLake for the Northern Hemisphere could suggest that, against expectations, change will be higher in mid latitudes compared to Arctic (results form CLIMo).

1.5 The determinants of ice formation and decay timing

The timing of ice phenology events is the result of complex set of climatic variables, location, elevation and morphological variables such as size and depth. The lake ice formation is determined by a moment when the heat loss at the surface of the lake exceeds the heat gained from solar radiation and convection in the lake. (Woo, 2012) The most significant differences in timing occur along the latitudinal gradient, as expected. Latitude is often used as proxy for the amount of shortwave solar radiation received at the surface. Closely linked to solar radiation is the air temperature which is considered the main determinant of variability in timing of phenological events. Next to the temperature, precipitation (snowfall) and wind speed, the lake morphology, especially size and depth of the lake, plays important role. The volume of the lake determines the heat storage capacity and in this way, the timing of freeze-up start. Large and deep lakes generally remain ice-free longer than smaller or shallower lakes at the same latitude and altitude. Ménard et al. (2002) even showed depth to be a determinant of freeze-up for Canadian lakes using thermodynamic model. The thickness of ice cover also varies slightly based on depth being thinner for deeper lakes. (Rouse et al., 2008) The decay of ice is less dependent on the lake morphological parameters and is determined rather by temperature patterns and local weather conditions, as well as the volume of inflow to the lake. Further the effects of temperature, precipitation, wind and large scale atmospheric circulation are discussed in detail.

1.5.1 Temperature

It has been well documented that temperature is the key variable that defines the lake ice phenology. Temperature during the months preceding the phenological event is often able to explain up to 60% - 70% of the variance in ice-on and ice-off timing. (Palecki & Barry, 1986; Robertson et al., 1992; David M. Livingstone et al., 2010) In past studies two main methods of relating ice freeze-up/break-up to air temperature at are used: 1) air temperature integrated over fixed period of time (strong correlation was observed for air temperature integrated over a 1 to 3 months before the event) 2) relation to the 0°C isotherm . Palecki & Barry, (1986) examined the ice-on and ice-off dates of four lakes in Northern Finland and discovered highest correlation values (0.8 and -0.8) for temperatures integrated over a period from October to November and the month of May, respectively. Commonly used freezing/thawing degree days method is based on integration of only negative/positive temperatures over given time period preceding freeze/thaw. Also simpler degree-day method is often used. It is based on analysis of long-term records around Northern Hemisphere and their relation to the increase in mean air temperature over the period recorded. It was estimated, that rise or decrease of 0.2°C results in 1 day change in mean date of phenological event.(Magnuson et al., 2000)

Linear response to air temperature was questioned by Weyhenmeyer, Meili, & Livingstone (2004). The study, conducted for 196 Swedish lakes with varying size and depth, shows nonlinear relationship between ice break-up and mean annual air temperature. To first approximation it is possible to express the annual cycle by a sinusoid characterized by an annual temperature mean (T_m) and amplitude (T_a). The duration of a period when temperature falls below 0°C can be then approximated as:

$$D \approx \left(\frac{1}{\pi}\right) arccos\left(\frac{T_m}{T_a}\right) \tag{1}$$

However, the arc cosine model is valid only for locations where the period of below freezing temperatures lasts between approximately 55 to 310 days. Assuming that the ice-on and ice-off days correspond linearly with the beginning and end of a period when the smoothed air temperature falls below 0°C, (supported by Duguay et al., (2006)) an equation defining the day of year when break-up occurs was developed taking into account regional temperature patterns in Sweden. Later it was adapted by Weyhenmeyer et al. (2011) on a set of 143 lakes around the Northern Hemisphere for period of 1961 to 1990. To increase the explanatory power dependence on latitude was tested as a proxy to solar radiation. The residuals for ice-on days showed strong relation to latitude unlike the ice-on dates. The timing of a lake break-up, freeze-up and ice duration can be then expressed as:

$$t_{ice-on} = \left(\frac{365.25}{2\pi}\right) arccos\left(\frac{T_m}{T_a}\right) + 55 \qquad (2)$$

$$t_{ice-off} = \left(\frac{365.25}{2\pi}\right) arccos\left(\frac{T_m}{T_a}\right) + (2\varphi - 55) \qquad (3)$$

$$D_{ice} = \left(\frac{365.25}{2\pi}\right) arccos\left(\frac{T_m}{T_a}\right) + (2\varphi - 110) \qquad (4)$$

Where T_m represents site-specific mean temperature and T_a amplitude in °C. ϕ is the latitude and together with the constants 55 or 110 gives the offset in days. The constant 365.25 stands for the length of year in days. The equations were able to explain 79% and 81% of variation in the mean ice-off dates and ice duration respectively, but only 28% of variation in mean ice-on dates. Better results were achieved earlier by S. E. Walsh et al. (1998), however, with the use of complex set of variables such as solar radiation, precipitation, humidity, long wave radiation, wind speed and lake morphology. This can be seen as evidence of lower dependency of freeze-up on temperature patterns.

1.5.2 Precipitation and wind

Apart from temperature precipitation either in the form of rain or snowfall have significant effect on the timing of phenological events. Increased snowfall in early winter mixing with cold surface water may form a snow slush layer and significantly advance the freeze-up timing. Once the ice cover has been established snow cover acts insulator retarding the heath exchange between air and ice and slowing the ice growth on water ice-interface. (Woo, 2012) However, in case cracks develop in the ice, the pressure of the overlying snow causes water to penetrate to snow-ice interface resulting in white ice (snow ice) formation and thus increase in ice cover thickness.(Adams & Roulet, 1980) During the spring period the insulating properties of the snow layer and its high albedo compared to black ice tend to delay the break up timing. According to Duguay et al. (2003) snowfall is responsible for much of the variation in thickness and break up timing. In case of rainfall, we can expect opposite effect.

Another climatic variable that plays important role is the wind. In early winter strong winds can prevent the formation of solid ice cover and delay the timing of complete freeze up. On the other hand once solid ice is formed the cooling effect of wind can positively influence the ice growth. During the break up period wind accelerates the decay of ice cover both mechanically and through increased heat exchange between ice and air. (Brown & Duguay, 2010; Woo, 2012) Blenckner et al. (2004) observed that variance in break up dates in Northern Scandinavia is strongly determined by direction and strength of zonal winds in January and May given by the phase of North Atlantic Oscillation.

1.5.3 Large-scale atmospheric circulation patterns

The large-scale atmospheric and oceanic oscillations are considered as important forcing factor of inter-annual variability air temperature and thus lake ice phenology. (Livingstone, 2000) The largest teleconnection patterns influencing the temperature and precipitation around the northern Hemisphere are North Atlantic Oscillation (NAO), Arctic Oscillation (AO), El Niño-La Niña/Southern Oscillation (ENSO), Pacific North American pattern (PNA), and Pacific Decadal Oscillation (PDO).

According to Hurrell (1996) the NAO and AO are associated with almost a third of inter-annual variance in winter and spring temperatures over Northern Europe, Northwest Asia and East Canada. The NAO index is defined as the difference between normalized Icelandic Low and Azores High pressures. The positive phase (representative of deepened Icelandic Low and stronger Azores High) is associated flow of warmer air masses from the Atlantic to Northern Europe causing warmer winters in North Eurasia and the flow of cold Arctic air to Canada and thus colder than normal conditions. (Figure 2) For the negative phase, the situation is opposite. (Girjatowicz, 2010) Significant correlation of the NAO/AO phase and the ice phenology has been detected for lakes in Scandinavia, and North-west Russia. (Blenckner et al., 2004; Blenckner et al., 2007; Girjatowicz, 2010; Livingstone, 2000; Yoo & D'Odorico, 2002) Livingstone (2000) observed 43% of shared variance between the break-up dates of two Finish lakes and NAO over the period of 1941-1990. Moreover the higher occurrence of positive NAO in last 50 years is consistent with the observed trends towards shorter ice duration in Scandinavia and Western Russia. (Prowse et al., 2011b)

The winter and spring temperature variability over much of North America is strongly determined by the ENSO, PNA and PDO. During the ENSO positive phase (El Niño) and positive phases of PNA and PDO defined by deepened Aleutian Low higher winter temperatures are observed and vice versa. Wang et al. (2012) observed 4-year cycles in lake ice phenology of Great lakes associated with the ENSO. Bonsal et al. (2006) investigated the effects of large-scale teleconnections on freshwater break-up and freeze-up in Canada for the period of 1950-

1999 and observed approximately 10 day shorter ice durations during positive phase opposed to negative phase.

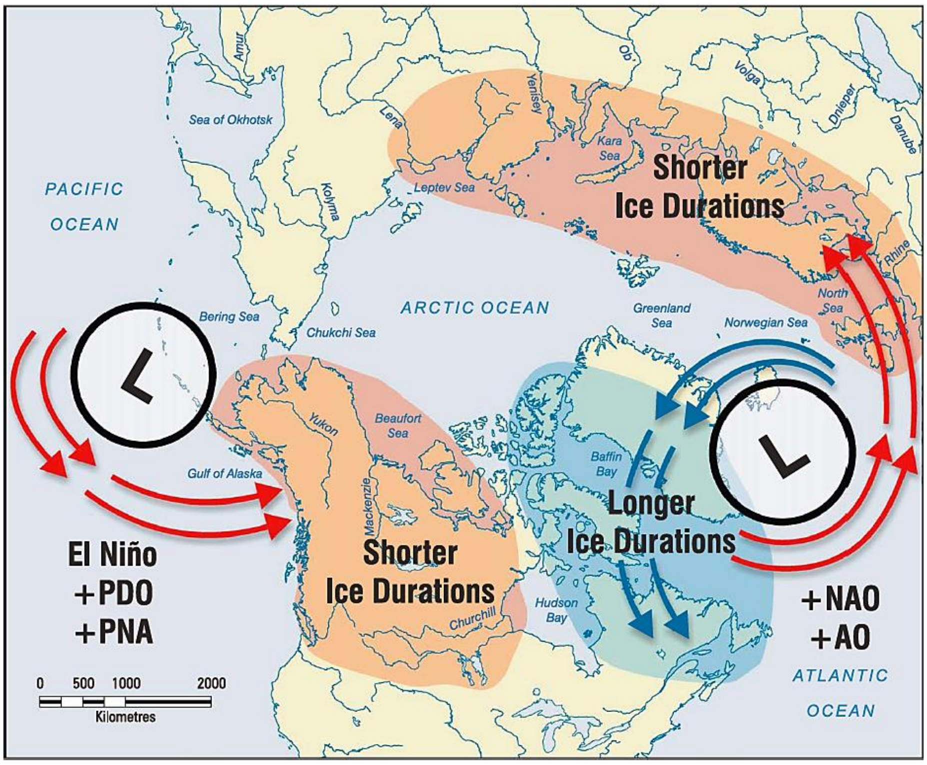

Figure 2: Cold-season (positive phases) atmospheric circulation patterns and the influence on lake ice duration. Negative phases are associated with opposite temperature changes and effects on ice duration. (Prowse et al., 2011b)

1.6 Monitoring and prediction

Historically the thaw and freeze dates of lakes and rivers were recorded for practical purposes such as transportation, or agriculture, for religious or cultural reasons of simply out of curiosity. (Magnuson et al., 2000) Monitoring programs collecting lake ice data in various countries are usually run by the national hydrological or environmental institutes such as the Swedish Meteorological and Hydrological institute (SMHI), Finish Environment institute (SYKE), Canadian Ice Service (CIS) or the National Snow and Ice Data Center (NSIDC). The number of monitored lakes has decreased globally since the beginning of the century. Most datasets are based on *in-situ* observations, currently for example the Canadian volunteer-based observation network IceWatch. However, monitoring in sparsely populated areas of far North observations is logistically challenging and expensive. Digital imagery cameras were found to be a useful tool for unattended ice monitoring.

(Brown & Duguay, 2012) For Canada the number of observation sites reached peak between 1950 and 1980 and decreased dramatically after that to only about 10% of 1980 numbers in 2000. Development of new technologies and methodologies for monitoring lake ice phenology using remote sensing will hopefully fill the gaps in lake ice records caused by the decrease in *in-situ* observations. The future trends of lake ice phenology under the changing climate conditions are also of interest for scientist around the world. Number of thermodynamics models were developed and used to simulate freeze a break-up dates under different climatic conditions. (Dibike et al., 2012; Menard et al., 2002; Mishra et al., 2011b; Walsh et al., 1998)

1.6.1 Remote sensing

Remote sensing provides a very effective and relatively low cost tool for monitoring lake ice over large or remote areas. First attempts to explore the feasibility of remote sensing observations for determining the lake freeze-up and break-up dates were most likely made by Maslanik & Barry (1987). This study, and many following, were based on visual interpretation of imagery from medium resolution optical sensors such as the NOAA Advanced Very High Resolution Radiometer (AVHRR), Defense Meteorological Satellite Program Operational Linescan System (DMSP OLS) or Geostationary Operational Environmental Satellite – Visible Infrared Spin-Scan Radiometer (GOESS-VISSR, ended 1996) that provided reasonable frequency of the observations and spatial resolution. (Palecki & Barry, 1986; Maslanik & Barry, 1987; Leshkevich et al., 1993; Wynne & Lillesand, 1993; Wynne et al., 1998)

Today, visual interpretation is often used with imagery from active remote sensing instruments such as Synthetic Aperture Radar (SAR) supplemented by optical sensor images. For example, the Canadian Ice Service uses mainly RADARSAT ScanSAR images supplemented by AVHRR, ERS2 imagery for monitoring of 135 lakes on weekly basis. (CIS, 2005) Active remote sensing have proven useful for lake ice monitoring at high latitudes where, low sun angle over the winter period and persistent cloud cover, pose a challenge for optical sensors. (Jeffries et al., 2005) Cook & Bradley (2010) used semi-automatic approach to derive ice-on and ice-off dates for two Arctic lakes from Canadian Space Agency RADARSAT-1 satellite. The imagery was first classified to ice and open water, the areas of open water were digitized and from the total lake area the percentage of ice cover was derived on approximately weakly basis. Another study has developed fully automatic method for extraction of lake ice maps from Envisat ASAR data. The method was tested for lake Päijänne, Finland and is based on object-based classification. The validation of algorithm results was done against *in-situ* observations and MODIS snow and ice product. The accuracy of classification was low due to speckle noise and similarity of backscatter from water and ice for certain lake surface states. Moreover the algorithm is suitable for use only in periods when both water and ice classes are present. (Hindberg et al., 2012) In addition the temporal resolution of current SAR sensors limits the use for operational monitoring and climate change studies where accurate estimations are needed.(Latifovic & Pouliot, 2007)

In 1997, the Interactive Multisensor Snow and Ice Mapping System (IMS) was developed and implemented to operational use. It provides the opportunity to combine multiple sources of remotely sensed data (NOAA polar orbiters (POES), NOAA geostationary (GOES) satellites, Japanese geostationary meteorological satellites (GMS), European geostationary meteorological satellites (METEOSAT), US Department of Defence polar orbiters, and Defence meteorological satellite program (DMSP), later also AVHRR, MODIS imagery was added). One of its many products is the daily snow and ice cover for entire Northern Hemisphere with 4km spatial resolution. (Helfrich et al., 2007) Medium resolution optical sensors such as Terra/Aqua MODIS (500m) or NOAA AVHRR (1000m) also offer daily snow and ice cover products but with considerably finer resolution than IMS. Brown & Duguay (2012) investigated the utility of IMS (2004 - 20012) and MODIS (MOD10A1, 2000 - 2012) daily snow and ice products for ice phenology monitoring in Quebec (189 lakes). Overall, the MODIS product performed better which was attributed to finer spatial resolution. The IMS does not provide information for small lakes that are not captured by the relatively coarse 4km resolution, moreover IMS has a tendency towards early freeze-up dates.

Apart from snow and ice products, MODIS and AVHRR also provide daily surface reflectance with 250m and 1000m resolution, respectively. The AVHRR has the longest available data record extending from 1979 to present and offers already more than 30 years of continuous observations. The MODIS sensor was launched considerably later (Terra-1999, Aqua-2002). Latifovic & Pouliot (2007) have used AVHRR daily top-of-atmosphere (TOA) reflectance in the near infrared band (NIR) data to derive four phenological variables for 42 Canadian lakes larger than 100km^2. The near infrared part of spectra is most suitable for ice phenology extraction, due to high water absorption and therefore, higher contrast between ice and open water. In addition, NIR band is less affected by atmospheric conditions compared to red band. Daily reflectance derived from optical imagery is generally used over smaller regions due to large amounts of data and difficulties connected with handling them. The authors, however, designed an automatic extraction method based on lake reflectance profile. The large dataset is then eliminated to a matrix containing reflectance values averaged over the area of each lake sorted by day since beginning of the observation (column) and latitude (row) of the lake. FUS, FUE, BUS and BUE were automatically extracted for each lake from the smoothed profile and compared to *in-situ* records and dates extracted from the profile visually. Agreement between the values based on visual and automatic extraction was high ($r^2 > 0.9$). Even though AVHRR reflectance data show great potential for monitoring lake ice phenology they require significant amount of pre-processing (Latifovic et al., 2005). Moreover, 1km spatial resolution is limiting the use for monitoring small Arctic lakes.

Majority of Arctic lakes has surface area smaller than 2 km^2 and finer spatial resolution is need in order to monitor changes in their phenological cycle. MODIS sensor does not provide equally long time series as the AVHRR, however, it provides spatial resolution of 250m for red and near infrared bands, enabling monitoring of lakes down to 1 km^2 in size. The major limitation in the use of MODIS for ice phenology monitoring is the cloud cover, as it is with any other optical sensor. MODIS sensor, however, is mounted on two polar orbiting satellites TERRA and AQUA with different overpass time. The surface reflectance is produced separately for TERRA and AQUA but it is possible to substitute missing or clouded scene from one with clearer scene from the other. Moreover in the polar areas orbits of the satellite overlap producing multiple observations for each pixel enabling production of daily surface reflectance composites with least cloud cover. Chaouch et al. (2012) applied threshold based decision tree to MOD09GQ NIR band to provide daily rive ice extent maps. The thresholds were derived empirically from frequency distribution analysis of 90 MOD09GQ images in NIR band covering the study area. Pixels were divided between clean river pixels and mixed river/land pixels and for each group two thresholds defining three classes (water, water-ice and ice) were implemented. The results were validated against classified Landsat imagery and shoved ice detection probability of 91%. As expected the accuracy is lower for the mixed land-river pixels due to high reflectance of land that leads to overestimation of ice fraction values. In spite of its advantages MODIS surface reflectance product have not yet been used for lake ice phenology monitoring.

1.6.2 Predictive models

Apart from direct monitoring various models simulating lake ice growth and decay have been developed over the past years. Three main types of lake ice models have been used – empirical and regression models, energy balance models and thermodynamic lake ice models. (Brown &

Duguay, 2012) Lake ice models provide an opportunity to study the effects of ongoing and future climate changes on the duration of lake ice cover and thus on the ecology in lake rich northern areas.

Empirical and regression models such as that used by Palecki & Barry (1986), Livingstone (1997) or Livingstone & Adrian (2009) attempt to define empirical relationship between air temperature and the freeze and break up dates. Some of these methods have been mentioned in section *1.3.1 Temperature*. Generally, some form of integrated temperature over certain time period is used (mean temperature, only negative/positive temperatures over fixed calendar of moving time period). Livingstone & Adrian (2009) developed new method based on air temperature probability function allowing to estimate the duration of ice cover. Unlike other mentioned methods it is valid even for areas where the ice cover duration is not continuous during the winter and does not require specific time period to be defined. The model was able to explain 93.5% of variability in ice cover duration for Lake Müggelsee in Northern Germany. For the other two mentioned studies the shared variance between the temperature variable and ice cover duration is lower (64% - 80%).

Number of one-dimensional thermodynamic models have been developed and used to simulate the future developments in lake ice phenology under various future climate scenarios. Findings of some of these studies have been discussed previously in the section *1.4.1 Future development.* Lake ice cover specific models are often derived for sea ice cover models. Examples of most used models successfully used to model evolution of lake ice phenology are the Canadian Lake Ice Model - CLIMo (Brown & Duguay, 2011, 2012; Duguay et al., 2003; Menard et al., 2002) or High Resolution Thermodynamic Snow and Ice – HIGHTSI (Launiainen & Cheng, 1998; Semmler et al., 2012). In addition, lake models containing submodels for ice cover have also been used. Examples include: Lake Ice Model Numerical Operational Simulator - LIMNOS (Vavrus et al., 1996; Walsh et al., 1998), FLake (Kheyrollah Pour et al., 2012; Semmler et al., 2012), MyLake (Gebre et al., 2013; Saloranta & Andersen, 2007). Some of mentioned models are discussed and compared in Brown & Duguay (2010) and Stepanenko et al. (2010). Lake ice models are invaluable supplement to operational monitoring as they provide deeper understanding of interactions between lake and climate.

1.7 Research problem and objectives

Since the 1980 the number of lakes for which lake ice phenology is monitored *in-situ* decreased almost tenfold. Remote sensing offers ideal tool to fill created information gap, however, visual interpretation, implemented currently for operational use, is time consuming and can only be used for limited number of lakes. It is unfeasible for monitoring of small lakes and over large areas. Small lakes bellow 2 km^2 are, however, the representative lake size in the Arctic. Automatic method for extraction of lake ice phenology over large areas would enable mapping of spatial trends in lake ice phenology within the Arctic. Deeper knowledge of spatial distribution of break-up timing would contribute, for example, to more accurate estimation of methane emissions, especially the timing of the spring pulse release and its distribution in space. Monitoring of small lakes is especially important as they are likely to emit more greenhouse gasses per unit area than large lakes.

In consequence to limited data availability, majority of lake ice phenology trend analyses have been performed only for large or otherwise important lakes. Apart from limited number of regional projects, no assessment of trends for Arctic lakes is available. It has been observed that the effects of climatic change may be more severe in the Arctic than anywhere else, however, further study is still needed. Lake ice phenology has been recognized as robust indicator of climate change. Therefore analysis of changes and trends in timing of lake ice cycle might help to enhance our understanding of climate change at high latitudes.

Number of studies explored the relationship between the timing of phenological events to climatic and non-climatic variables. However, further investigation is needed to determine if the forcing variables differ from location to location.

1.7.1 Research Objectives

The aim of this study is to quantify the changes in lake ice phenology in the Arctic in relation to changing climate over the period from 2000 to 2013.

Specific objectives:

1. To develop an algorithm for automatic extraction of the dates of phenological events.
2. To map spatial distribution of timing of the phenological events in five study areas evenly distributed around the circumpolar Arctic.
3. To analyze lake ice phenology trends in the Arctic.
4. To estimate the role of various climatic variables on the timing of phenological events.

1.7.2 Research Questions

Objective 1:

- Is MODIS surface reflectance 250m product suitable for ice phenology monitoring at high latitudes?

Objective 2:

- What is the spatial distribution of timing of the phenological events?

Objective 3:

- Are there significant trends in the timing of phenological events for Arctic lakes?
- What is the spatial distribution of these trends, if they exist?

Objective 4:

- Which climatic factors determine the timing of phenological events?

2 Data

2.1 Study Areas

Five study areas have been selected for this study. All selected areas are located above 65° N and are approximately evenly distributed around the Arctic land mass. (Figure 2) The main selecting criteria were high concentration of lakes and relative climatic homogeneity within the area.

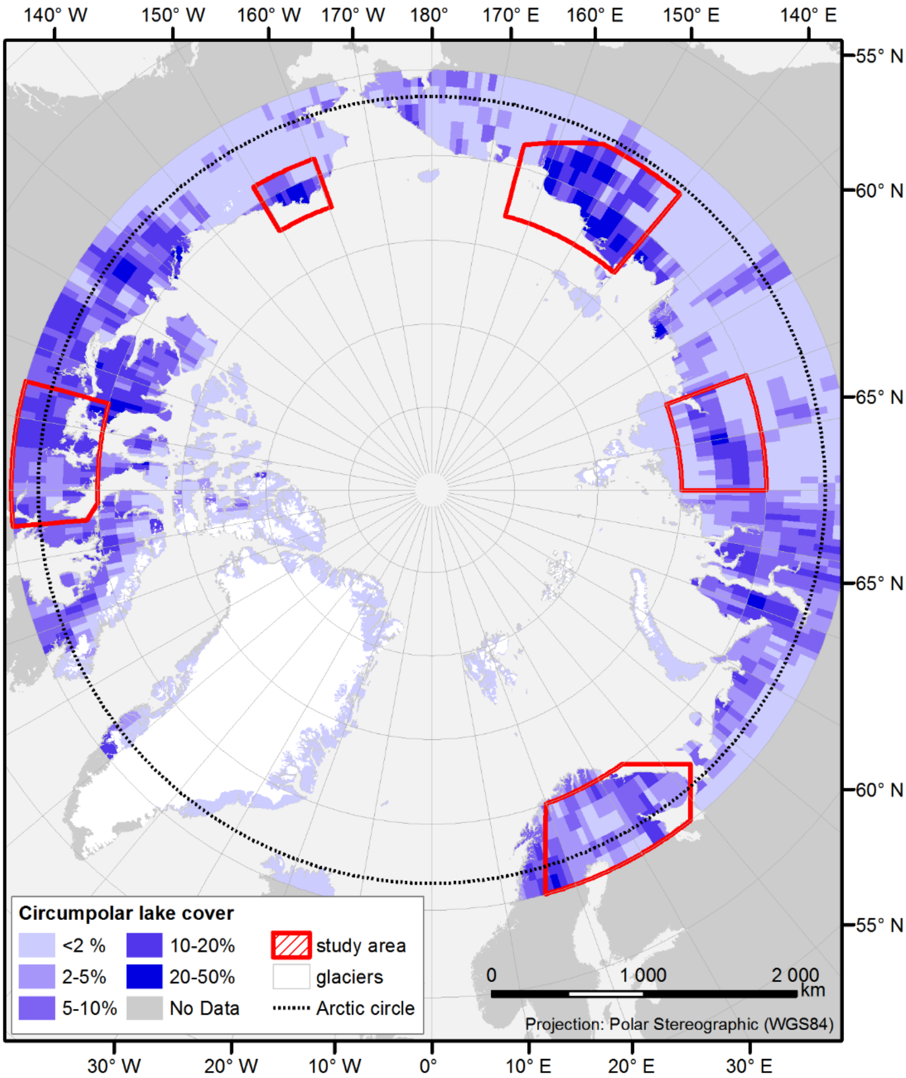

Figure 3: Study areas and circumpolar lake cover percentage (Paltan Lopez, 2013)

Only lakes larger than 1 km^2 are included in the study due to 250m resolution of the satellite data. 3223 such lakes, most of glacial origin, has been identified in the North European study area. The range of mean annual temperatures within the area is -5.1°C to 2.4°C (for base period 1950-1990) with mean temperature of -1.1°C. The average annual precipitation reaches 512mm (range 400 – 750 mm). (Figure 3 & 4) Compared to other four study areas Northern Europe is considerably more heterogeneous in many aspects (elevation, land cover, permafrost extent, etc.). It is also warmer and wetter as an effect of warm Gulf Stream running alongside the western coast. The western part of Scandinavia was excluded from the study due to much large increase in precipitation caused by the Scandes. Northern Europe is the only study area for which *in-situ* observations were available.

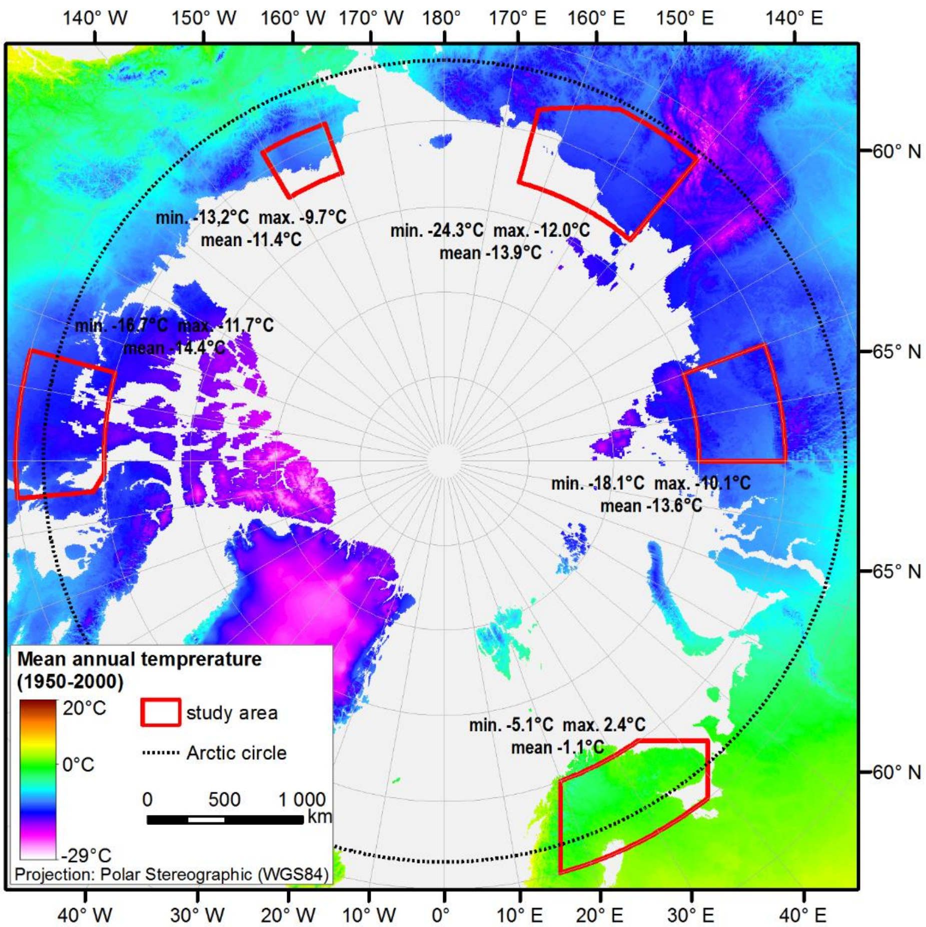

Figure 4: Map of mean annual temperature (1950-2000) (Hijmans et al., 2005)

The other four areas although different in many aspects, have similar characteristics regarding mean annual temperatures and rainfall. (Figure 3 & 4) Unlike North Europe where only isolated patches can be found in mountainous areas, continuous permafrost is present through their entire extent. (Figure 5) East Canada, Alaska and Taymyr study areas are open scrubland plains, with small elevation differences. Similarly as North Europe, Russian Yedoma study area, contains variety of land cover classes from bare rock to woodlands in low altitude locations.

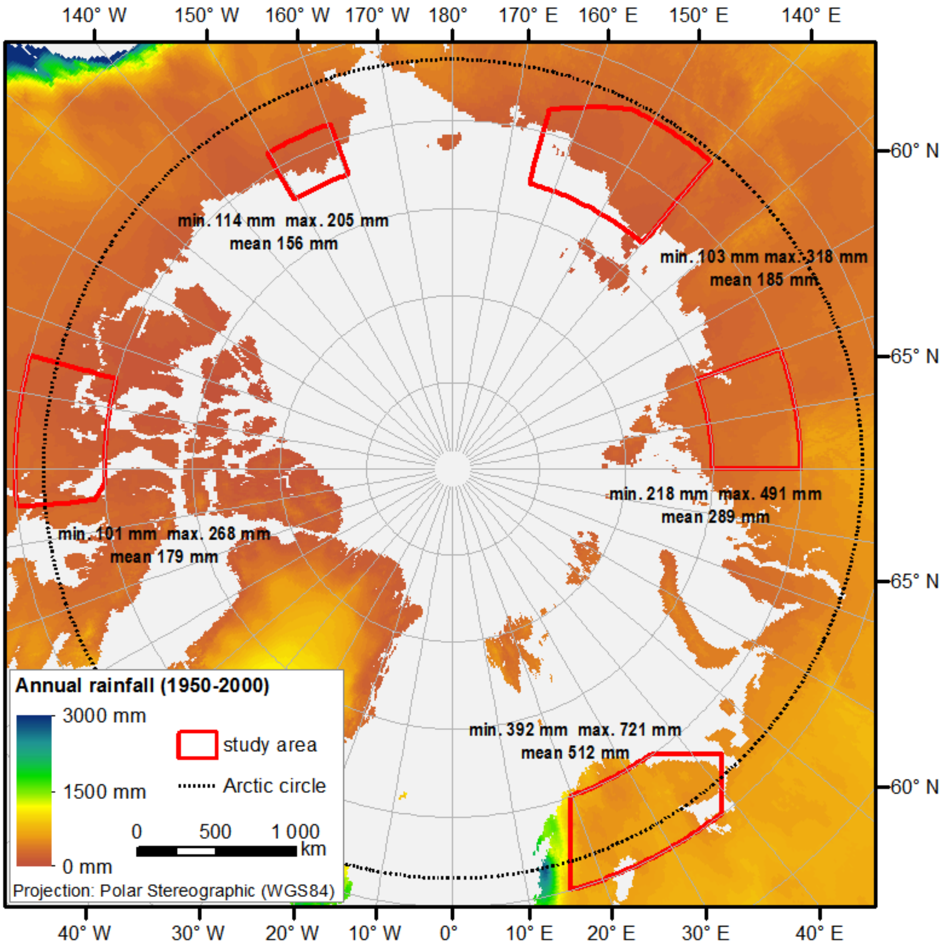

Figure 5: Map of average annual rainfall (1950-2000) (Hijmans et al., 2005)

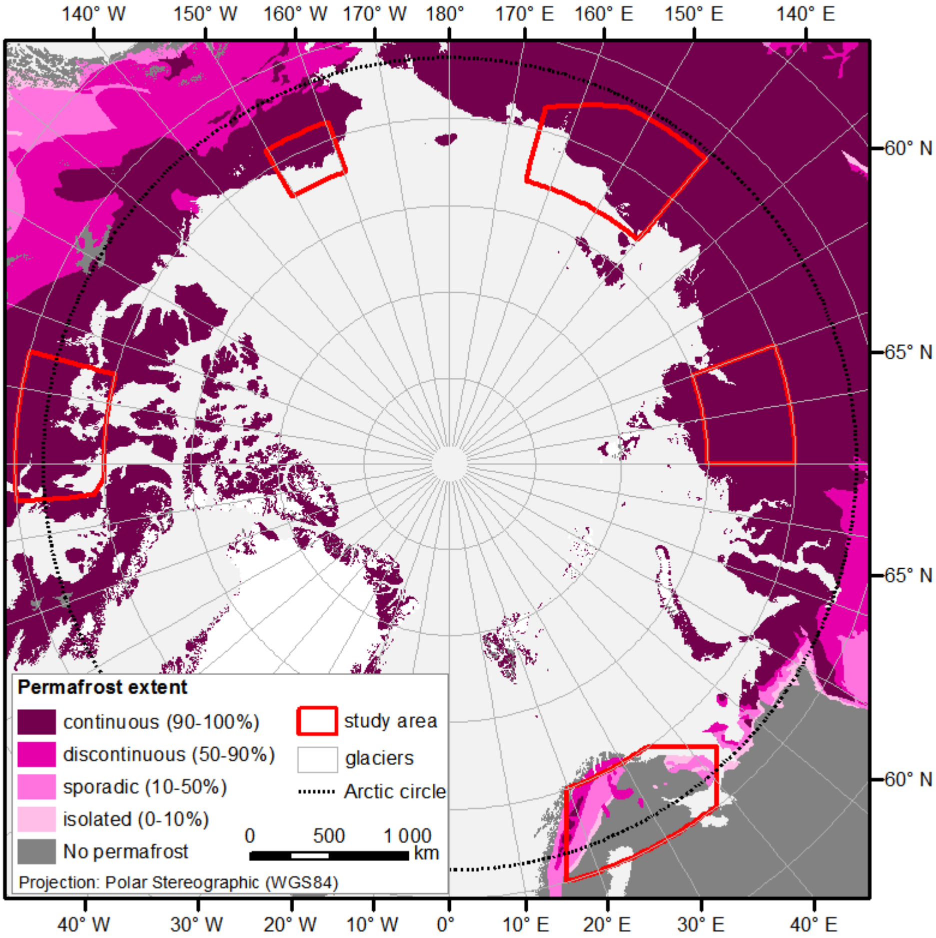

Figure 6: Map of circumpolar permafrost extent (Brown, et al., 1998)

2.2 Satellite data

The study is based on data acquired by the Moderate Resolution Imagining Spectroradiometer (MODIS) on board of the Terra satellite. Terra (EOS AM-1) satellite was launched in December of 1999 by NASA and has been collecting data since 24th of February, 2000 in to a sun-synchronous orbit with 98.5° inclination, at 705 km above Earth's surface. It is passing over equator north to south at approximately 10:30 a.m. with period of 99 min. Twin MODIS sensor has been launched on board Aqua satellite later in 2002 in similar orbit but with later overpass at 1:30 p.m.

MODIS acquires daily global coverage in 36 bands, 29 of which have 1km pixel resolution at nadir, 5 have 500m resolution and 2 bands in red and near-infrared parts of spectrum with resolution of 250m. 40 products are generated from the raw data, 18 combined form Terra and Aqua satellites and 22 separate. In this study daily MODIS Surface Reflectance Daily L2G Global 250m product (MOD09GQ) with generated from Terra MODIS data will be used. As level 2G-lite product MOD09GQ has been corrected for the effects of atmospheric scattering and absorption, composited on daily level, gridded into Modis sinusoidal projection and cut into approximately 10°x10° granules. (Vermote et al., 2011) Table 1 shows bands present in the original HDF file.

Science Data Sets	Units	Data Type	Fill Value	Valid Range	Scale Factor
Number of observations within a pixel	none	8-bit signed integer	-1	0-127	NA
250m Surface Reflectance Band 1 (620-670 nm)	Reflectance	16-bit signed integer	-28672	-100 - 16000	0.000 1
250m Surface Reflectance Band 2 (841-876 nm)	Reflectance	16-bit signed integer	-28672	-100 - 16000	0.000 1
250m Reflectance Band Quality	Bit Field	16-bit unsigned integer	2995	NA	NA
Observation Coverage	Percent	8-bit signed integer	-1	0 - 100	0.01

Table 1: Science Data Sets for MOD09GQ (Vermote et al., 2011)

The quality band of this product however, is not populated and the 1 km Reflectance Data State band from the MODIS Surface Reflectance Daily L2G Global 500 and 1km product (MOD09GA) must be used instead. The characteristics of each pixel are coded into 16 bit field and converted in to decimal number in range 0-65535. The state band includes information on cloud state, cloud shadow, water and ice/snow presence, aerosol quantity etc.

For purposes of this study only the near infrared band (band 2) of MOD09GQ product was used. Near infrared part of spectra is most suitable for lake ice detection due to very low reflectance of open water and therefore higher contrast between water and ice conditions. Moreover the NIR band is not as heavily affected by atmospheric conditions as the red wavelengths. (Latifovic & Pouliot, 2007)

Both, the MOD09GQ and MOD09GA products, were acquired for period from 2000 to 2013 from the NASA EOSDIS using the Reverb Echo online tool. As all optical sensor MODIS imagery is affected by low illumination in polar areas during fall and winter period. For this reason data from 31st November to February (date depending on study area) were not downloaded. However, the start and end of the low illumination period changes with latitude and additional correction is needed remove affected pixels.

2.3 Lake Datasets

Several datasets have been used to create underlying lake database for this study. Originally only two datasets were considered. The Global Lake and Wetland Database (GLWD) developed by Lehner & Döll (2004) is considered the most complete global database available. It is widely used for regional and global studies but majority of small (>10km^2) lakes mainly in the Arctic is omitted or their area is underestimated. The Arctic lakes that are mapped show slight shift in their location which makes the database unsuitable to use as a mask for satellite data extraction. Recently new more accurate database became available do the arctic areas created by Paltan (2013). The New Arctic Lake Geodatabase (NALGDB) is based on Landsat imagery and covers area north of 65°N. With overall accuracy of 78% and 30m resolution it is the most complete and accurate database of Arctic lakes available. While small lakes are mapped very wellmost of the large lakes are missing. This is most likely due to extremely long and raged boundaries large Arctic lakes and their misclassifivation as rivers when creating the database.

Due to drawbacks of both databases it was decided that their combination would yield the best results. As the more accurate, the NALGDB was used as a base, only lakes not included in NALGDB were then selected from the GLWD and both were merged into a final dataset for each study area. The result was checked visually over Google Earth.

For part of North Europe (Sweden and Finland) study area the Corine Landover (CLC2000) lakes class was used. The CLC2000 lake class shows accuracy > 95% and has 30m resolution. (EEA, 2006) It is, however, available only for European Union countries. For the remainder of Europe study area same approach as for other study areas was adopted. For all study areas the final lake dataset was projected to Polar Stereographic projection as defined by the WGS84 datum with central meridian at 0° E and latitude of origin at 71°N. (EPSG code 7215)

2.4 In-situ data – validation data

The availability of *in-situ* observations for lake ice phenology events is very limited in Arctic areas. Only two datasets were acquired, both within the North Europe study area.

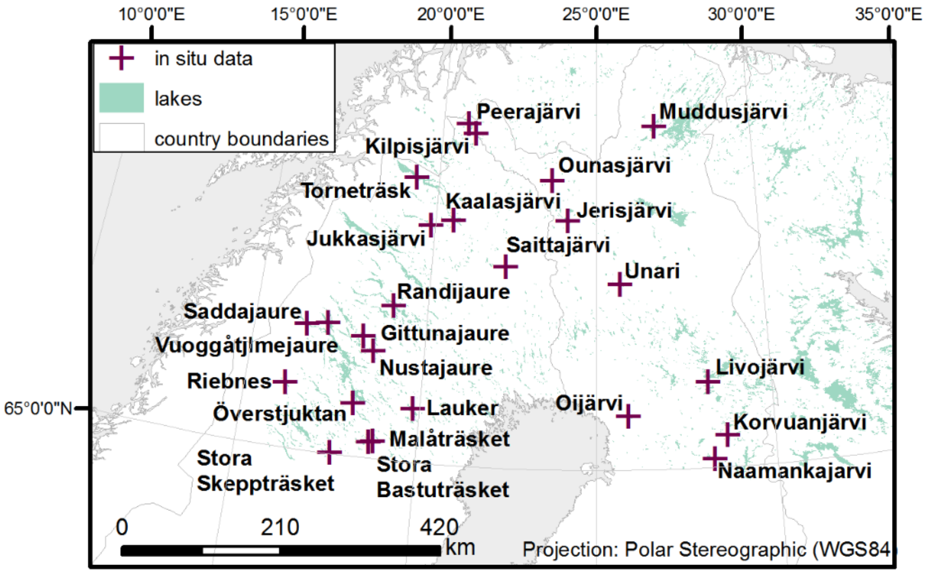

Figure 7: Map of lakes for which in-situ observations are available

Freeze-up end (FUE) and break-up end (BUE) time series were obtained from the Finish Environmental institute (SYKE) and the Swedish Hydrological and Meteorological Institute (SHMI) for 25 lakes. Some of the time series are incomplete. In total 288 observations for break-up and 275 for freeze-up.

2.5 Climatic data

Daily temperature, wind speed, precipitation and snow depth data were acquired form ERA Interim dataset. Gridded ERA Interim is global atmospheric reanalysis dataset produced by the European Centre of Medium-range Weather Forecast (ECMWF). It covers period from 1st January 1979 till present and continues to be produced in near-real time. The data were downloaded in NetCDF format for the entire Arctic. It has been selected due to its spatial resolution of 0.75 degrees, which one of the finest for global climatic datasets. Mean daily data are not available, therefore, it was calculated form 6-hourly data.

3 Methodology

With the exception of final phenology feature extraction, all processing was done using scripts developed in R or MATLAB environment. All used codes are included in Appendix.

3.1 Phenology feature extraction

The first objective of this study is to develop automatic method for extraction of lake ice phenology events from surface reflectance data. As mentioned earlier, visual interpretation is still the most widely used method for operational ice phenology extraction. Latifovic & Pouliot (2007) suggested the use of lake surface reflectance temporal profiles which greatly reduces the volume of data to be processed.

3.1.1 Surface reflectance profile preparation

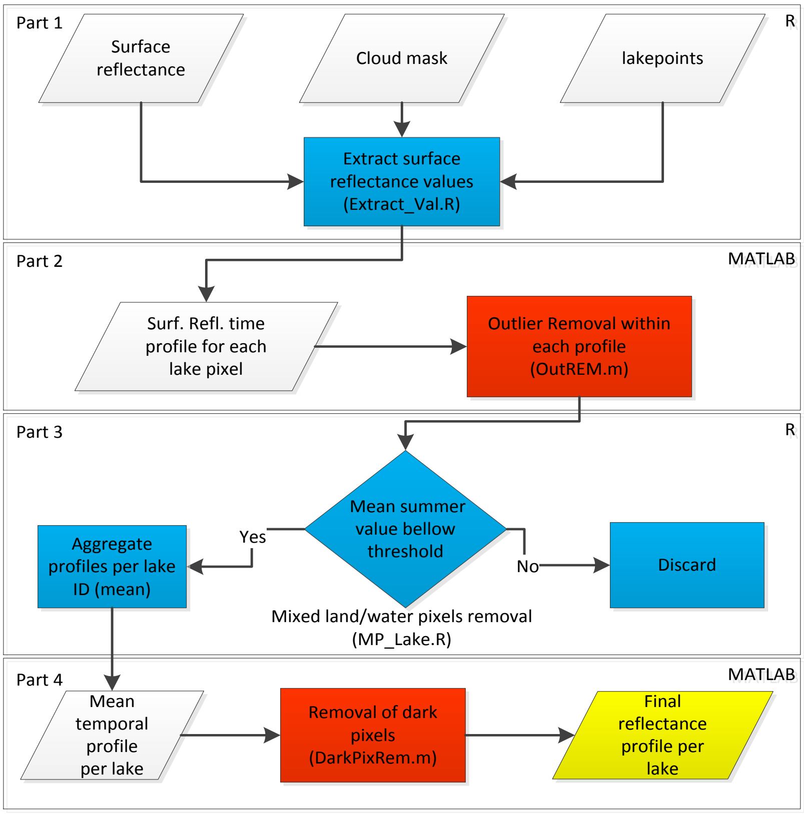

Figure 8: Flowchart for preparation of the final surface reflectance temporal profile. Each part is described below.

3.1.1.1 Satellite data pre-processing

Due to significant size of the original MOD09GQ and MOD09GA tiles the data had to be downloaded and processed continuously. This was done using the Aria2 download utility and MODIS Reprojection Tool (MRT) via script developed in R. (see Appendix A) The MRT was developed by the NASA EOS Land Processes Distributed Active Archive Center (LP DAAC) to "*support higher level MODIS Land products, which are distributed as Hierarchical Data Format-Earth Observing System (HDF-EOS)*[2] *files projected to a tile-based Sinusoidal grid*." (LP DAAC, 2011) For each of the five study areas the NIR and state band were extracted, mosaicked, reprojected and resampled to 250m. All data were converted into single band TIFF images of equal extent. In total, nearly 90 000 MOD09GQ and MOD09GA HDF tiles were processed over all study areas.

The state data were reclassified to binary cloud mask using code developed in Python based on ESRI ArcGIS functionality. (see Appendix B) The reclassification was based on look-up table in ASCII format containing all state values (0-65535) and corresponding class. Classes were assigned to decimal values based on corresponding 16bit binary number and State QA description available from Vermote et al. (2011). Only cloud free pixels with no cloud shadow and no cirrus cloud detected were classified as 1 all lower quality pixels were classified as 0.

3.1.1.2 Lake point dataset

First step was preparation of lake-point dataset for each study area containing coordinates, at which values from the satellite data will be extracted. Each lake in the final study area polygon lake dataset was assigned unique ID sorted by latitude of its centroid. It is crucial that the points correspond to the centroid of cells in the resampled imagery. The polygon dataset was then converted to a raster aligned with the pre-processed satellite data and with corresponding cell size (250m) and the unique lake ID was used as the value field. The raster was then converted to a point dataset where each point represents the centre of a raster cell. Apart from unique lake ID each point was assigned a unique point ID and set of coordinates in Polar Stereographic projection. The point dataset was the exported into an ASCII format.

3.1.1.3 Profile extraction (Part 1)

The surface reflectance values were extracted using code developed in R and using the R raster package. (see Appendix C) The extraction was done for one year at a time. The inputs were surface reflectance daily images, corresponding cloud mask data, both sorted by day of year, and the lake point dataset. The output is in a form of ASCII file, where each row represents daily reflectance values for one input point. As mentioned in the *Data* section, imagery from winter period from 1st December to 8th – 26th February (depending on location) was omitted from processing completely due to low illumination. (Figure 9) Moreover, in several years number of dates is not available due to MODIS technical issues. In such cases and in case of cloud cover No Data value (NaN) was assigned. All profiles show 75 – 88% of missing data. Although MODIS data are only available from 24th February 2000, the 1st January 2000 is considered as first day of study (DOS).

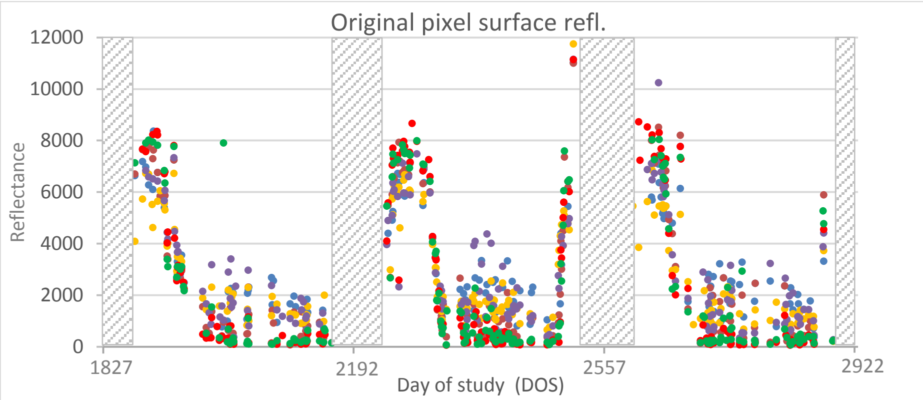

Figure 9: Example 3-year surface reflectance profiles for single lake covered by 8 pixels (output of Part 1). Grey rectangles show periods for which data were not processed due to low sun illumination affecting the reflectance values.

3.1.1.4 Outlier removal (Part 2)

Although low quality pixels were masked prior to value extraction, residual outliers still may be present within the profile. These values would affect the final mean lake profile and therefore a method to detect and remove them was devised in MATLAB. (see Appendix D) Leys et al. (2013) discussed the use of mean absolute deviation from median as the most robust dispersion measure in the presence of extreme values and therefore, ideal for their detection. Unlike frequently used standard deviation from mean, MAD is not so strongly affected by extreme values. Moreover, it is applicable even to very small sample sizes. Median absolute deviation (MAD) was calculated within each window. MAD is be defined as follows:

$$MAD = b * M_i(|x_i - M_j(x_j)|) \tag{5}$$

where b is a constant defining the data distribution, here b= 1.4826 for normal distribution was used. $M_j(x_j)$ is the median of *n* original values and M_i is the median of absolute deviations from the sample median. The MAD was calculated in 11-day moving window. A conservative threshold of 3*MAD form window median was selected to define outliers. The central value of the sample exceeds this threshold it is assigned NoData value. However, only under the condition that minimum of 5 non-missing values are present in the window.

3.1.1.5 Removal of mixed pixel profiles (Part 3)

Large number of pixels for which the reflectance profile was extracted correspond co mixed land/water cover. The reflectance of land is much higher than that of open water in summer and, depending on snow cower, it can be lower than ice reflectance in winter. Therefore, mixed pixels would introduce errors into the final mean lake profile and affect the accuracy of feature extraction. The effect of mixed reflectance can be seen in Figure 9 (violet and yellow time series).

The first step was to define an open water reflectance threshold. 1000 reference open water points were selected from the dataset over high resolution imagery (Google Earth). For each of these points mean reflectance for two summer months July and August was calculated over 5 years. July and August were selected as months, for which the lakes are most likely to be ice free. In case some of residual ice on some lakes, the 3rd quartile (390.4) of all the means was then set as the open water reflectance threshold.

The next step was to identify and remove the mixed pixel profiles. For all the points in the original dataset summer means for 5 years were calculated, similarly as for the reference points before. All means were then compared to the threshold. If the summer means were higher than the threshold for more than 2 out of 5 years, the profile was discarded. Due to raged coastlines of Arctic lakes almost 75% of profiles were removed. For approximately 40% of lakes, depending on the study area, all the profiles were removed eliminating them completely from the analysis.

In final step all pixel profiles per lake were averaged to produce mean surface reflectance profile per lake. (Figure 10) The whole process was done using algorithm written in R. (see Appendix E)

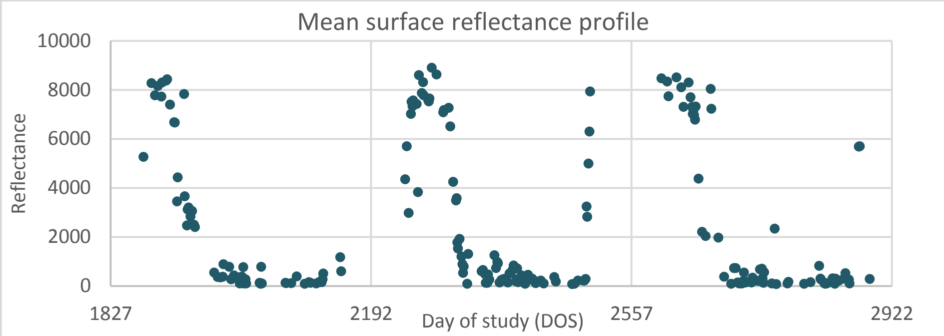

Figure 10: Example 3-year mean lake surface reflectance profile

3.1.1.6 Removal of 'dark pixels' (Part 4)

In fall, the effect of low sun illumination on the surface reflectance starts to be visible from mid-October, depending on the study area. (16th October for Taymyr area, 29th October for Alaska and 3rd of November for Northern Europe and East Canada). The affected portion of image gradually increases till the end of November and decreases again from to clear image on 8th – 26th February depending on study area. The reflectance values of affected 'dark' pixels are close to 0, however, the state data do not classify these pixels as low quality. For spring, all affected images can be removed from analysis because the start of break-up period occurs considerably later. On the other hand, the freeze-up period in the Arctic occurs when the sun radiation decreases during October and November. Therefore, it is crucial for estimation of FUS and FUE to preserve as many valid observations for November as possible.

In this period majority of profiles show rise in reflectance values suggesting start of ice formation on the lake, and then sudden drop to near 0. To remove the 'dark pixel' values a function was built in MATLAB. (Appendix F) To identify affected values, the maximum value in 7 days preceding the first affected image is compared to all following values. All lower values are then removed. If all values in the 7 day window are missing, the open water reflectance 390 calculated previously is set as the lower limit.

Up to this point all processing was done on yearly basis. The last step in the reflectance profile preparation is combining all the years into one 14 year time series. For easier manipulation, MATLAB *.mat* file was created containing the time series (variable Y), list of days of study (1-5114, variable X) and variable Lakes_ID containing the list of ID for all lakes in Y. This file is used as input for further processing.

3.1.2 Phenology events extraction

The extraction of lake ice phenology was done using the TIMESAT software developed at Lund University for analysis of time series data. In this study the MATLAB based version of the software was used.

3.1.2.1 TIMESAT input file preparation

TIMESAT requires input data in predefined format – An ASCII file where each row corresponds to one time series, the header contains number of years in time series (*nyear)*, number of observations per year (*nptyear*) and number of time series (rows) in file (*nts*). For all years the number of observations must be equal and the total number of values for series must be equal to *nyear*nptyear.* The preparation was done using MATLAB. (Appendix G)

14 year time series have proven too long for TIMESAT to process. Therefore, the data have been divided into two 7-year parts. The software extract seasonality for n-1 seasons from the centre, therefore the data for first and last season would not be extracted. To overcome this issue, the time series were padded with one dummy year on each side containing mirrored data for first/last year. Finally the amount of missing data within the time series (75-88%) posed an issue, therefore, linear interpolation was used to fill the gaps (Figure 11) and all remaining missing data must be assigned numerical value (-999).

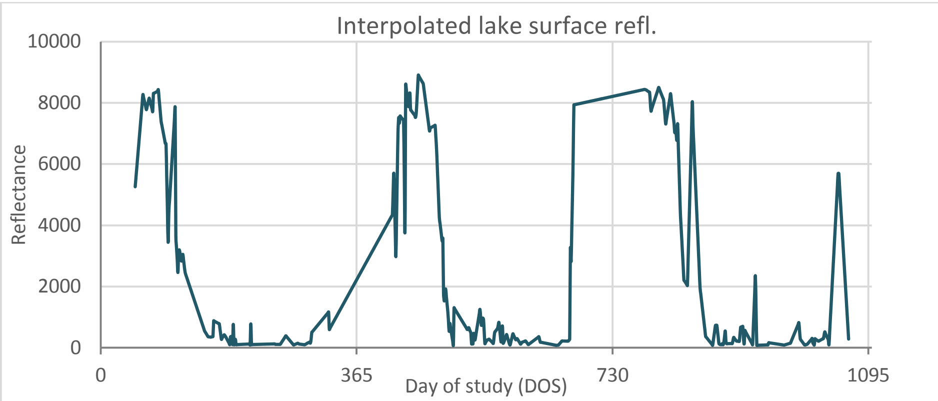

Figure 11: Example 3-year interpolated profile

3.1.2.2 TIMESAT

TIMESAT implements three methods based on least-squares fit to upper/lower envelope: adaptive Stavitzky-Golay filter, fit to asymmetric Gaussian and fit to double logistic functions. For purposes of this study the third method, consisting of double logistic functions fitted to data in intervals around local minima and maxima, was used. (Jönsson & Eklundh, 2004)

Many time series contain residual extreme values that seriously affect the function fits. TIMESAT provide three methods for spike and outlier removal. The first method is based on moving window median. In the second method weights are assigned to values based on STL-decomposition. The third method allows to modify the weights based on ancillary data. Although outliers have been removed from the time series during pre-processing it still contains large amount of spikes. Based on visual comparison the best results are achieved using the second method.

TIMESAT was developed to process predominantly NDVI time series and therefore its functionality allows adaptation to upper envelope on scale 1 to 10. In first step the fit parameters are calculated by solving set of equations with weights obtained from STL-decomposition. In second step weights for values below the model function are deceased by certain factor (adaptation strength). (Eklundh & Jönsson, 2012) Therefore, although adaptation strength bellow 1 is not in the intended input range, it can be used to force adaptation to lower envelope. The adaptation process can be repeated 2 times. The adaptation strength is set specific to phenological feature and will be discussed below.

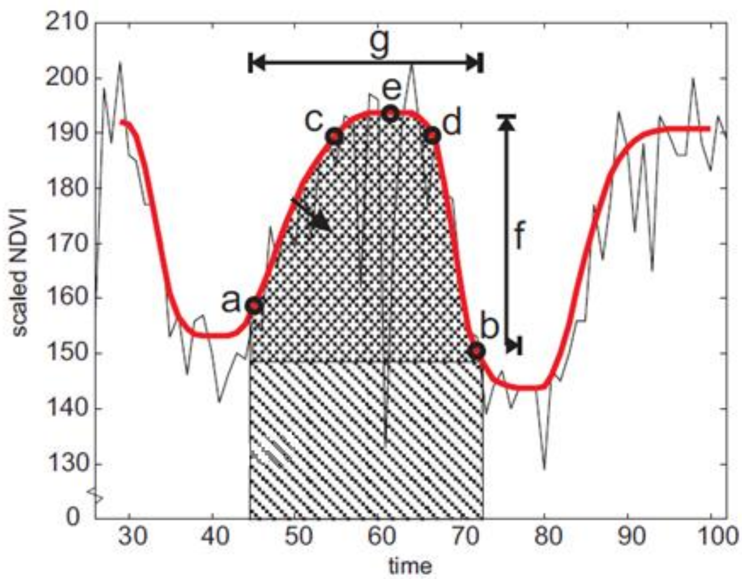

Figure 12: Example extracted seasonality - points (a) and (b) mark start and end of the season. Points (c) and (d) give the 80 % levels. (e) displays the point with the largest value. (f) shows the seasonal amplitude and (g) the length. (Eklundh & Jönsson, 2012)

Other settings include number of seasons per year (*season par.* = 1), *data range* (0-10000), and *Force to minimum*, this setting forces set value to minimum (not used). The seasonality extraction is based either on absolute value or on % of seasonal amplitude measured from minimum level. (Eklundh & Jönsson, 2012) Extracted seasonal parameters include: time of start and end of the season, amplitude, length, base value (minimum), maximum value, etc. (Figure 12) However, for the purposes of this study, only the start and end of season are of importance.

The setting files are prepared in TIMESAT *TSM_GUI* utility which allows visual comparison of various setting on the fitted function and extracted seasonality. Although the MATLAB based version of TIMESAT was used the utility to process data (*TSF_process*) is FORTRAN based. MATLAB based version is also available (*TSM_process*), however due to considerably lower speed it is only suitable for small datasets.

3.1.2.3 FUS and BUE dates extraction

Features in lower portion of the reflectance profile, such as start of freeze-up (FUS) and end of break-up (BUE) are more strongly affected by erogenous bright values (cloud, haze, etc.). (Latifovic & Pouliot, 2007) To overcome this issue they were from fit adapted to lower envelope. The fit to lower envelope was achieved by setting adaptation strength to 0.1 (increasing weights of values bellow the initial function fit).

The position of both events was based on visual inspection of profiles. For BUE three reference lakes have been selected from the *in-situ* data (18 observations). It was observed that the date of BUE corresponds approximately to 12% amplitude. The freeze-up start (start of season) was defined at 5% of amplitude. (Figure 13) The complete settings file is included in Appendix G.

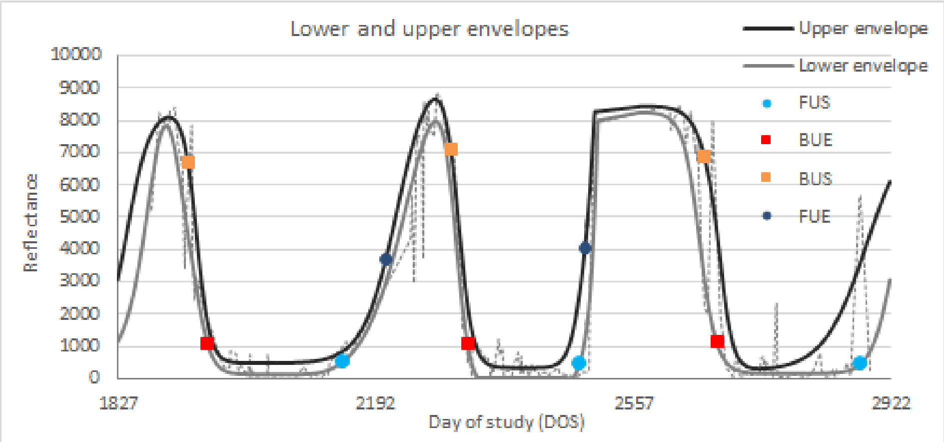

Figure 13: Example fit to lower and upper envelopes with position of extracted phenological events

3.1.2.4 FUE and BUS dates extraction

Features in the upper part of reflectance spectrum, freeze-up end (FUE) and break-up start (BUS), are more susceptible to dark artifacts (e.g. shadow) Therefore, fit to upper envelope was used to extract these events (adaptation strength set to maximum - 10).

Due to low sun illumination, FUE is not a distinct feature in most profiles. The best position (40% amplitude) was, similarly as for BUE, judged by visual comparison with *in-situ* data for three reference lakes. For BUS the best position was estimated at 80 % of amplitude. (Figure 13) The complete settings file is included in Appendix G.

3.1.2.5 Seasonality processing

Extracted seasonality for all events is exported into an ASCII file using the TIMESAT *TSM_printseasons* tool. Original data have been divided into two 7 year blocks and padded with dummy year at both sides. The output files must be therefore combined and the extracted dates must be corrected for the added first year. The freeze-up and break-up dates must be processed separately due to shift in season number (break dates for seasons 1999/2000 to 2012/2013, freeze dates for 2000/2001 to 2013/2014). The output file differentiates lakes only using row number of input file, therefore Lake ID must be assigned to all records again. For some lakes and seasons the parameters were not extracted, most likely due to error in original data or low amplitude value. Lakes for which less than 13 events (seasons) years were extracted have been removed from further analysis. Next step was to convert dates given in day of study (1-5114) to day of year. The final seasonality for each phenological event was saved into separate CSV file. The processing was done in MATLAB and the codes for both freeze and break-up can be found in Appendix I.

Final step was to create one final seasonality dataset containing all phenological events and lake characteristics, such as ID, latitude, longitude and area in km^2. This was again done in MATLAB (Appendix J). In this step values that would present problems in later analysis, such as freeze-up dates occurring before freeze-up start, were removed. Also dates of freeze-up occurring later than day 400 were removed, such values do not correspond to reality in Arctic areas and are caused by errors in extraction due to missing surface reflectance data (large gap during the freeze-up period that cannot be accurately reconstructed with linear interpolation). Output combined seasonality file serves as input for further processing.

3.2 Comparison with in-situ observations

Estimated dates for freeze-up end and break-up end were compared to available *in-situ* data. Agreement between *in-situ* observations and estimates was measured using simple correlation, coefficient of determination R^2, mean absolute error (MAE) and root means sugared error (RMSE).

3.3 Trend analysis

The third objective of this study was to analyse trends in lake ice phenology in the Arctic. To achieve this goal trends for each lake were analysed over the studied period 2000-2014. To analyse the time series for trend the Theil-Sen slope approach was used. (Sen, 1968) The rank based, nonparametric Mann Kendall test was used to assess the significance of trend. To correct for serial correlation the approach proposed by Zhang et al. (2000) was used. The trend analysis was performed using the 'zyp' (Zhang + Yue-Pilon trends) package in R environment developed by Bronaugh & Werner (2013). Trends statistically significant at minimum of 95% confidence level were visualized using ESRI ArcMap software.

3.4 Regression

The last objective of the study was to relate extracted phenological dates to climatic and non-climatic variables. Based on literature review 6 variables were included in the analysis. These were temperature, wind speed, snow depth, precipitation, latitude and surface area of the lake. For climatic variables several variations were tested as described in Table 2 below.

Group	Variable	Definition
Size	SL	Size combined with latitude – size alone cannot be used (timing is different for lakes of same size at different latitude)
Latitude	Lat.	
Amplitude	Amp.	Annual temperature amplitude (temp. profile smoothed in 31 day window)
0° isotherm	0is	0° C isotherm timing (temp. profile smoothed in 31 day window)
Temperature	T7	Mean temperature over 7 days before event
	T15	Mean temperature over 15 days before event
	T30	Mean temperature over 30 days before event
	T45	Mean temperature over 45 days before event
	Ti7	Integrated negative/positive temperature for 7 days before the event
	Ti15	Integrated negative/positive temperature for 15 days before the event

	Ti30	Integrated negative/positive temperature for 30 days before the event
	Ti45	Integrated negative/positive temperature for 45 days before the event
Wind Speed	W7	Mean wind speed over 7 days before the event
	W15	Mean wind speed over 15 days before the event
	Wm7	Maximum wind speed in 7 days before the event
	Wm15	Maximum wind speed in 15 days before the event
Snow Depth	S7	Snow depth 7 days before event
	S15	Snow depth 15 days before event
	S30	Snow depth 30 days before event
	S45	Snow depth 45 days before event
Precip itation	TP7	Mean precipitation over 7 days before the event
	TP15	Total precipitation over 15 days before the event
	TP30	Total precipitation over 30 days before the event

Table 2: Climatic and non-climatic variables for which effect on phenology event timing was tested.

The available daily climatic data were downloaded in NetCDF format for the entire Arctic. First data for one day were exported to Esri ArcMap as a raster layer. The raster was converted into point file where each point corresponded to the centroid if one cell and unique id was assigned to each point (*Rasterpoint_ID.csv*). Then the point dataset was converted back to raster of unique IDs and saved as TIFF image (*Raster_ID.tif*). The *Rasterpoint_ID.csv* file was the used to extract time series of daily climatic data for each cell into an ASCII file in R (same approach as when time series for surface reflectance were extracted) For temperature no mean daily data were available so 4 datasets of 6-hourly data were downloaded. Once time series for all 4 were extracted a daily mean time series was calculated. Similarly for wind speed only separate north to south component and east to west component were available and total winds had to be calculated. The raster values were extracted at lake centroids using R.

The 0.75 degrees resolution of climatic data, corresponds approximately to 111 km in in south to north direction and 47-29 km depending on latitude in east to west direction. Therefore many lakes with different timing of phenological events fall within one cell, this would strongly affect the results of the regression analysis. For this reason, it was necessary to first upscale the phenology to same spatial resolution. Size is one of selected explanatory variables, therefore it should be preserved during the aggregation. First lakes were divide into 16 size categories (<5 km^2, 5-15 km^2 … >200 km^2). Lake ice phenology was then aggregated based on size category and cell ID with median as the output value. All discussed steps were performed in R.

All climatic variables described in Table 2 were calculated and added to the data based on cell ID using a routine written in MATLAB (Appendix K). For annual temperature amplitude and 0 degree isotherm the daily temperatures were processed using TIMESAT Stavitzky-Golay filter with window size of 31 days and start and end of season set to absolute value 0. The output seasonality file then contains annual amplitude and the timing of fall 0 isotherm as end of season and timing of spring 0 isotherm as start of season.

In total 23 variables served as input to multiple stepwise linear regression based on Akaike information criterion (AIC) performed in R. To avoid autocorrelation between climatic variables, such as between T30 and T45 only one variable from each group was selected. However, temperature amplitude and 0 isotherm are not included in separate groups since they show no significant correlation with other variables or each other, and they hold different information. In first step simple regression was calculated for each variable and the model with lowest AIC was selected. All variables in the same group were then excluded from testing in the next step and so on. In total five steps were performed. Explanatory power of each variable vas explored using IBM SPSS Statistics software. In final model only variables explaining more than 1% of variability were retained.

4 Results and Discussion

4.1 Extracted phenological events

The first objectives of this research were to develop an algorithm for automatic extraction of the dates of phenological events and to map spatial distribution of Arctic lake ice phenology for five sample study areas. The results are presented and discussed in following section.

4.1.1 Automated phenology extraction

The entire process from MODIS imagery download to the phenology extraction has been successfully automated, although, it still requires significant amount of user control. The use of lake surface reflectance temporal profiles greatly reduces the data volume from thousands of images to single matrix containing all necessary information.

MODIS near infrared surface reflectance pixel profiles were extracted for over 18 000 lakes over the five study areas and extend over 13 years (5114 days) from 1st January of 2000 to 31st December 2013 corresponding to 13 full and 2 half ice seasons. However, more than large portion of the lakes were eliminated during the removal of mixed land/water pixel profiles. Lakes in postglacial landscapes with resistant bedrock, such as the Precambrian shields of Northern Europe and North East Canada, tend to have long and ragged boundaries. (Vincent & Laybourn-Parry, 2008) For some lakes this results in all overlying pixels classified as mixed. For Northern Europe and East Canada study areas 1809 out of 3223 lakes (56%) and 3021 out 5965 lakes (50%) of the original lake dataset, respectively, remained for further analysis. In study regions of Russian Yedoma, North Alaska Taymyr, lakes have slightly more regular shapes determined mostly by thermokarsting processes. For Alaska and Yedoma study areas this resulted in smaller portion of lakes being removed. 1308 lakes out of 1740 (75%) for Alaska and 6039 lakes out of 9759 (62%) for Yedoma were kept after the mixed pixel profiles removal. The highest portion of lakes have been eliminated for the Taymyr area (only 1237 out of 2547 (48%) lakes kept). The ragged boundary is not the only reason for all pixel profiles for a lake to be classified as mixed. Only lakes with area larger than 1 km^2 were included in the study. This corresponds to 4 pixels in the surface reflectance data. For lakes with surface area close to 1 km^2 (more than 50%) this can result in no clear water pixel available over the lake. Although many of Arctic lakes are very shallow (1-2m) the lake bottom reflectance has negligible impact on reflectance values in near infrared part of spectra. (Tolk et al., 2000)

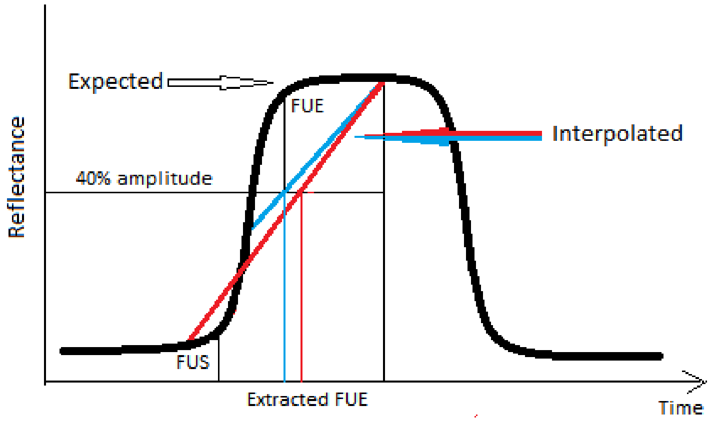

Figure 14: Simplified seasonal lake reflectance profile with expected (true) shape and interpolated shape if FUS occurs before the last available reflectance value (blue line) or after (red line.)

Due to low winter illumination and extensive cloud cover all profiles exhibit 75-88% of missing data, with especially long gap from October to February depending on latitude. Many Arctic lakes completely freeze after the sun illumination decreases to minimum. Limited data availability during the freeze-up period complicates the extraction of the freeze up start and end timing. Linear interpolation used to fill the gaps does not provide good results, as the true surface reflectance in this period does not follow a linear relationship with time. (Figure 14) If the ice cover starts to form before the last available surface reflectance value at least part of the true shape is retained in the data. However, if FUS occurs after, the possibilities of accurate extraction become very limited. In total ice phenology has been extracted for approximately 13 414 lakes.

4.1.2 Comparison to *in-situ* data

To assess the accuracy of extracted phenology the results were compared with *in-situ.* Ground observation data for period 2000 to 2013 were available only for the timing of freeze-up end (FUE) and break-up end (BUE) for selected lakes in North European study area. In total *in-situ* data for 25 lakes were available resulting in 288 *in-situ* observations for BUE and 274 for FUE. *In-situ* data for other study areas and for freeze-up start (FUS) and break-up start (BUS) were not available.

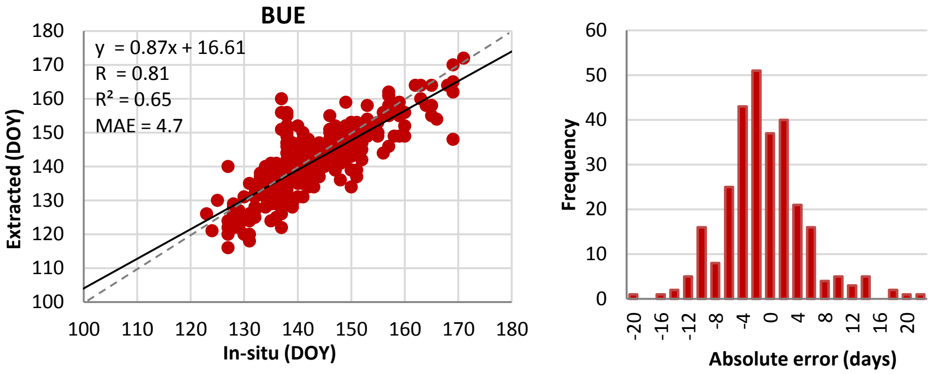

Figure 15: Comparison end of break-up dates observed in-situ and extracted (right) Absolute error distribution (left). Black line corresponds to linear trend-line and dashed line represents 1:1 relationship.

The extracted dates for BUE showed good agreement with the corresponding *in-situ* observations with mean absolute error (MAE) of 4.7 days and root mean squared error (RMSE) of 6.16 days. BUE occurs in period from end of April to mid-June when sun is high in the sky and incoming radiation if high. The BUE is defined as first day when water is completely free of ice, however, it is not uncommon for lakes to repeatedly break up and refreeze in the course of spring which then makes accurate date extraction difficult.

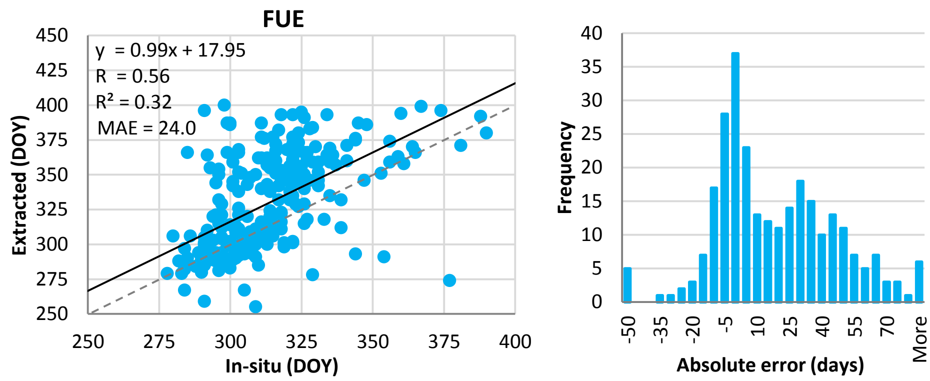

Figure 16: Comparison end of freeze-up dates observed in-situ and extracted (right) Absolute error distribution (left). Black line corresponds to linear trend-line and dashed line represents 1:1 relationship.

The results for FUE are not very encouraging with coefficient of determination of 0.32. The mean absolute error equals 24.0 days and RMSE is 32.8 days. (Figure 16) This can be mostly explained by the fact that FUE generally corresponds with low sun illumination and therefore surface reflectance values for the period of freeze-up are not available, as was discussed previously. The extracted values are consistently overestimated by approximately 17 days. This is most likely combined effect of interpolated data and extraction method. (Figure 14) As discussed in previous section the accuracy of used extraction method (40% amplitude) is determined by the value of last available surface reflectance observation. To assess, if errors are lake specific, the extracted variability within time series was analysed. The median standard deviation within the FUE time series is 31 days, as opposed to 7.5 days for BUE. The high standard deviation of FUE values suggests that there is no lake specific bias in the extraction method.

Not all uncertainty can be attributes to the extracted values, small part can be most likely attributed to the *in-situ* observations. Arctic lake ice often breaks and refreezes several times during course of fall and spring and multiple dates are, therefore, recorded for FUE and BUE. The last dates are the considered as the true timing of complete freeze over and break-up. However it has been pointed out by the SYKE that occasionally the dataset contains other than the last date. The accuracy, however was not given. *In-situ* observations are generally point based and often it is not possible to see entire lake from a single viewpoint.

For extracted dates of freeze-up start and break-up start the variability within time series was also assessed. For BUS the median standard deviation is 7.6 days (in agreement with BUE). For FUS (19.0 days) it is lower than for FUE, however, still higher than expected.

Due to low confidence given to extracted dates for both freeze-up events, further analysis will be performed only for start and end of break-up period.

4.1.3 Spatial distribution of BUS and BUE dates

In this section the spatial trends within study areas will be discussed. In all study areas the spatial distribution of break-up start and break-up end timing corresponds well to that simulated by Walsh et al. (1998) for northern Hemisphere.

4.1.3.1 Northern Europe

Figure 17 (p.66) and Figure 1818 (p.67) show spatial distribution of timing of break-up start and break-up end, respectively, for the North European study area. Of all the study areas Northern Europe shows the highest variability in dates for both BUS and BUE. The timing of break-up start ranges from end of March for lakes around the Gulf of Bothnia and western coast of White Sea to end of May for high altitude lakes in the Scandes and the plains of north east Kola Peninsula. The timing of break up end ranges from late May to end of June and exceptionally to early July for small high latitude lakes on the western slope of Scandes. The duration of break-up period varies from three to six weeks with median of 27 days. (On following maps same colour for BUS and BUE represents break-up period of 25 days)

The spatial distribution is a result of multiple factors. Apart from the expected latitudinal trend, strong determination by altitude can be observed. For example, lakes in the depression of Tuloma River, break up much sooner than surrounding higher located lakes. When inspected individually, if the lakes show considerably later break-up, it is due to permanently shadowed location on northern slope of a mountain. Several river segments have been misclassified as lakes and appear as points of very early break-up compared to surrounding lakes. Examples of misclassified river segments are visible on eastern tip of Kola Penisula and on dammed rivers in Scandes and in delta of Tuloma River.

The results of trend analysis for individual will be reported in next section, however the study area as a whole does not show any significant trend.

4.1.3.2 East Canada

The spatial distribution of break-up start and break-up end timing for East Canada study area is shown in Figure 19 (p. 68) and Figure 20 (p. 69), respectively. The timing of break-up start ranges from mid-May (warm years) south-west part to late June for north islands (e.g. King William Island, Victoria Island). The timing of break up end ranges from early June to late July. The duration of break-up period shows the same range as for previously discussed Northern Europe study area with slightly higher median of 28 days.

The spatial distribution follows south-west to north-east gradient following regional temperature patterns. There are no significant differences in altitude, land cover, precipitation or precipitation and temperature appears to be the sole determinant of spatial distribution.

The results of trend analysis for individual will be reported in next section, however the study area as a whole does not show any significant trend. Number of seemingly randomly scattered lakes in east part of the study area shows earlier break-up than surrounding lakes. Upon closer inspection over high resolution imagery, these lakes have light green-brown colour. Without further information it is impossible confidently determine the reason.

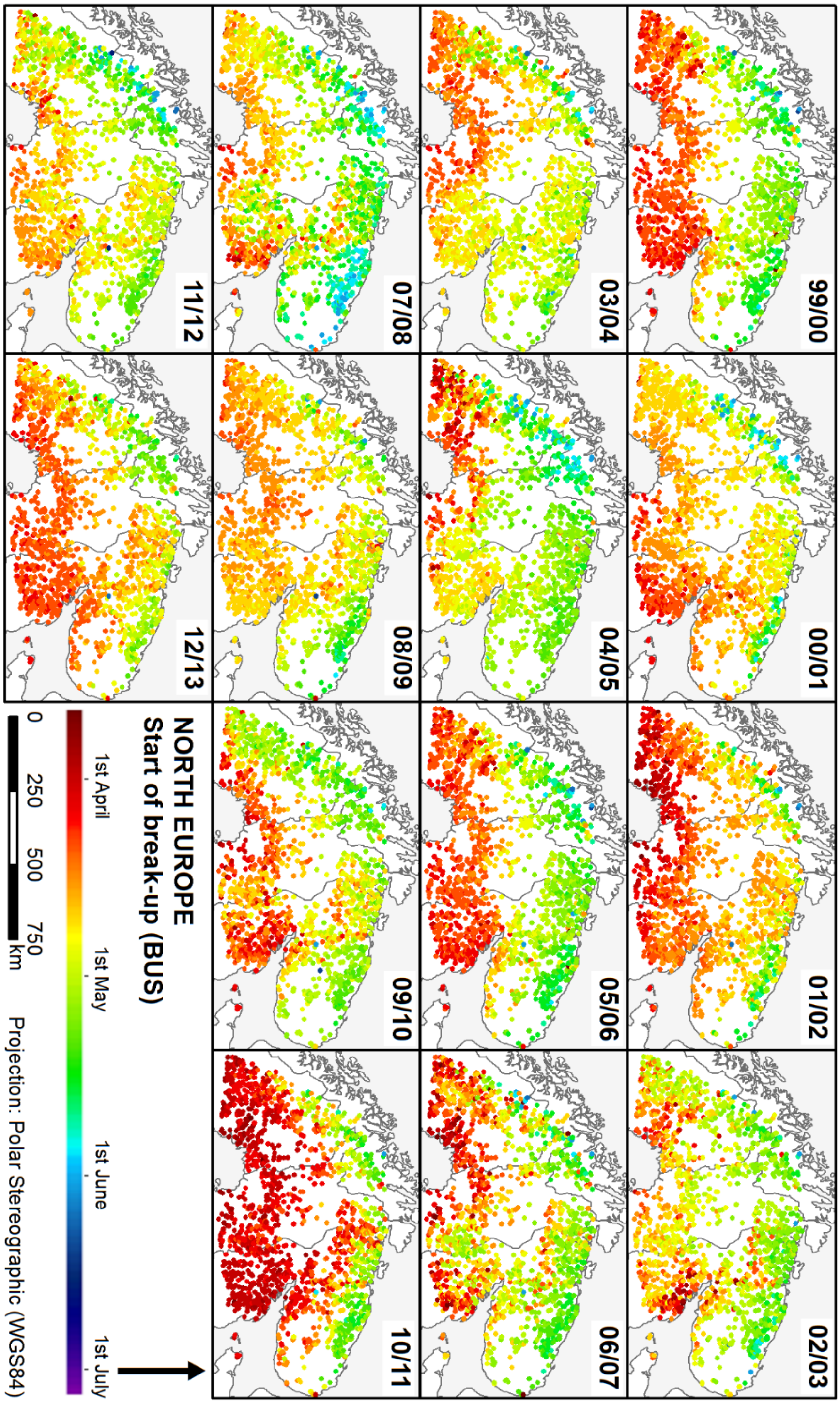

Figure 17: Timing of break-up start (BUS) for North Europe study area for seasons 1999/2000 to 2012/2013

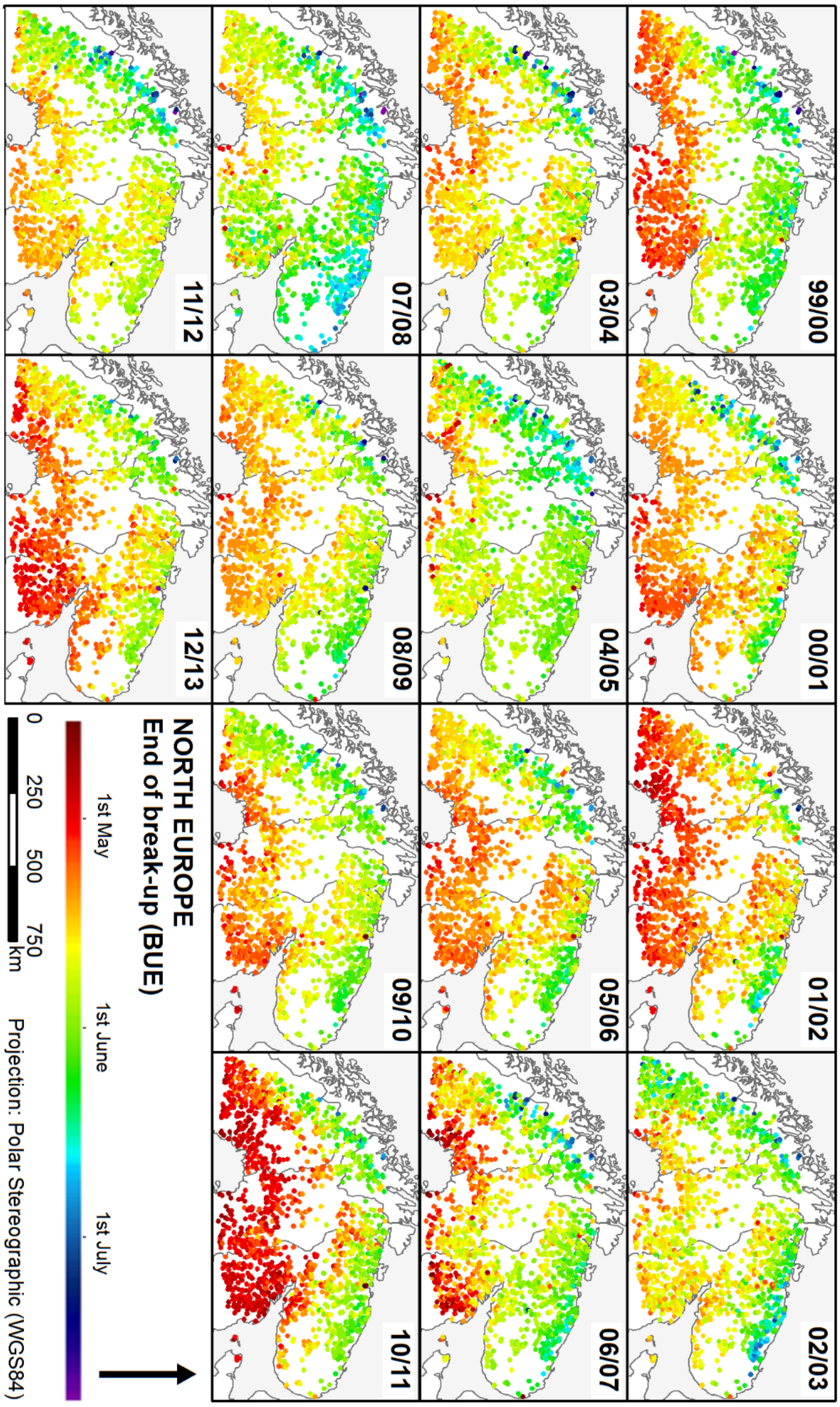

Figure 18: Timing of break-up end (BUE) for North Europe study area for seasons 1999/2000 to 2012/2013

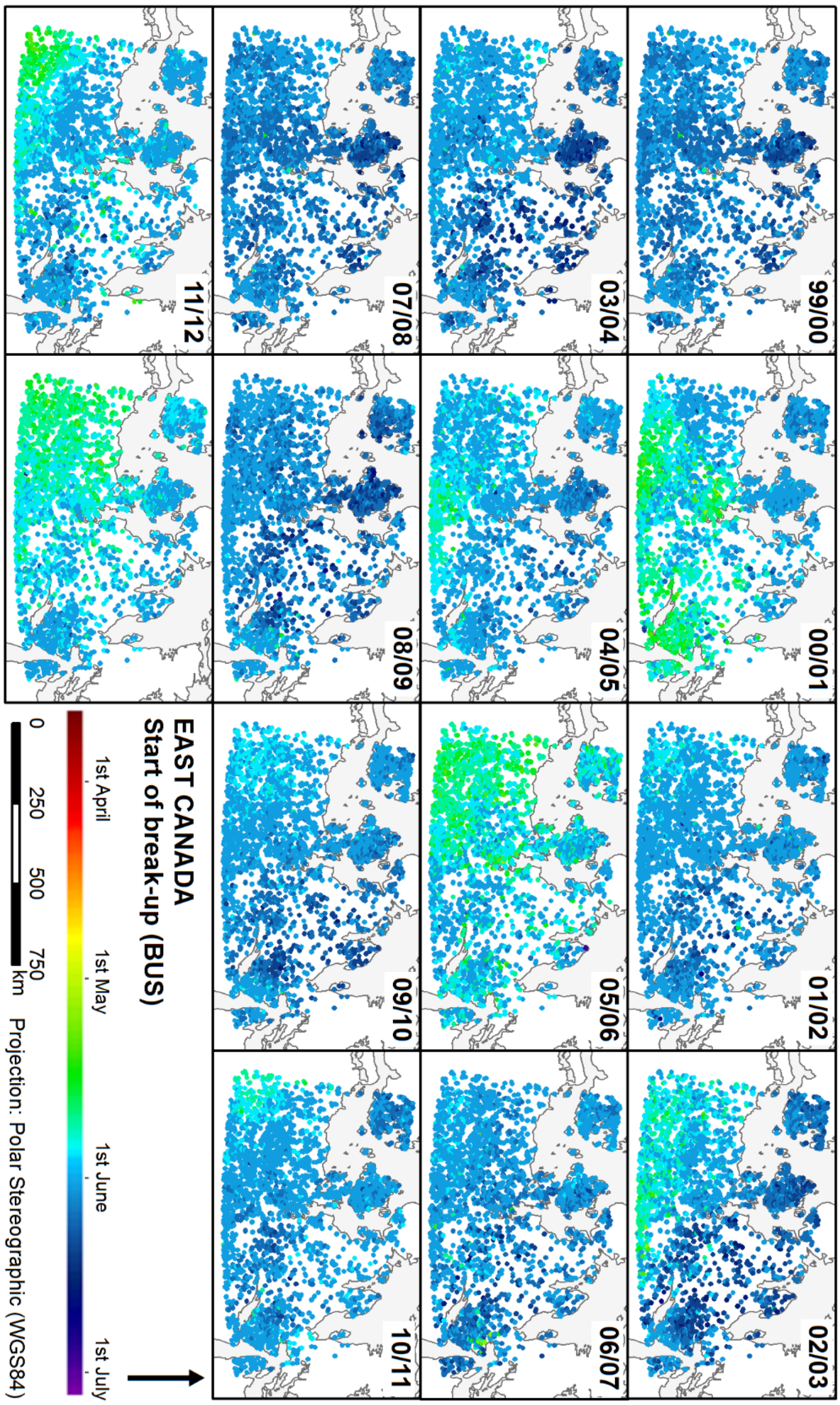

Figure 19: Timing of break-up start (BUS) for East Canada study area for seasons 1999/2000 to 2012/2013

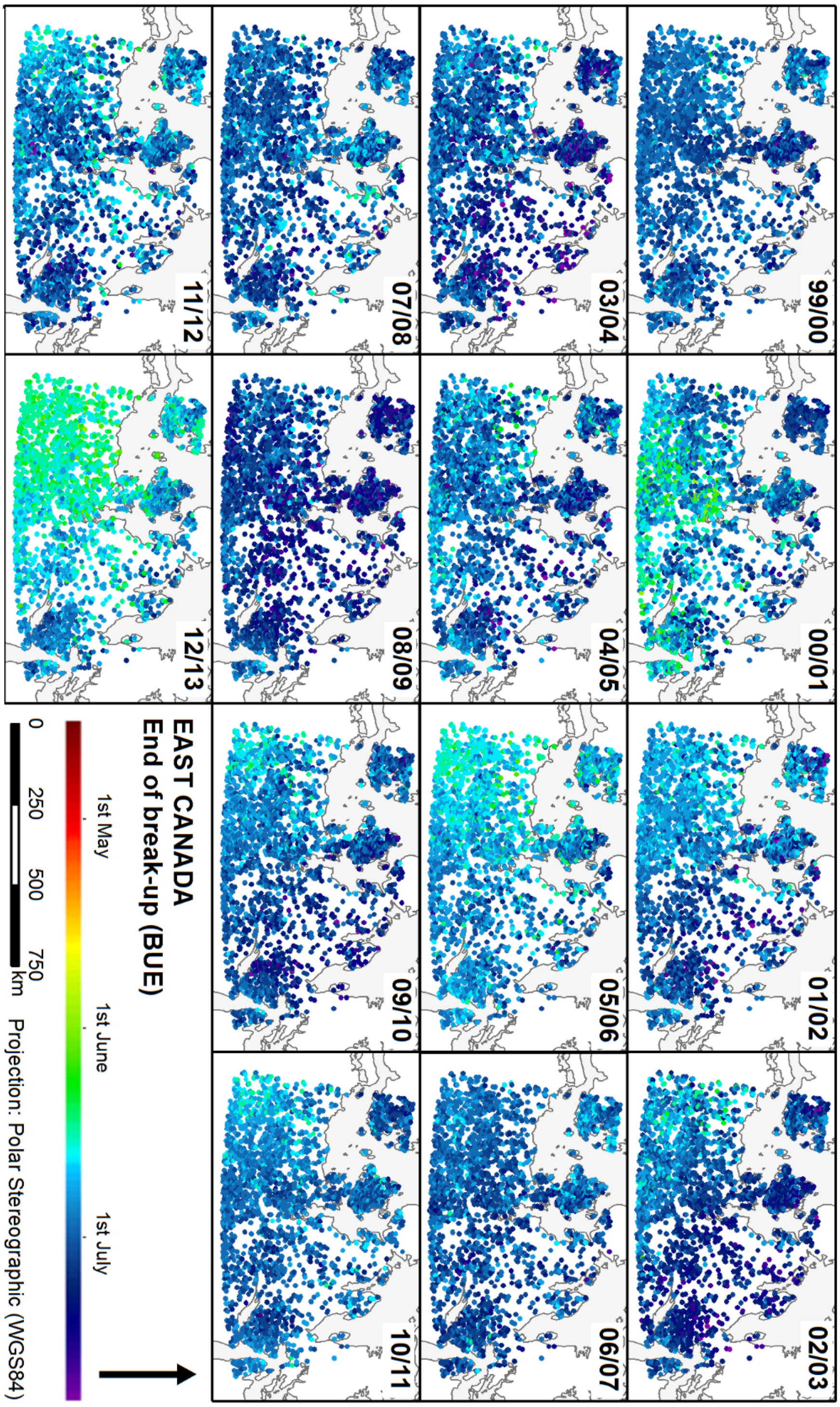

Figure 20: Timing of break-up end (BUE) for East Canada study area for seasons 1999/2000 to 2012/2013

4.1.3.3 Alaska

Figure 21 (p. 71) and Figure 22 (p.72) show spatial distribution of timing of break-up start and break-up end, respectively, for the study area located on Northern Slope of Alaska (NSA) near Barrow. The timing of break-up start ranges from beginning of May to mid-June for lakes along the north-eastern coast. The timing of break up end ranges from early June to early July. The median duration of break-up period (27 days) is equal to the duration observed for Northern Europe study area. Similarly as for East Canada, the spatial distribution follows south-west to north-east gradient. The ice phenology of have been observed to correlate with distance from the coast. Lakes closer to the coast retain the ice cover longer than lakes further inland. (Hinkel et al., 2012) For several years outlying values appear that could not be explained and are most likely a result of an error in extraction.

4.1.3.4 Taymyr

The spatial trends in timing of break-up start and break-up end for study area located south of Lake Taymyr are shown in Figure 23 (p.73) and Figure 24 (p.74), respectively. Break-up start occurs from late May in southernmost regions to late June in the north. Break up end ranges from early June to mid-July. The median duration of break-up period of 27 days is again equal to previous locations. The spatial distribution follows south-east to north-west gradient that can be explained as combination of temperature and altitude. Many lakes appear as outliers from the predominant trend, showing very late break-up in southern locations. When inspected closely over high resolution imagery no differenced from surrounding lakes were discovered. Further analysis of local conditions and surface reflectance profiles would be necessary to determine the causes.

4.1.3.5 Yedoma

Figure 25 (p.75) and Figure 26 (p.76) show spatial distribution of timing of break-up start and break-up end, respectively, for the Russian Yedoma area. The spatial distribution follows expected latitudinal trend combined with the effect of elevation. The timing of break-up start ranges from end of April for lakes low lying lakes in the basin of Kolyma River to late June for northern lakes. The timing of break up end ranges from late May to early July. The median duration of break-up period of 25 is slightly lower than for other study areas. Due to time constraint the spatial patterns of break-up period duration were not analysed.

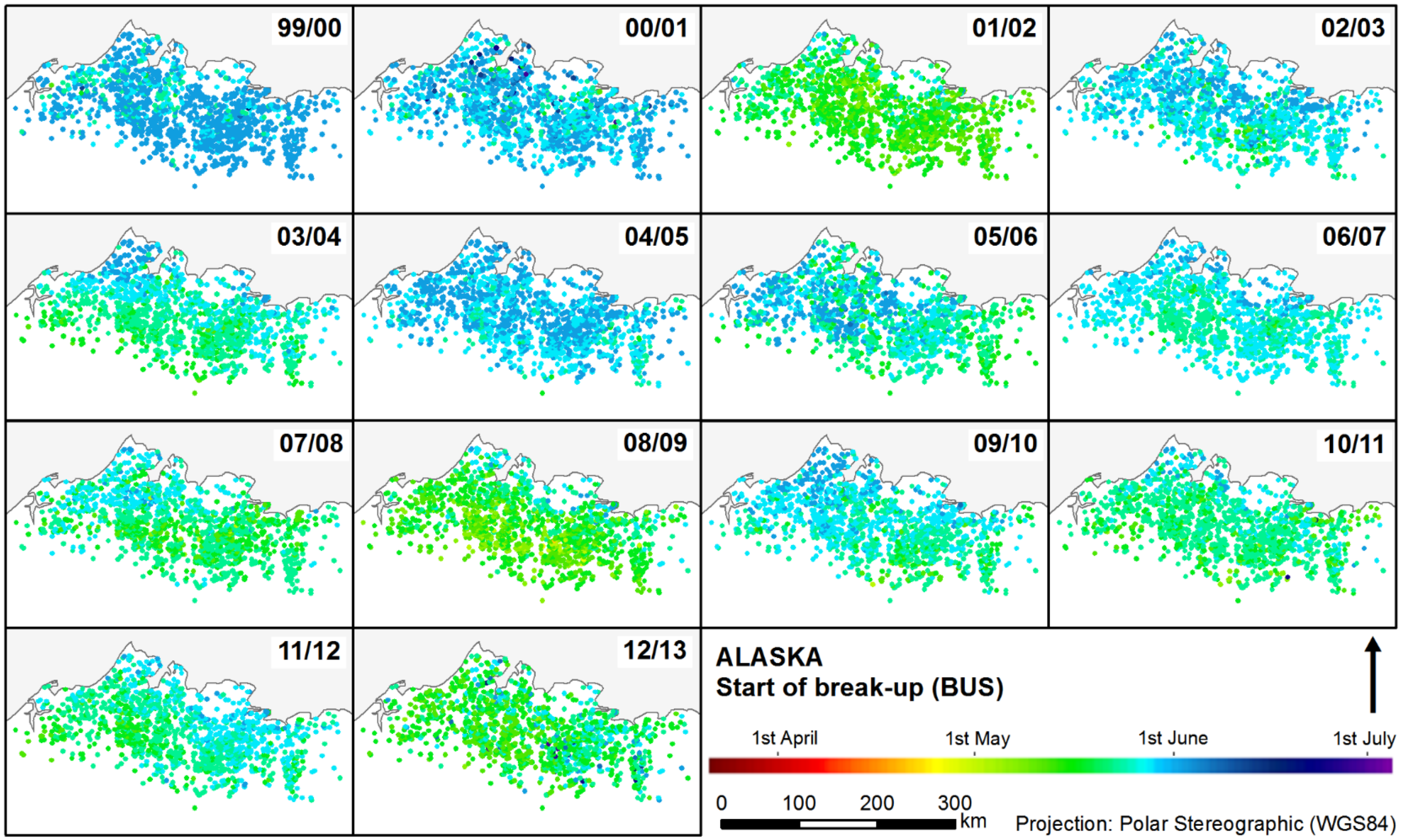

Figure 21: Timing of break-up start (BUS) for Alaska study area for seasons 1999/2000 to 2012/2013

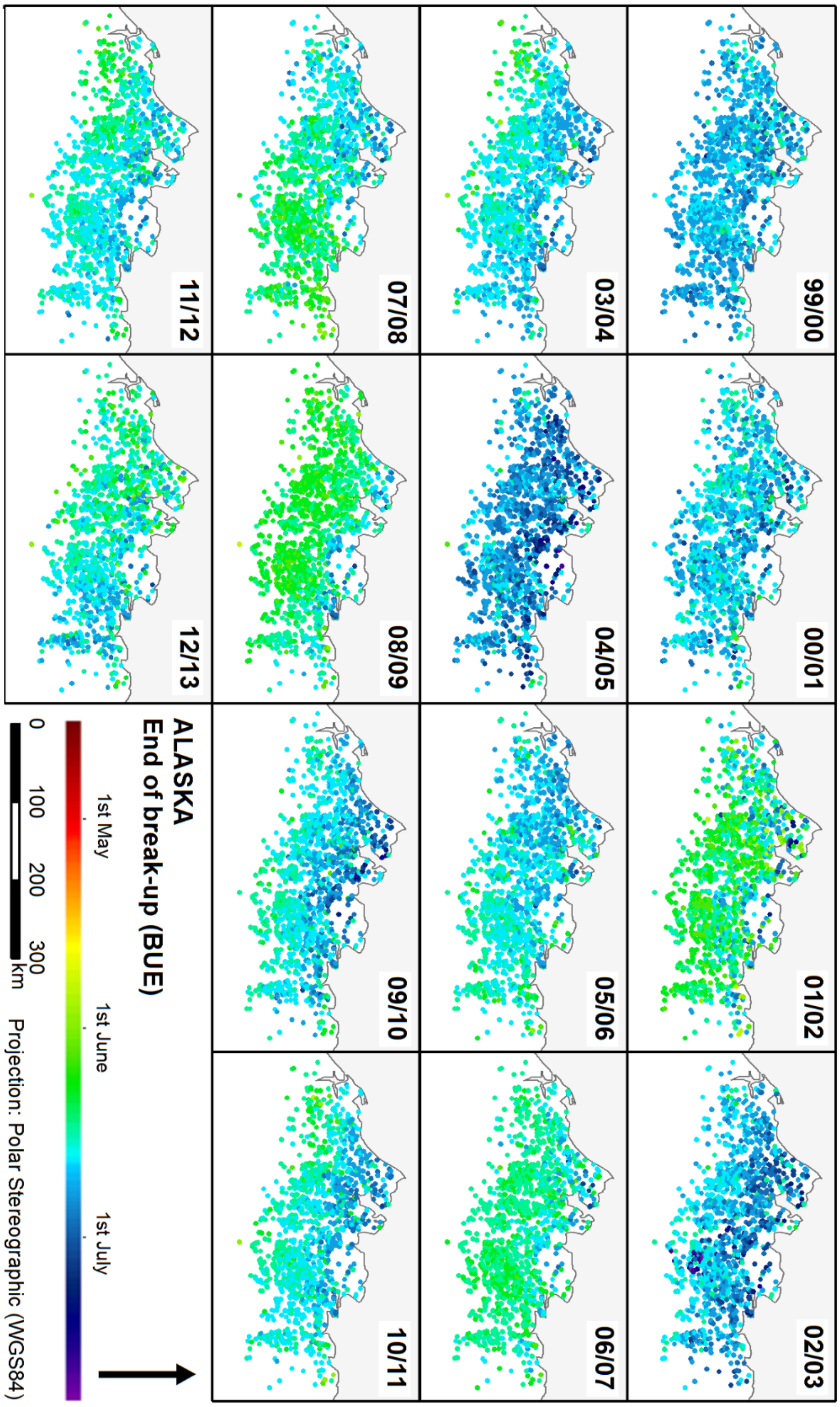

Figure 22: Timing of break-up end (BUE) for Alaska study area for seasons 1999/2000 to 2012/2013

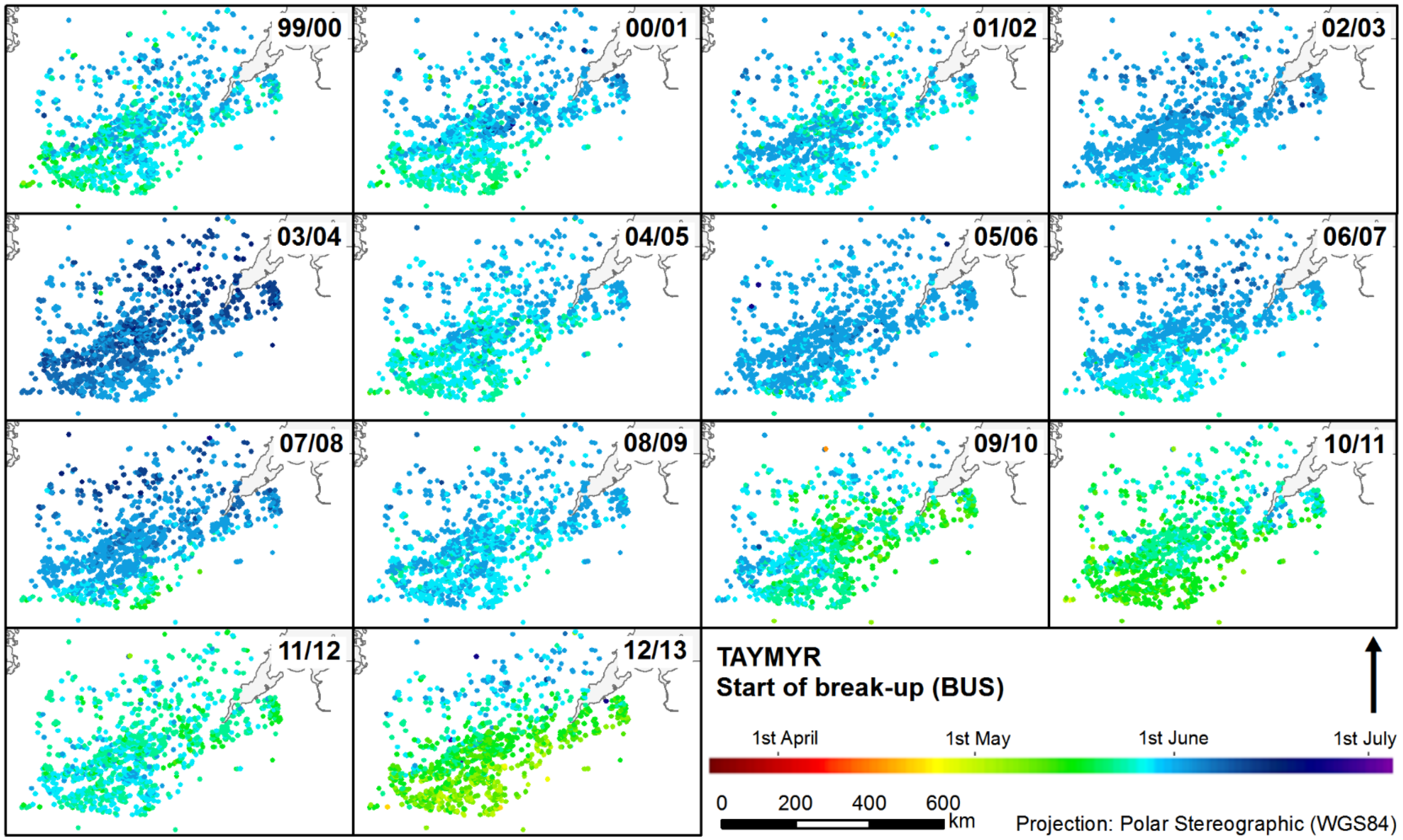

Figure 23: Timing of break-up start (BUS) for Taymyr study area for seasons 1999/2000 to 2012/2013

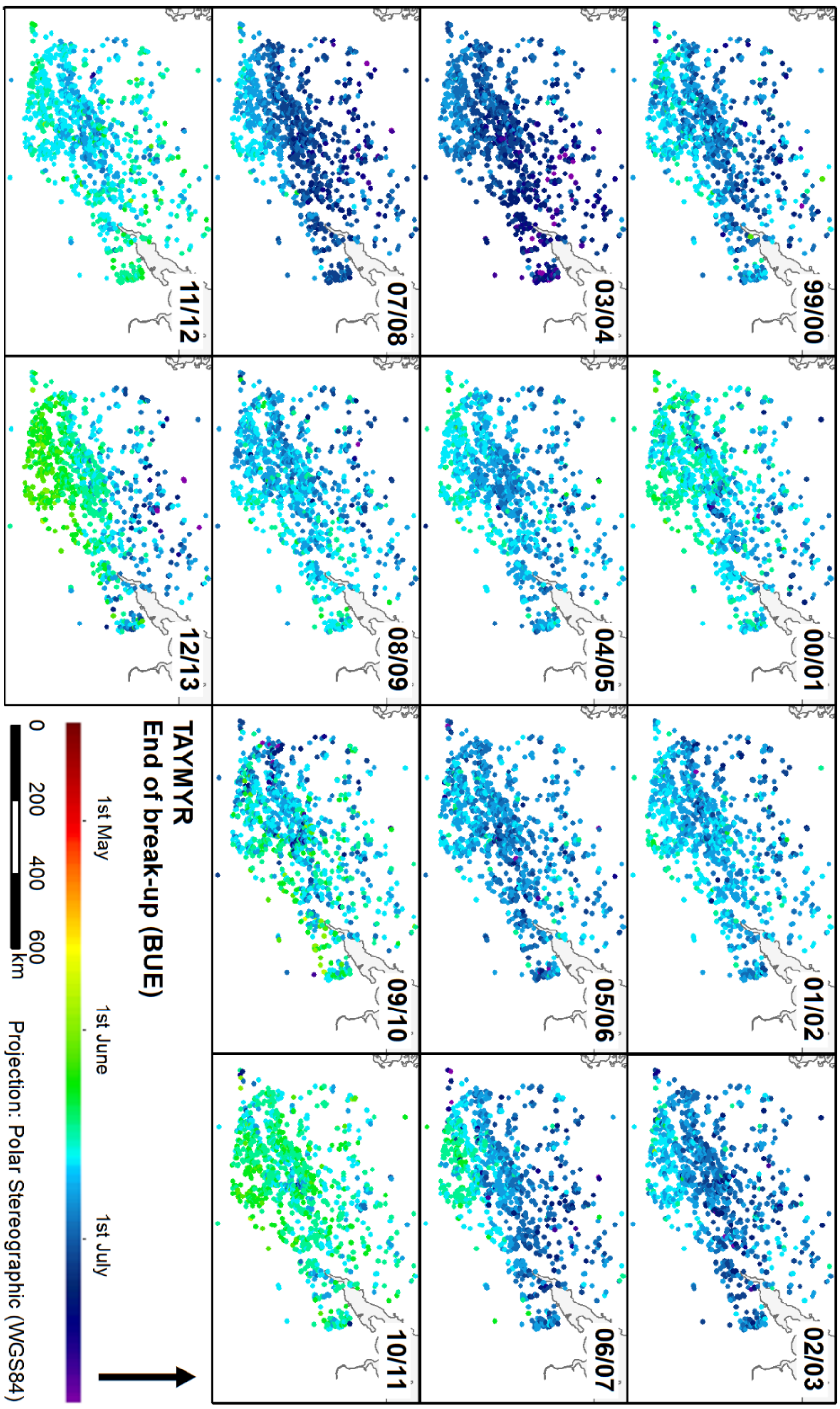

Figure 24: Timing of break-up end (BUE) for Taymyr study area for seasons 1999/2000 to 2012/2013

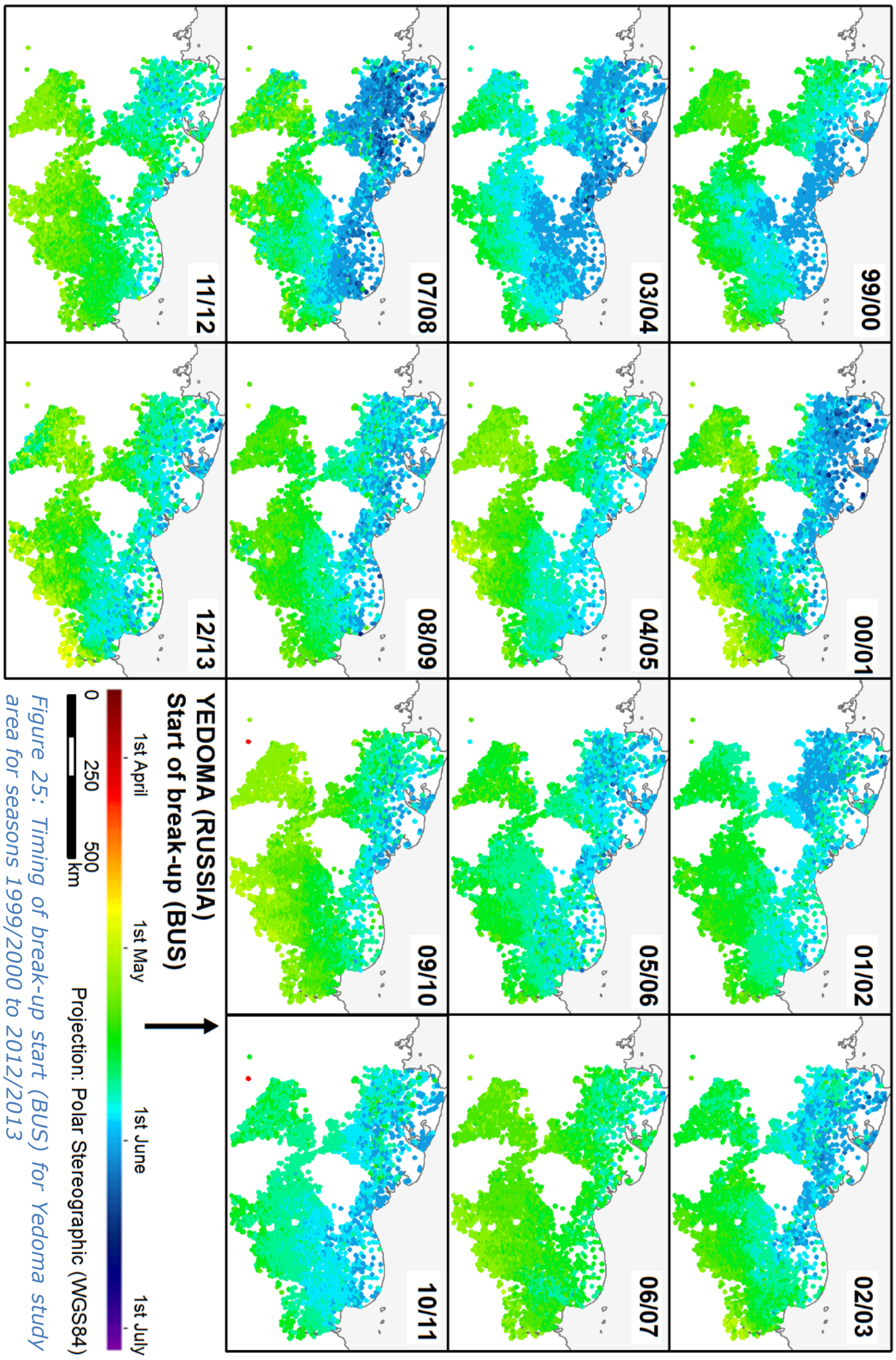

Figure 25: Timing of break-up start (BUS) for Yedoma study area for seasons 1999/2000 to 2012/2013

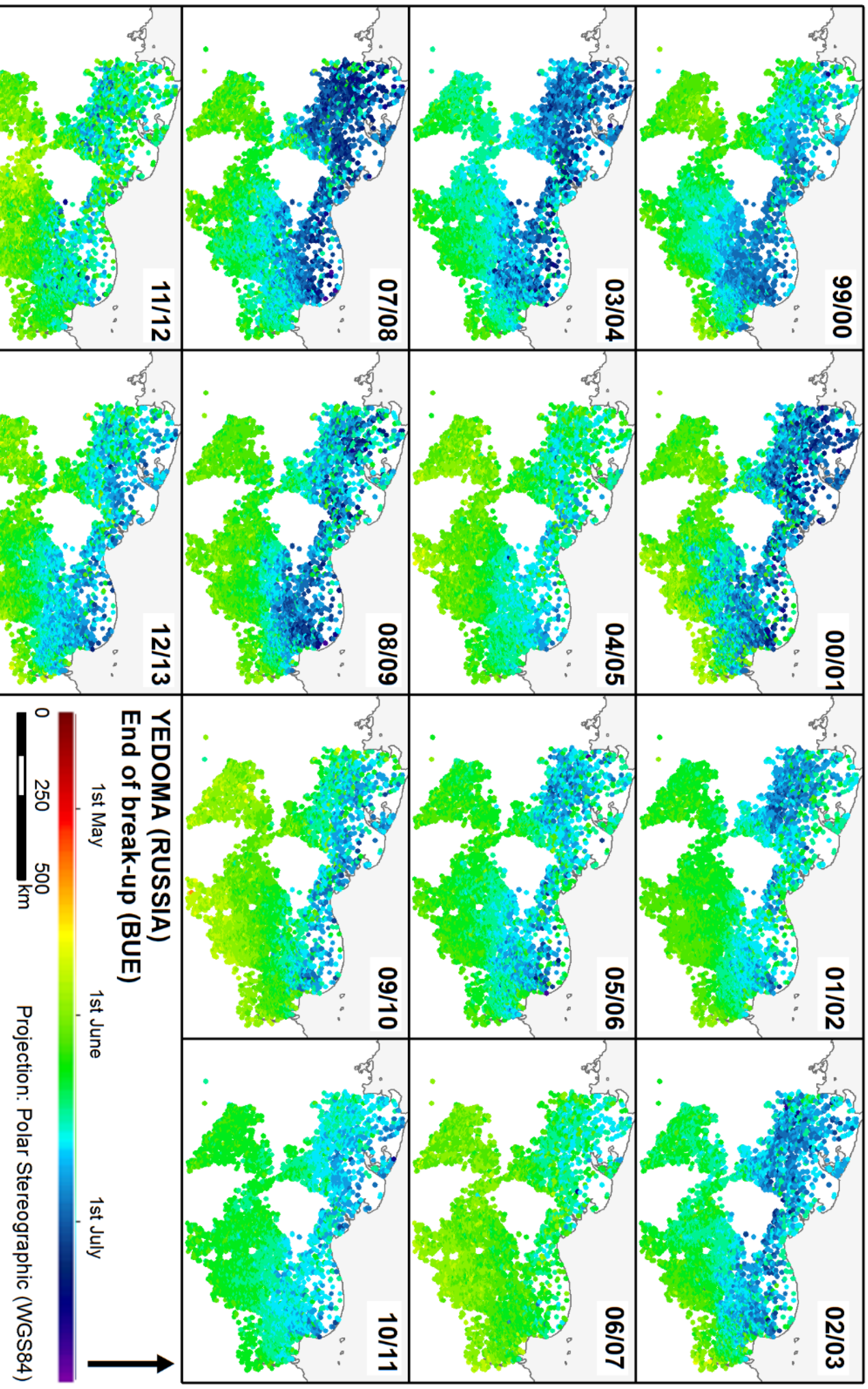

Figure 26: Timing of break-up end (BUE) for Yedoma study area for seasons 1999/2000 to 2012/2013

4.2 Trend analysis

The trend analysis was performed for the timing of break-up start and break-up end over the period 2000 - 2013, however, the emphasis is given on the break-up end (BUE) as the extracted values are the most reliable based on comparison with *in-situ* observations. No measure of reliability is available for BUS, however, the start of break- up appears as distinct feature in the surface reflectance profiles. For many lakes no significant trend was observed, which can be partly attributed to very short time period. The series length is found to strongly affect the significance of a trend tested by Mann-Kendall test.

4.2.1 Northern Europe

Very small number significant trends have been identified for the North Europe. Out of 1809 lakes for which the BUE trend was analysed 18 lakes show trend statistically significant at 95% confidence level (Figure 27) 14 lakes show trend towards earlier break-up between -0.5 to -1.3 days per year with mean of -0.89 days/year.

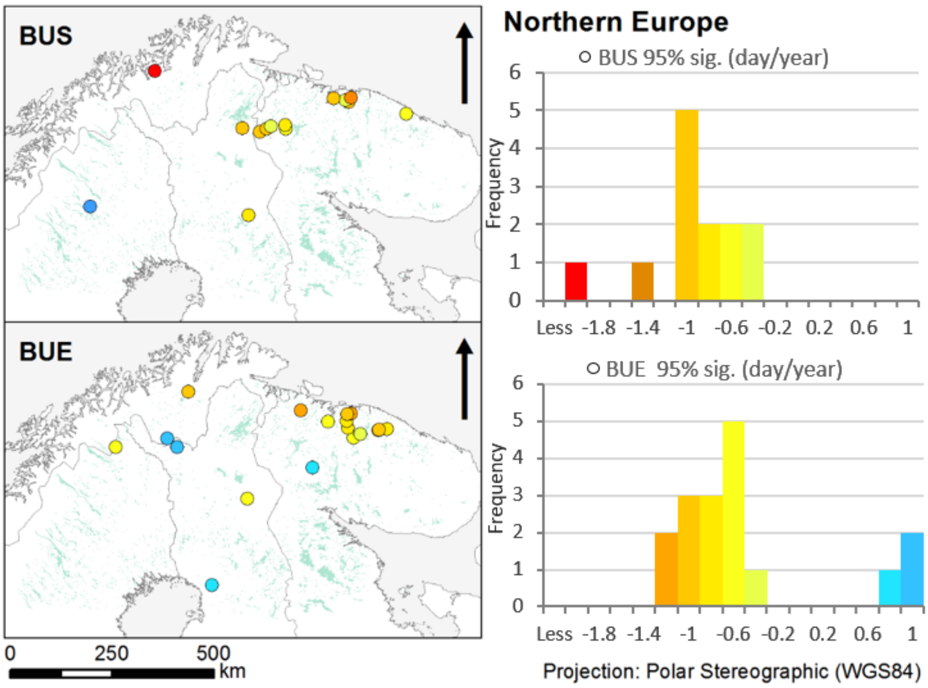

Figure 27: Northern Europe study area - Trends significant at min. 95% confidence level for BUS and BUE.

Majority of lakes showing significant negative trend are located in the northern half of the study area and clustered along the northern coast of the Kola Peninsula. 4 lakes show positive shift toward later break-up at rate -0.72 to -0.85 days per year.

For the start of break-up period 15 lakes show significant trend at 95% conf. level. (Figure 28) For all, but one lake, trend towards earlier break-up start was observed. The negative rate of change varies between -0.5 and -1.5 days/year. The mean rate of change for lakes with negative significant trend is -0.84 days/year. For one lake, trend toward later break-up was observed at rate of 1.09 days/year. It is Lake Randijaure for which *in-situ* data were available. The *in-situ* record for this lake does not indicate any significant trend and therefore the identified positive trend might be attributed to error in extracted dates.

Interestingly only five lakes show significant trend for both BUS and BUE events and then only if trends significant at 90% confidence level are included. All five lakes are relatively small from 2-4 km^2 and are located along the northern coast of Kola Peninsula. The inter-annual variability between the BUE and BUS dates follows similar pattern for each lake. (Figure 28)

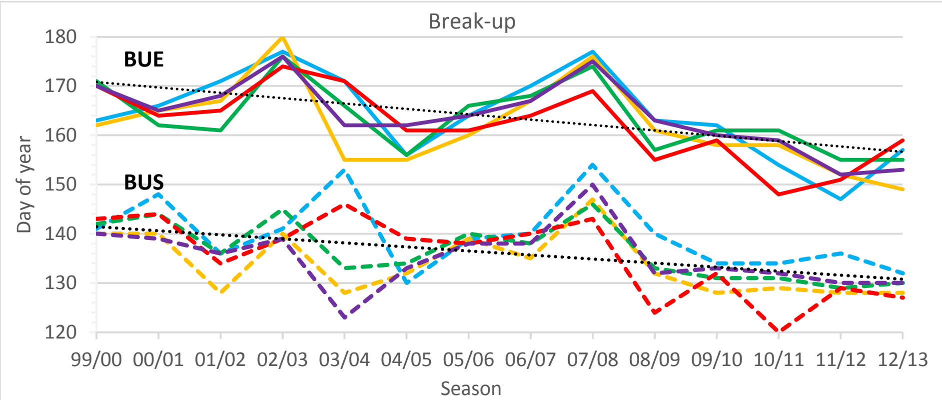

Figure 28: Example of time series for break-up events for 5 lakes showing significant trend for both events (90% conf. int.)

Limited number of studies have inspected trends for European lakes north of 65° latitude. Korhonen (2006) reports significant negative trends for southern and central part of Finland (1885-2000), however, no significant trends were observed for analyzed lakes in Lapland. Similarly, Efremova et al. (2013) reported earlier break-up for lakes in South Karelia (1950-2009) but no significant trends for one analyzed northern lake. Other studies including high latitude European lakes report the same results. (Benson et al., 2012; T. Blenckner et al., 2004) Only Lei et al. (2012) reported a significant trend for Kilpisjärvi, relatively large, deep lake located in mountainous area on border between Norway and Sweden. All mentioned studies however used solely available *in-situ* observations and concentrated on small number of large lakes over different and much longer time period than studied here. In agreement with literature, no significant trends were found for these lakes. All lakes showing a trend have relatively small surface area and majority is located at the north-east coast of Kola peninsula, which is the only part of the study area showing significant trend towards higher mean annual temperatures for period of 2000-2013.

4.2.2 East Canada

For East Canada trends for 3021 lakes were analysed and 79 show shift towards earlier end of break-up statistically significant at 95%. (Figure 29) The mean rate of change for is -1.05 days/year (-0.25 to -2.0). For four lakes significant positive trend was identified between 0.84 and 1.18 days/year. The number of identified trends (54) is slightly lower for the beginning of break-up period with average change of 0.92 days/year (-0.4 to -2.26). Only one lake with positive trend was identified.

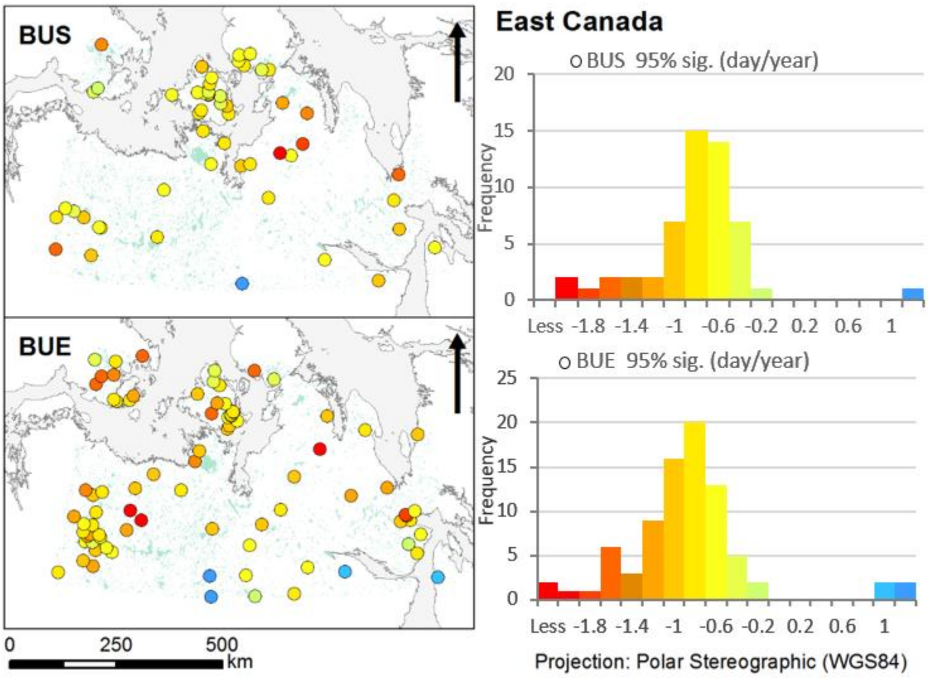

Figure 29: East Canada study area - Trends significant at min. 95% confidence level

The highest concentration of lakes with significant negative trend for either BUS or BUE was observed for northern islands (Victoria and King

William – latest break-up) and for south-east of the study area (earliest break-up). Number of lakes located on the northern islands show significant shift for the timing of whole break-up period (both BUS and BUE).

Observed trends are in agreement with trends reported in literature for Canadian Arctic. Latifovic & Pouliot (2007) analysed ice phenology of 6 Canadian high latitude lakes using AVHRR surface reflectance data and identified mean trend of -0.99 days/year over the period 1985-2004.

4.2.3 Alaska

Out of 1237 lakes, analysed for study area located on Alaskan Arctic coastal plain, 115 indicated change towards earlier timing BUE with mean rate of -0.88 days/year (Figure 30).

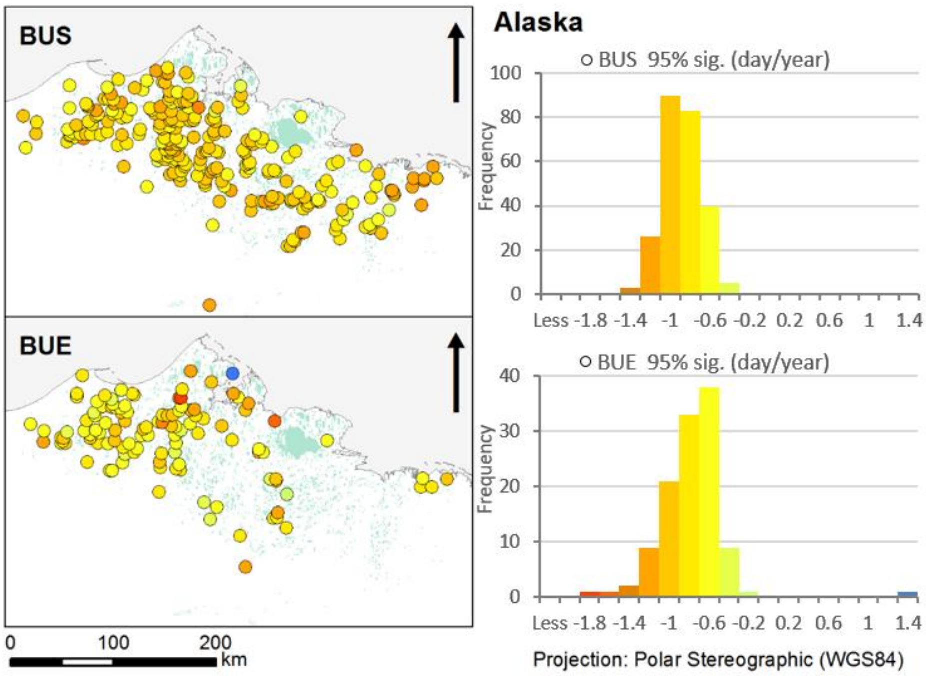

Figure 30: Alaska study area – Trends significant at min. 95% confidence level

The number of identified trends (247) significantly increased for the start of break-up period. The observed rate of change (mean 0.94 days/year) was relatively homogenous for the entire area. One lake located about 1km from the coast of Admiralty Bay, showed significant positive trend of 1.4 days/year for BUE. Upon investigation of the BUE time series extracted for this lake, it appears to break-up significantly earlier than surrounding lakes in some years while in others it follows the same timing as other lakes in the area.

Recently Surdu et al. (2014) examined the response of lakes in northern part of this study area, near Barrow point, to observed change in climatic conditions in this region for period 1950 to 2011. They identified overall shift in break-date by 17.7-18.6 days over the study period corresponding to 0.29-0.3 days/year, change much slower than observed here. The reason for this difference can be explained significant difference in the length of studied period. Moreover, the period 1950 to 1976 has been reported significantly colder than the period 1977 to present (Figure 31), therefore the trends in ice phenology including this period would necessarily be lower than trends for period after 1977. (Alaska Climate Research Center, 2014)

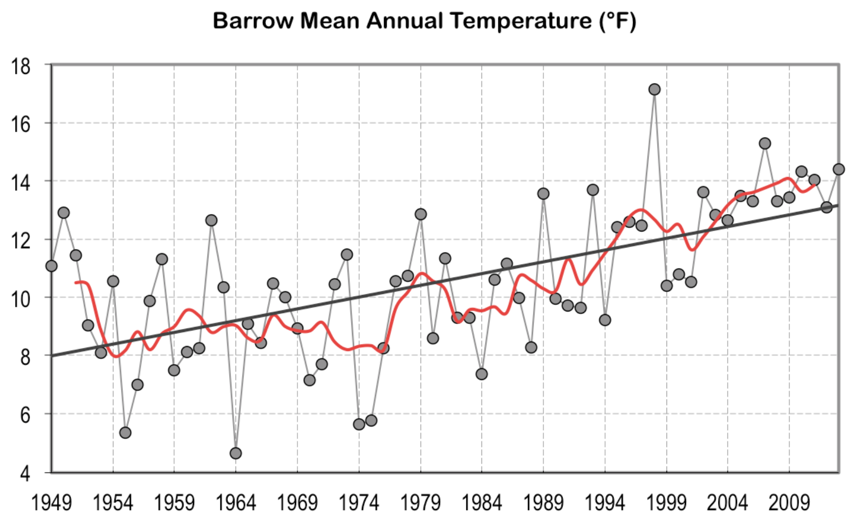

Figure 31: Annual mean temperature (1950-2011) at National Weather Service station, Barrow, Alaska. (Alaska Climate Research Center, 2014)

4.2.4 Taymyr

Out of lakes 1237 south of Lake Taymyr in Northern Russia 143, have shown significant trend toward earlier break-up end and 414 toward earlier break-up start. For 134 lakes trend was identified for both, BUS and BUE. The mean rate of change, 1.10 days/year (-0.5 to 2.27 days/year) for BUE and 1.40 days/year (-0.52 to -3.14 days/year) for BUS, is higher than observed for other study areas. Lakes showing trend are located mainly in lowlands north of Central Siberian Plato. Area of higher rate of change can be discerned in the around the delta of Kolyma river. Number of seemingly randomly distributed lakes shows extreme rates of change <-1.8 days/year. Upon simple inspection over high resolution imagery, no explanation was found. Four lakes show very high values for BUS<-3 days/year, such rapid change is unlikely and is most likely product of an uncertainties in extraction process. No positive trends were observed in the area.

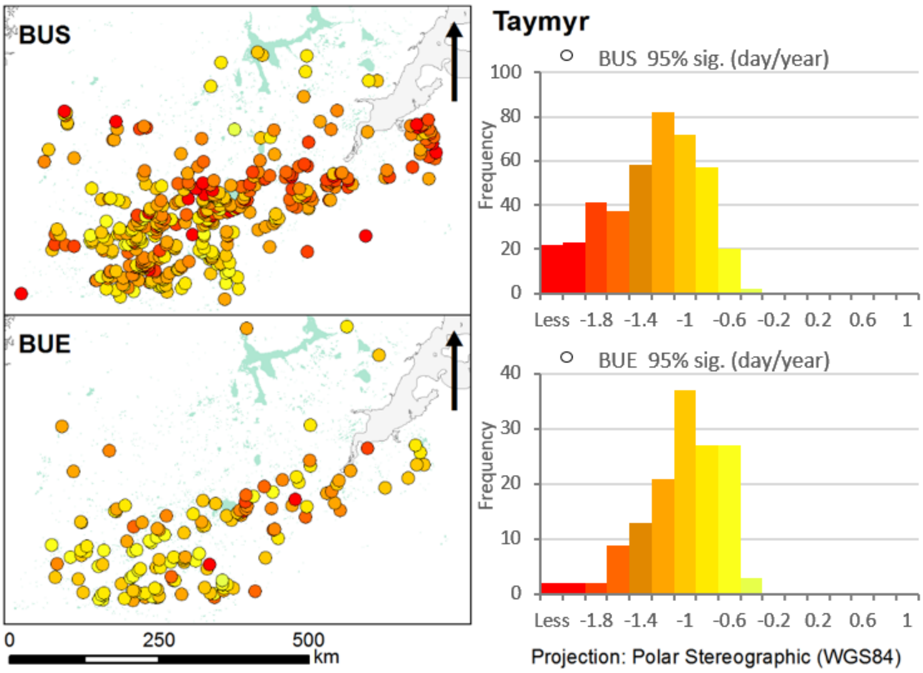

Figure 32: Taymyr study area – Trends significant at min. 95% confidence level

4.2.5 Yedoma

For the Yedoma region in north eastern Siberia 6039 lakes were analysed, 532 lakes show significant trend for break-up towards earlier date at mean rate -0.95 days/year (-0.33 to -3.09 days/year). 9 lakes have positive trend in rage from 0.62 to 1.68 days/year. For the start of break up period, significant negative trend was identified for 310 lakes with mean rate of -0.85 days/year. BUS -0.33 to -1.55 days/year. Only one lake showed positive trend of 0.83 days/year. If extreme and positive values are disregarded weak south-east to north-west trend in trend magnitude can be detected. The spatial distribution of positive values appears random even upon closer inspection. 99 lakes had significant negative trends for both, BUS and BUE.

The Yedoma area was the only study area where sufficient number of large lakes (>20km^2) shoved significant trend, to analyse dependence of trend magnitude on lake size. However, no statistically significant difference was observed between magnitude of change for small lakes (<2 km^2) and large lakes (>20 km^2).

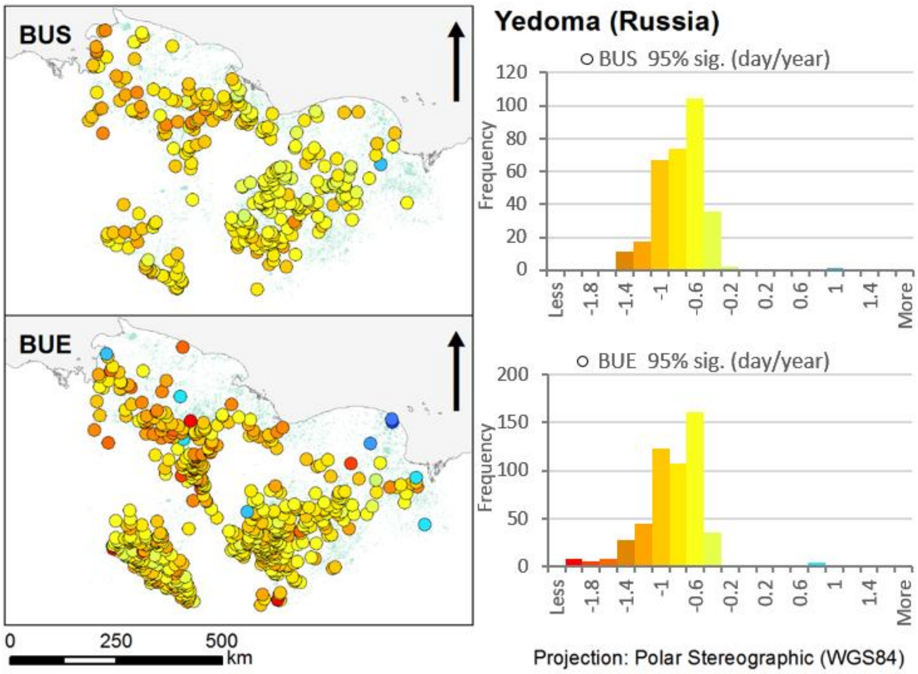

Figure 33: Yedoma study area – Trends significant at min. 95% confidence level

4.3 Regression

The formation and decay of lake ice cover are in their essence very different processes governed by different factors. The final objective of this research was to determine which climatic and non-climatic factors play the main role in determining the timing of each phenological and whether their influence is different between different study areas. In total eight factors were considered, air temperature, precipitation, snow depth, wind speed, size and latitude. For each factor several variables were tested and are somatised in Table 2 (*Methodology*). As stated previously, due to low confidence given to extracted timing of freeze-up start and end, only break-up start and break-up end will be discussed.

In agreement with literature the major determinant of end of break-up period is temperature averaged over 45 (Temp. 45) days before the event. Although the portion in variance explained ty temperature over 45 days alone, is lower than reported by D. M. Livingstone (1997) or Palecki & Barry (1986) who reported up to 70% explained variability. The timing of spring 0°C isotherm as characteristics of annual temperature pattern explained additional portion of variance for all locations except Alaska. For Yedoma timing of 0° isotherm accounted for greater portion of variance than Temp. 45. Apart from temperature latitude improved the explanatory power of the model significantly for Europe and Alaska and Yedoma but explained little variance for other areas. This is supported by findings of Weyhenmeyer et al. (2011) who used latitude as a proxy for solar radiation for a model tested on European lakes.

Significant differences have been found between study areas. For East Canada only temperature variables explained 78% of variability with no detectable effect of latitude. On contrary, for Alaska larger portion of variability was explained by latitude than by temperature. For Yedoma the main determining variable was the timing of spring 0°C isotherm.

	Temp. 45	Latitude	0° isotherm	Amplitude	Precip. 15	Adj. R^2
Europe	0.52	0.18	0.07	0.03	-	0.80
East Canada	0.61	-	0.08	0.03	0.04	0.78
Alaska	0.34	0.37	-	0.04	-	0.75
Taymyr	0.48	0.03	0.24	-	-	0.75
Yedoma	0.28	0.12	0.42	-	-	0.82

Table 3: Summary of main variable determining the end of break-up period and portion of explained variance within the model.

The start of break-up period showed similar results as BUE, however for Alaska and Taymyr temperatures over shorter (15 and 30 days) period have more impact. Interestingly for break-up start in Alaska latitude seem to play no significant role and were not able to explain any variance in BUS timing. Even larger variance in BUS dates in Yedoma is explained by timing of spring 0°C isotherm.

	Temp. 45	Latitude	0° isotherm	Temp. 15	Temp. 30	Precip. 30	Adj. R^2
Europe	0.58	0.15	0.03	-	-	-	0.75
East Canada	0.58	0.02	0.13	-	-	-	0.73
Alaska	-	-	0.12	-	0.37	0.04	0.55
Taymyr	-	-	0.20	0.52	-	-	0.72
Yedoma	0.17	0.08	0.57	-	-	-	0.82

Table 4: Summary of main variable determining the start of break-up period and portion of explained variance within the model.

Overall up to 82% (Yedoma) of variance in both break-up start and break-up end dates have been explained by combination of temperature and latitude. Climatic data used in this study were not lake specific but had rather coarse spatial resolution of 0.75 degrees. This poses a problem especially in heterogeneous areas such as Northern Europe.

5 Conclusions and Recommendations

The first objective of this research was to develop automated method to extract Arctic lake ice phenology from MODIS surface reflectance data. Due to time constraint five sample areas were selected approximately evenly distributed around the circumpolar Arctic. The entire process from MODIS imagery download to the phenology extraction has been successfully automated, although, it still requires significant amount of user control. The use of lake surface reflectance profiles greatly reduces the volume of data to be processed and is the only feasible way of analysing lake ice phenology over large areas from optical satellite data.

Comparison of derived phenology and in-situ observations showed that MODIS surface reflectance product with 250m resolution can be successfully used to extract the timing of break-up with good accuracy. However, low sun illumination during fall and winter period hampers the usability of optical imagery for the retrieval of freeze-up timing. Therefore, the agreement with in-situ observations for freeze-up period was low and the results have not been used for further analysis. The results for break-up period could still be improved by implementing more sophisticated extraction method. Overall, while the developed method has a good potential, there is still room for significant improvement such integrating all processing steps into one.

The spatial distributions of break timing has been successfully mapped. In most areas it has been found to follow expected south-north gradient, however, the effect of altitude appears equally important. In East Canada spatial trend following temperature patterns rather than latitude has been observed and confirmed by following regression analysis. Deeper knowledge of spatial distribution of break-up timing can contribute to more accurate estimation of methane emissions from Arctic lakes, especially the timing of the spring pulse release and its distribution in space.

Limited knowledge exists regarding changes in lake phenology of small Arctic lakes. Trend analysis performed on the extracted data showed predominant trend towards earlier break-up in all areas. In agreement with findings reported in literature the least number of trends has been observed for Northern Europe, while in the other areas large number of strong trends were observed. The identified trends were higher overall than trends reported before, however, this can be partly attributed to different time period. For some areas, weak spatial trend towards increase in the rate of change with latitude was observed. However, further monitoring is necessary to develop sufficiently long record to make conclusive decision.

Final objective was to determine the role of various climatic variables on timing of phenological events. The results have confirmed findings reported for mid-latitude lakes. Temperature is the major determinant for the timing of break-up start and end of break-up period for all the studied locations. Up to 60% of variance on timing of break-up end can be explained by temperature over 45 proceeding days alone. Additional information about the annual temperature cycle can further improve the portion of explained variance in all areas except Europe. Latitude plays the second most important role in break-up timing for all areas except East Canada, where break-up appears to be solely dependent on temperature patterns. Use of higher resolution climatic data would most likely significantly improve the explained variability and identify other determining variables, effects of which are not detected due to coarse resolution.

Overall the main objectives of this study have been achieved successfully. Many improvements can be made to the extraction process including alternative ways of detecting freeze-up timing during the period of low sun elevation. The new Sentinel-1 SAR sensor might offer opportunities for incorporating radar data and interesting direction for further research.

References

Adams, W. P., & Roulet, N. T. (1980). Illustration of the Roles of Snow in the Evolution of the Winter Cover of a Lake. *ARCTIC*, *33*(1), 100–116. doi:10.14430/arctic2550

Alaska Climate Research Center. (2014). Temperature Changes in Alaska. Retrieved May 26, 2014, from http://climate.gi.alaska.edu/ClimTrends/Change/TempChange.html

Benson, B. J., Magnuson, J. J., Jensen, O. P., Card, V. M., Hodgkins, G., Korhonen, J., … Granin, N. G. (2012). Extreme events, trends, and variability in Northern Hemisphere lake-ice phenology (1855–2005). *Climatic Change*, *112*(2), 299–323. doi:10.1007/s10584-011-0212-8

Bieniek, P. A., Bhatt, U. S., Rundquist, L. A., Lindsey, S. D., Zhang, X., & Thoman, R. L. (2011). Large-Scale Climate Controls of Interior Alaska River Ice Breakup. *Journal of Climate*, *24*(1), 286–297. doi:10.1175/2010JCLI3809.1

Blenckner, T., Adrian, R., Arvola, L., Järvinen, M., Nõges, P., Nõges, T., … Weyhenmeyer, G. A. (2010). The Impact of Climate Change on Lakes in Northern Europe. In G. George (Ed.), *The Impact of Climate Change on European Lakes* (pp. 339–358). Springer Netherlands.

Blenckner, T., Adrian, R., Livingstone, D. M., Jennings, E., Weyhenmeyer, G. A., George, D. G., … Teubner, K. (2007). Large-scale climatic signatures in lakes across Europe: a meta-analysis. *Global Change Biology*, *13*(7), 1314–1326. doi:10.1111/j.1365-2486.2007.01364.x

Blenckner, T., Järvinen, M., & Weyhenmeyer, G. A. (2004). Atmospheric circulation and its impact on ice phenology in Scandinavia. *Boreal environment research*, *9*(5), 371–380.

Bonsal, B. R., & Prowse, T. D. (2003). Trends and Variability in Spring and Autumn 0 °C-Isotherm Dates over Canada. *Climatic Change*, *57*(3), 341–358. doi:10.1023/A:1022810531237

Bonsal, B. R., Prowse, T. D., Duguay, C. R., & Lacroix, M. P. (2006). Impacts of large-scale teleconnections on freshwater-ice break/freeze-up dates over Canada. *Journal of Hydrology*, *330*(1-2), 340–353. doi:10.1016/j.jhydrol.2006.03.022

Bronaugh, D., & Werner, A. (2013). Zhang + Yue-Pilon trends package. Retrieved from http://cran.r-project.org/web/packages/zyp/index.html

Brown, L. C., & Duguay, C. R. (2010). The response and role of ice cover in lake-climate interactions. *Progress in Physical Geography*, *34*(5), 671–704. doi:10.1177/0309133310375653

Brown, L. C., & Duguay, C. R. (2011). The fate of lake ice in the North American Arctic. *The Cryosphere*, *5*(4), 869–892. doi:10.5194/tc-5-869-2011

Brown, L. C., & Duguay, C. R. (2012). Modelling Lake Ice Phenology with an Examination of Satellite-Detected Subgrid Cell Variability. *Advances in Meteorology*, *2012*, 1–19. doi:10.1155/2012/529064

Callaghan, T. V., Johansson, M., Key, J., Prowse, T., Ananicheva, M., & Klepikov, A. (2011). Feedbacks and Interactions: From the Arctic Cryosphere to the Climate System. *Ambio*, *40*(Suppl 1), 75–86. doi:10.1007/s13280-011-0215-8

Callaghan, T. V., Johansson, M., Prowse, T., Olsen, M. S., & Reiersen, L.-O. (2012). Arctic Cryosphere: Changes and Impacts. *AMBIO*, *40*(S1), 3–5. doi:10.1007/s13280-011-0210-0

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

CIS. (2005). *Manual of Standard Procedures for Observing and Reporting Ice Conditions (MANICE)* (9th ed.). Canadian Ice Service – Environment Canada

Cook, T. L., & Bradley, R. S. (2010). An Analysis of Past and Future Changes in the Ice Cover of Two High-Arctic Lakes Based on Synthetic Aperture Radar (SAR) and Landsat Imagery. *Arctic, Antarctic, and Alpine Research*, *42*(1), 9–18. doi:10.1657/1938-4246-42.1.9

Dibike, Y., Prowse, T., Bonsal, B., Rham, L. de, & Saloranta, T. (2012). Simulation of North American lake-ice cover characteristics under contemporary and future climate conditions. *International Journal of Climatology*, *32*(5), 695–709. doi:10.1002/joc.2300

Dibike, Y., Prowse, T., Saloranta, T., & Ahmed, R. (2011). Response of Northern Hemisphere lake-ice cover and lake-water thermal structure patterns to a changing climate. *Hydrological Processes*, 2942–2953. doi:10.1002/hyp.8068

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

Duguay, C. R., Prowse, T. D., Bonsal, B. R., Brown, R. D., Lacroix, M. P., & Menard, P. (2006). Recent trends in Canadian lake ice cover. *Hydrological Processes*, *20*(4), 781–801. doi:10.1002/hyp.6131

EEA. (2006). *The thematic accuracy of Corine land cover 2000 - Assessment using LUCAS* (No. 7/2006). Copenhagen: European Environment Agency.

Efremova, T., Palshin, N., & Zdorovennov, R. (2013). Long-term characteristics of ice phenology in Karelian lakes. *Estonian Journal of Earth Sciences*, *62*(1), 33–41. doi:10.3176/earth.2013.04

Eklundh, L., & Jönsson, P. (2012). TIMESAT 3.1 Software Manual. Lund University. Retrieved from http://www.nateko.lu.se/timesat

Gebre, S., Boissy, T., & Alfredsen, K. (2013). Sensitivity of lake ice regimes to climate change in the nordic region. *The Cryosphere Discussions*, *7*(2), 743–788.

Girjatowicz, J. P. (2010). Effects of the North Atlantic Oscillation on water temperature in southern Baltic coastal lakes. *Annales de Limnologie - International Journal of Limnology*, *47*(1), 73–84. doi:10.1051/limn/2010031

Hassol, S. J. (2005). *Arctic Climate Impact Assessment - Scientific Report*. Cambridge University Press.

Helfrich, S. R., McNamara, D., Ramsay, B. H., Baldwin, T., & Kasheta, T. (2007). Enhancements to, and forthcoming developments in the Interactive Multisensor Snow and Ice Mapping System (IMS). *Hydrological Processes*, *21*(12), 1576–1586. doi:10.1002/hyp.6720

Hindberg, H., Malnes, E., Asalmi, H., & Mattila, O.-P. (2012). Validation of Envisat ASAR based lake ice maps on lake Päijänne. In *2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 5238–5241). New Tork, USA: IEEE.

Hinkel, K. M., Lenters, J. D., Sheng, Y., Lyons, E. A., Beck, R. A., Eisner, W. R., … Potter, B. L. (2012). Thermokarst Lakes on the Arctic Coastal Plain of Alaska: Spatial and Temporal Variability in Summer Water Temperature: Arctic Lakes Water Temperature. *Permafrost and Periglacial Processes*, *23*(3), 207–217. doi:10.1002/ppp.1743

Hurrell, J. W. (1996). Influence of variations in extratropical wintertime teleconnections on northern hemisphere temperature. *Geophysical Research Letters*, *23*(6), 665–668. doi:10.1029/96GL00459

IPCC. (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. (T. F. Stocker, D. Qin, G.-K. Pattner, M. Tignor, S. K. Allen, J. Boschung, … P. M. Midgley, Eds.). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Jeffries, M. O., Morris, K., & Kozlenko, N. (2005). Ice Characteristics and Processes, and Remote Sensing of Frozen Rivers and Lakes.

In C. R. Duguay & A. Pietroniro (Eds.), *Remote Sensing in Northern Hydrology: Measuring Environmental Change* (pp. 63–90). American Geophysical Union.

Jensen, O. P., Benson, B. J., Magnuson, J. J., Card, V. M., Futter, M. N., Soranno, P. A., & Stewart, K. M. (2007). Spatial analysis of ice phenology trends across the Laurentian Great Lakes region during a recent warming period. *Limnology and Oceanography*, *52*(5), 2013–2026. doi:10.4319/lo.2007.52.5.2013

Jönsson, P., & Eklundh, L. (2004). TIMESAT—a program for analyzing time-series of satellite sensor data. *Computers & Geosciences*, *30*(8), 833–845. doi:10.1016/j.cageo.2004.05.006

Jorgenson, M. T., & Shur, Y. (2007). Evolution of lakes and basins in northern Alaska and discussion of the thaw lake cycle. *Journal of Geophysical Research: Earth Surface*, *112*(F2), n/a–n/a. doi:10.1029/2006JF000531

Kheyrollah Pour, H., Duguay, C., Martynov, A., & Brown, L. C. (2012). Simulation of surface temperature and ice cover of large northern lakes with 1-D models: a comparison with MODIS satellite data and in situ measurements. *Tellus A*, *64*(0). doi:10.3402/tellusa.v64i0.17614

Kirillin, G., Leppäranta, M., Terzhevik, A., Granin, N., Bernhardt, J., Engelhardt, C., … Zdorovennov, R. (2012). Physics of seasonally ice-covered lakes: a review. *Aquatic Sciences*, *74*(4), 659–682. doi:10.1007/s00027-012-0279-y

Korhonen, J. (2006). Long-term changes in lake ice cover in Finland*. *Nordic Hydrology*, *37*(4-5), 71–78. doi:10.2166/nh.2006.019

Kropáček, J., Maussion, F., Chen, F., Hoerz, S., & Hochschild, V. (2013). Analysis of ice phenology of lakes on the Tibetan Plateau from MODIS data. *The Cryosphere*, *7*(1), 287–301. doi:10.5194/tc-7-287-2013

Latifovic, R., & Pouliot, D. (2007). Analysis of climate change impacts on lake ice phenology in Canada using the historical satellite data record. *Remote Sensing of Environment*, *106*(4), 492–507. doi:10.1016/j.rse.2006.09.015

Latifovic, R., Trishchenko, A. P., Chen, J., Park, W. B., Khlopenkov, K. V., Fernandes, R., … Cihlar, J. (2005). Generating historical AVHRR 1 km baseline satellite data records over Canada suitable for climate change studies. *Canadian Journal of Remote Sensing*, *31*(5), 324–346. doi:10.5589/m05-024

Launiainen, J., & Cheng, B. (1998). Modelling of ice thermodynamics in natural water bodies. *Cold Regions Science and Technology*, *27*(3), 153–178. doi:10.1016/S0165-232X(98)00009-3

Lehner, B., & Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, *296*(1-4), 1–22. doi:10.1016/j.jhydrol.2004.03.028

Lei, R., Leppäranta, M., Cheng, B., Heil, P., & Li, Z. (2012). Changes in ice-season characteristics of a European Arctic lake from 1964 to 2008. *Climatic Change*, *115*(3-4), 725–739. doi:10.1007/s10584-012-0489-2

Leshkevich, G. A., Schwab, D. J., & Muhr, G. C. (1993). Satellite environmental monitoring of the great-lakes - a review of NOAA's great-lakes CoastWatch program. *Photogrammetric Engineering and Remote Sensing*, *59*(3), 371–379.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766

Livingstone, D. M. (1997). Break-up dates of alpine lakes as proxy data for local and regional mean surface air temperatures. *Climatic Change*, *37*(2), 407–439. doi:10.1023/A:1005371925924

Livingstone, D. M. (2000). Large-scale climatic forcing detected in historical observations of lake ice break-up. In W. D. Williams (Ed.), *Verhandlungen der Internationalen Vereinigung für Theoretische und Angewandte Limnologie* (Vol. 27(5), pp. 2775–2783). Stuttgart, Germany.

Livingstone, D. M., & Adrian, R. (2009). Modeling the duration of intermittent ice cover on a lake for climate-change studies. *Limnology and Oceanography*, *54*(5), 1709–1722. doi:10.4319/lo.2009.54.5.1709

Livingstone, D. M., Adrian, R., Blenckner, T., George, G., & Weyhenmeyer, G. A. (2010). Lake Ice Phenology. In G. George (Ed.), *The Impact of Climate Change on European Lakes* (pp. 51–61). Springer Netherlands.

LP DAAC. (2011). *MODIS Reprojection Tool User's Manual* (No. 4.1). NASA Land Processes Distributed Active Archive Center. Retrieved from https://lpdaac.usgs.gov/tools/modis_reprojection_tool

Magnuson, J. J., Robertson, D. M., Benson, B. J., Wynne, R. H., Livingstone, D. M., Arai, T., … Vuglinsky, V. S. (2000). Historical Trends in Lake and River Ice Cover in the Northern Hemisphere. *Science*, *289*(5485), 1743–1746.

Maslanik, J. A., & Barry, R. G. (1987). Lake ice formation and breakup as an indicator of climate change: Potential for monitoring using remote sensing techniques. In *The Influence of Climate Change and Climate Variability on the Hydrologic Regime and Water Resources* (Vol. 168, pp. 153–161). Presented at the Vancouver Symposium 1987, Vancouver, Canada: IAHS Press.

Menard, P., Duguay, C. R., Flato, G. M., & Rouse, W. R. (2002). Simulation of ice phenology on Great Slave Lake, Northwest

Territories, Canada. *Hydrological Processes*, *16*, 3691–3706. doi:10.1002/Hyp.1230

Mishra, V., Cherkauer, K. A., Bowling, L. C., & Huber, M. (2011b). Lake Ice phenology of small lakes: Impacts of climate variability in the Great Lakes region. *Global and Planetary Change*, *76*(3–4), 166–185. doi:10.1016/j.gloplacha.2011.01.004

NSIDC. (2012). Global Lake and River Ice Phenology - Documentation. Retrieved September 4, 2013, from http://nsidc.org/data/docs/noaa/g01377_lake_river_ice/#doc_info

Palecki, M. A., & Barry, R. G. (1986). Freeze-up and Break-up of Lakes as an Index of Temperature Changes during the Transition Seasons: A Case Study for Finland. *Journal of Climate and Applied Meteorology*, *25*(7), 893–902. doi:10.1175/1520-0450(1986)025<0893:FUABUO>2.0.CO;2

Pienitz, R., Doran, P. T., & Lamoureux, S. F. (2008). Origin and geomorphology of lakes in the polar regions. In *Polar lakes and rivers: limnology of Arctic and Antarctic aquatic ecosystems* (pp. 25–42). Oxford University Press.

Prowse, T., Alfredsen, K., Beltaos, S., Bonsal, B., Duguay, C., Korhola, A., … Weyhenmeyer, G. A. (2011a). Arctic Freshwater Ice and Its Climatic Role. *AMBIO*, *40*(1), 46–52. doi:10.1007/s13280-011-0214-9

Prowse, T., Alfredsen, K., Beltaos, S., Bonsal, B., Duguay, C., Korhola, A., … Weyhenmeyer, G. A. (2011b). Past and Future Changes in Arctic Lake and River Ice. *AMBIO*, *40*(S1), 53–62. doi:10.1007/s13280-011-0216-7

Robertson, D. M., Ragotzkie, R. A., & Magnuson, J. J. (1992). Lake ice records used to detect historical and future climatic changes. *Climatic Change*, *21*(4), 407–427. doi:10.1007/BF00141379

Rouse, W. R., Blanken, P. D., Duguay, C. R., Oswald, C. J., & Schertzer, W. M. (2008). Climate-Lake Interactions. In M. Woo (Ed.), *Cold Region Atmospheric and Hydrologic Studies. The Mackenzie GEWEX Experience* (pp. 139–160). Springer Berlin Heidelberg.

Saloranta, T. M., & Andersen, T. (2007). MyLake—A multi-year lake simulation model code suitable for uncertainty and sensitivity analysis simulations. *Ecological Modelling*, *207*(1), 45–60. doi:10.1016/j.ecolmodel.2007.03.018

Schwartz, M. D. (2003). *Phenology: An Integrative Environmental Science*. Springer.

Semmler, T., Cheng, B., Yang, Y., & Rontu, L. (2012). Snow and ice on Bear Lake (Alaska) – sensitivity experiments with two lake ice models. *Tellus A*, *64*(0). doi:10.3402/tellusa.v64i0.17339

Sen, P. K. (1968). Estimates of the Regression Coefficient Based on Kendall's Tau. *Journal of the American Statistical Association*, *63*(324), 1379. doi:10.2307/2285891

Sharma, S., Magnuson, J. J., Mendoza, G., & Carpenter, S. R. (2013). Influences of local weather, large-scale climatic drivers, and the ca. 11 year solar cycle on lake ice breakup dates; 1905–2004. *Climatic Change*, *118*(3-4), 857–870. doi:10.1007/s10584-012-0670-7

Simpson, J. A. (Ed.). (2002). *Oxford English Dictionary: Version 3.0 : Upgrade Version*. Oxford University Press, Incorporated.

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

Stepanenko, V. M., Goyette, S., Martynov, A., Perroud, M., Fang, X., & Mironov, D. (2010). First steps of a Lake Model Intercomparison Project LakeMIP. *Boreal environment research*, *15*(2).

Surdu, C. M., Duguay, C. R., Brown, L. C., & Fernández Prieto, D. (2014). Response of ice cover on shallow lakes of the North Slope of Alaska to contemporary climate conditions (1950–2011): radar remote-sensing and numerical modeling data analysis. *The Cryosphere*, *8*(1), 167–180. doi:10.5194/tc-8-167-2014

Tolk, B. L., Han, L., & Rundquist, D. C. (2000). The impact of bottom brightness on spectral reflectance of suspended sediments. *International Journal of Remote Sensing*, *21*(11), 2259–2268. doi:10.1080/01431160050029558

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

Vermote, E. F., Kotchenova, S. Y., & Ray, J. P. (2011). MODIS Surface Reflectance User's Guide. MODIS Land Surface Reflectance Science Computing Facility. Retrieved from http://modis-sr.ltdri.org/products.html

Vincent, W. F., Callaghan, T. V., Dahl-Jensen, D., Johansson, M., Kovacs, K. M., Michel, C., … Sharp, M. (2011). Ecological Implications of Changes in the Arctic Cryosphere. *AMBIO*, *40*(1), 87–99. doi:10.1007/s13280-011-0218-5

Vincent, W. F., & Laybourn-Parry, J. (2008). *Polar lakes and rivers: limnology of Arctic and Antarctic aquatic ecosystems*. Oxford University Press.

Walsh, J. E. (2005). Cryosphere and Hydrology. In *Arctic Climate Impact Assessment - Scientific Report* (pp. 183–242).

Cambridge University Press. Retrieved from http://www.acia.uaf.edu/pages/scientific.html

Walsh, S. E., Vavrus, S. J., Foley, J. A., Fisher, V. A., Wynne, R. H., & Lenters, J. D. (1998). Global patterns of lake ice phenology and climate: Model simulations and observations. *Journal of Geophysical Research: Atmospheres*, *103*(D22), 28825–28837. doi:10.1029/98JD02275

Walter, K. M., Smith, L. C., & Chapin, F. S. (2007). Methane bubbling from northern lakes: present and future contributions to the global methane budget. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *365*(1856), 1657–1676. doi:10.1098/rsta.2007.2036

Walter, K. M., Zimov, S. A., Chanton, J. P., Verbyla, D., & Chapin, F. S. (2006). Methane bubbling from Siberian thaw lakes as a positive feedback to climate warming. *Nature*, *443*(7107), 71–75. doi:10.1038/nature05040

Wang, J., Bai, X., Hu, H., Clites, A., Colton, M., & Lofgren, B. (2012). Temporal and Spatial Variability of Great Lakes Ice Cover, 1973–2010*. *Journal of Climate*, *25*(4), 1318–1329. doi:10.1175/2011JCLI4066.1

Weyhenmeyer, G. A., Livingstone, D. M., Meili, M., Jensen, O. P., Benson, B., & Magnuson, J. J. (2011). Large geographical differences in the sensitivity of ice-covered lakes and rivers in the Northern Hemisphere to temperature changes. *Global Change Biology*, *17*(1), 268–275. doi:10.1111/j.1365-2486.2010.02249.x

Weyhenmeyer, G. A., Meili, M., & Livingstone, D. M. (2004). Nonlinear temperature response of lake ice breakup. *Geophysical Research Letters*, *31*(7), n/a–n/a. doi:10.1029/2004GL019530

Weyhenmeyer, G. A., Meili, M., & Livingstone, D. M. (2005). Systematic differences in the trend toward earlier ice-out on Swedish lakes along a latitudinal temperature gradient. In *Verhandlungen der Internationalen Vereinigung für Theoretische und Angewandte Limnologie* (Vol. 29(1), pp. 257–260). Stuttgart, Germany.

Woo, P. D. M. (2012). Cold Lakes. In *Permafrost Hydrology* (pp. 305–346). Springer Berlin Heidelberg.

Wynne, R. H., & Lillesand, T. M. (1993). Satellite observation of lake ice as a climate indicator: initial results from statewide monitoring in Wisconsin. *Photogrammetric Engineering and Remote Sensing*, *59*.

Wynne, R. H., Lillesand, T. M., Clayton, M. K., & Magnuson, J. J. (1998). Satellite monitoring of lake ice breakup on the Laurentian Shield (1980-1994). *G*, *Photogrammetric Engineering and Remote Sensing. 64*(6), 607–617.

Yoo, J. C., & D'Odorico, P. (2002). Trends and fluctuations in the dates of ice break-up of lakes and rivers in Northern Europe: the effect of the North Atlantic Oscillation. *Journal of Hydrology*, *268*(1–4), 100 – 112. doi:10.1016/S0022-1694(02)00161-0

Zhang, X., Vincent, L. A., Hogg, W. D., & Niitsoo, A. (2000). Temperature and precipitation trends in Canada during the 20th century. *Atmosphere-Ocean*, *38*(3), 395–429. doi:10.1080/07055900.2000.9649654

Appendix

Appendix A: Download and pre-processing of MODIS data (R)

```
#### DOWNLOAD, MOSAIC AND REPROJECT MODIS DATA
# title          : MODIS_DRM.R
# author         : T. Smejkalova
# reference      : adapted from T. Hengl
#                  (http://spatial-analyst.net/wiki/
#                  index.php?title=Download_and_resampling
#                  _of_MODIS_images)
#                  - see for mosaic and reprojection code settings
# last update    : Southampton, UK, 18 April 2014.
# inputs :
#   ASCII files containing link to each tile have to be
#   downloaded from: http://reverb.echo.nasa.gov/reverb/
#       - for each year separately
#       - recommended name 'year_data.txt'
#       - must contain all tiles to mosaic
#       - pixel size to resample to
################################################################
# location of the Modis Reprojection Tool:
MRT <- 'D:\\MRT\\bin\\'
# location of the working directory containing:
  # text files for each year
  # aria2c.exe download utility
workd <- 'D:\\R_Modis\\'
setwd(workd)
# perpetual years
year <- c(2001,2002,2003,2005,2006,2007,2009,2010,2011,2013)
# leap years
# year <- c(2000,2004,2008,2012)
######################### year ################################
for (y in 1:length(year)){
  shell(cmd = paste('cd ',workd,sep=""))
  shell(cmd=paste('aria2c -i ', year[y], '_data.txt', sep=""))
  # data processed only for 8th Feb. to 31st Nov.
  s = year[y]*1000+39
  e = year[y]*1000+334 # 335 for leap years
  for (i in s:e)
    {
    # set pattern corresponding to input data product and
    tile1 <- list.files(getwd(), pattern=paste
                ('MOD09GA.A',i,'.h12v01.005.*.hdf', sep=''))
    tile2 <- list.files(getwd(), pattern=paste
                ('MOD09GA.A',i,'.h12v02.005.*.hdf', sep=''))
    # tile3 <- list.files(getwd(), pattern=paste
    #           ('MOD09GA.A',i,'.h22v02.005.*.hdf', sep=''))
    # Mosaic tiles:
    mosaicname = file(paste(MRT, "TmpMosaic.prm", sep=""),
                  open="wt")
    write(paste(workd, tile1, sep=""), mosaicname)
    write(paste(workd, tile2, sep=""), mosaicname, append=T)
    # write(paste(workd, tile3, sep=""), mosaicname, append=T)
    close(mosaicname)

    # generate temporary mosaic (see MRT manual)
```

```
    # "0 0 1 0 0 " processes only third and out of 5 bands
    # - if input is MOD09GQ product - NIR surface reflectance
    # "0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0" process band 2
    # - if input MOD09GA - state band
    shell(cmd = paste('cd ',MRT,sep=""))
    shell(cmd=paste(MRT, 'mrtmosaic -i ', MRT, 'TmpMosaic.prm -s
      "0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0" -o ',
      workd, 'TmpMosaic.hdf', sep=""))
      # resample to Polar stereographic projection:
      filename = file(paste(MRT, "mrt", i, ".prm", sep=""), open="wt")
      write(paste('INPUT_FILENAME = ', workd, 'TmpMosaic.hdf', sep=""),
            filename)
      write('  ', filename, append=TRUE)
      write('SPECTRAL_SUBSET = (1)', filename, append=TRUE)
      write('  ', filename, append=TRUE)
      write('SPATIAL_SUBSET_TYPE = OUTPUT_PROJ_COORDS',
            filename, append=TRUE)
      write('  ', filename, append=TRUE)
      # subset coordinates in output projection
      write('SPATIAL_SUBSET_UL_CORNER = ( -1150000.0 2150000.0 )',
            filename, append=TRUE)
      write('SPATIAL_SUBSET_LR_CORNER = ( -700000.0 1850000.0 )',
            filename, append=TRUE)
      write('  ', filename, append=TRUE)
      write(paste('OUTPUT_FILENAME = ', workd, i, '.tif', sep=""),
            filename, append=TRUE)
      write('  ', filename, append=TRUE)
      write('RESAMPLING_TYPE = NEAREST_NEIGHBOR',
            filename, append=TRUE)
      write('  ', filename, append=TRUE)
      write('OUTPUT_PROJECTION_TYPE = PS', filename, append=TRUE)
      write('  ', filename, append=TRUE)
      write('OUTPUT_PROJECTION_PARAMETERS = (', filename, append=TRUE)
      # see available projections and projection parameters in MRT Manual
      write(' 0.0 0.0 0.0', filename, append=TRUE)
      write(' 0.0 0.0 71.0', filename, append=TRUE)
      write(' 0.0 0.0 0.0', filename, append=TRUE)
      write(' 0.0 0.0 0.0', filename, append=TRUE)
      write(' 0.0 0.0 0.0 )', filename, append=TRUE)
      write('  ', filename, append=TRUE)
      write('DATUM = WGS84', filename, append=TRUE)
      write('  ', filename, append=TRUE)
      write('OUTPUT_PIXEL_SIZE = 250', filename, append=TRUE)
      write('  ', filename, append=TRUE)
      close(filename)
      # Mosaic the images to get the whole area
      shell(cmd=paste('cd ',MRT, sep=''))
      shell(cmd=paste(MRT, 'resample -p ', MRT, 'mrt', i, '.prm',
sep=""))
      # delete all HDF files!
      unlink(paste(getwd(), '/', 'TmpMosaic.hdf', sep=''))
      unlink(paste(MRT, 'mrt', i,'.prm',sep=""))
      unlink(paste(getwd(), '/', tile1, sep=''))
      unlink(paste(getwd(), '/', tile2, sep=''))
      # unlink(paste(getwd(), '/', tile3, sep=''))
    }
  }

# End of Script!
```

Appendix B: Reclassify to cloud mask (Python)

```
#### Reclassify raster by ASCII (standalone)
# Name: recode_py.py
# Author: ESRI (adapted by T. Smejkalova)
# Original code accessible from:
#       http://help.arcgis.com/en/arcgisdesktop
#       /10.0/help/index.html#//009z000000sp000000.htm
# Description: Reclassifies the values in a
# all rasters in folder based on ASCII table.
# Requirements: Spatial Analyst Extension

# Import system modules
import arcgisscripting
import arcpy
from arcpy import env
from arcpy.sa import *
import os

# arcpy.env.overwriteOutput = True

# Create the Geoprocessor object
gp = arcgisscripting.create()

# Set environment settings
env.workspace = "D:/DATA"
count = 0

# Get list of rasters from workspace
rasterList = arcpy.ListRasters("*", "All")

# sort alphabetically
rasterList.sort()
# table to recode by in ASCII format
inASCIIfile = "D:/DATA/recode.txt"

for inRaster in rasterList:

    outRaster = "D:/DATA/" + inRaster[: -4] + "_recl.tif"
    # Check out the ArcGIS Spatial Analyst extension license
    arcpy.CheckOutExtension("Spatial")
    # Process: ReclassByASCIIFile
    gp.ReclassByASCIIFile_sa(inRaster, inASCIIfile, outRaster, "NODATA")
```

Appendix C: Surface reflectance time series extraction (R)

```
## EXTRACT VALUES FOR LAKE PIXELS
# title          : Extract_values.R
# author         : T. Smejkalova
# last update    : Southampton, UK, 22 April 2014.
# Description    :
  # - Uses the raster package to stack reflectance
  #   tiffs for one year into one raster object,
  # - The same is done for the cloud mask tiffs.
  # - The two then mask operation is applied where
  #   each reflectance layer is masked by corresponding
  #   cloud mask layer (pixels with value 0 are masked
  #   out in the reflectance layer)
```

```
  # - Pixel values are extracted at provided coordinates in
  #   Polar stereographic projection (lat0 = 71N)
  # - Values written in a .csv file
# inputs       :
  # CSV file containing point ID and X and Y coordinates
  # CSV file containing day of study, day of year and date
# date :
t2000 <- 1:366
t2001 <- 367:731
t2002 <- 732:1096
t2003 <- 1097:1461
t2004 <- 1462:1827
t2005 <- 1828:2192
t2006 <- 2193:2557
t2007 <- 2558:2922
t2008 <- 2923:3288
t2009 <- 3289:3653
t2010 <- 3654:4018
t2011 <- 4019:4383
t2012 <- 4384:4749
t2013 <- 4750:5114

############################################################
library(raster)
library(plyr)
year <- c(2000,2001,2002,2003,2004,2005)
location <- 'WS'
workd <- paste('C:\\data_',year[1], sep = '')
setwd(workd)
total<- read.csv('Date.csv')

# Create point coordinate file
pointCoordinates=read.csv('lakepoints_WS.csv')
coordinates(pointCoordinates)= ~ X+ Y

for (i in 1:length(year)) {
  # Working directory
  workd <- paste('C:\\data_', year[i], '\\', sep='')
  setwd(workd)

  ##########################################################
  ## STACKING AND MASKING
  # list of rasters to stack
  files <- list.files(pattern='sur_refl_b02_1.tif')
  mask <- list.files(pattern='state_1km_1_recl.tif')
  # Visual check if layers correspond
  test<-data.frame(files=substr(files[1:length(files)],1,10),
                   mask=substr(mask[1:length(mask)],1,10))

  # Stacking layers
  rasStack_refl = stack(files)
  rasStack_mask = stack(mask)

  # Cloud Mask
  Refl_stack <- mask(rasStack_refl, rasStack_mask, maskvalue = 0 )
  # remove all unnecessary data from workspace
  rm(files,mask,rasStack_refl,rasStack_mask)

  #########################################################
  ## VALUE EXTRACTION
  # Extract values based on coordinates
```

```
  X=extract(Refl_stack, pointCoordinates)

  ##########################################################
  ## ASSIGN NAMES TO CLUMNS AND WRITE CSV
  # Assign names to columns and add columns for NoData days
    coln <- colnames(X)
    doy <- substr(coln[1:length(coln)],6,8)
    # colname<- gsub('_','-',colnam)
    # doy<-strptime(colname, '%F')$yday+1
    colnames(X) <- doy

  # Setting column names to Day of study
  d <- get(paste('t', year[i], sep=''))
  dat<-date[d, 1]
  namevector <-as.character(dat)
  Y <- data.frame(matrix(data=NA, nrow = 0, ncol=length(namevector)))
  colnames(Y) <-namevector
  Y <- rbind.fill.matrix(X,Y)
  Y<-Y[, c(namevector)]
  dos<- date[d, 2]
  colnames(Y)<-dos

  csvname <- paste(year[i],'_', location, '.csv', sep='')
  values<- data.frame(pointCoordinates, Y)
  write.table(values,file=csvname, append=F, sep= ",",
              row.names = FALSE, col.names=TRUE, na = 'NaN')
  rm(X, coln, doy, dat, namevector, Y, dos, csvname, values)
}

# End of script
```

Appendix D: Outlier removal (MATLAB function)

```
function [Y] = OutREM(I, k, NoData) % Version 1.5
% OutREM removes outliers from time series data with missing values.
% Author : T.Smejkalova
% Last updated: 22. February, 2014
% Syntax:
%     [Y] = OutREM(I,k);
%     [Y] = OutREM(I,k,NoData);
% Input:
%     I    -  Name of the csv file to process as string - without '.csv'
%             extension.
%     k    -  Window semi-length. A positive scalar. Window length = 2k+1
%     NoData - residual NoData value from original imagery to remove
%                (Optional). If not set then Default = -28672.
% Output:
%     Y     - Matrix with removed outliers
%     .csv file - containing Y and original coordinates and Point and
%                 Lake IDs
% Description:
%     within a sliding window of user defined length 2k+1, identifies
%     values
%     that are further than 3 times the value of Median Absolute
%     Deviation (MAD) from the Median, calculated, and replaces them with
%     NaNs.
% Uses subfunction OutREM_Static(x)
%     OutREM_Static replaces outlying values with NaN.
%     Function calculates median and Median absolute deviation within a
%     window and checks whether central value falls within
```

```
%     Median +- 3*MAD interval if it does, it returns original value, if
%     not and the count of non-missing values in the window is larger
%     than 4, it returns NaN.
% Example:  csv. file to process is 'test.csv'
%           k = 5 ...length of the window will be 11
%           NoData = -28672 (Default)
%           Y = OutREM_MAD('test', 5);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Error checking 1
clc;
if ~nargin
 error('OutREM:Inputs','There are no inputs.')
elseif nargin<2 || isempty(k)
 error('OutREM:Inputs', 'The semi-length k of moving window must be set')
end
k = round(k);
if k==0
  error('OutREM:Inputs', 'k must be larger than 0')
end
if nargin<3 || isempty(NoData)
    NoData = -28672;
    fprintf('NoData vaue equals -28672. \n')
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Load user defined data in .csv format
IN = strcat(I, '.csv');
%line = input('first line to import (1)');
%col = input ('first column to import (4)');
X = csvread(IN,1,4);
% remove residual noData values and replace with NaN
X(X == NoData) = NaN;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Error checking 2
ndim = ndims(X);
if (ndim ~= 2)
 error('OutREM:Inputs','Input .csv does not contain a vector or matrix.')
end
N = size(X,2);
if 2*k+1>N
 warning('OutREM:Inputs',...
 'Window size must be less or equal as the number of elements in a row.')
End

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Apply filter in a 1xn window (row-wise)
%  - matrix needs to be padded with k NaN values before and after the
last
%    column
X = padarray(X, [0, k], NaN);
n = (2*k)+ 1;
Y = nlfilter(X, [1,n], @OutREM_Static);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Combine with coordinate values, Lake_IDs a and Pixel_IDs
r = size(X,1);
c = size (X,2);
Y = Y(:, (k+1):(c-k));
C = csvread(IN,1,0, [1 0 r 3]);
V = [C Y];
```

```
%% Write new .csv file
filename = strcat(I,'_OR', '.csv');
dlmwrite(filename, V, 'precision', 12);
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% OutREM_Static subfunction
function [y] = OutREM_Static(x)
    %% Calculate median and MAD of a window ignoring NaNs
    Median = nanmedian(x, 2);
    MAD = mad(x, 1, 2);
    MADp = (1.4826)*MAD; % constant for normal distribution

    %% Calculate number of non-missing values within a window
    NM = sum(~isnan(x),2);

    %% Identify the central value
    n = size(x, 2);
    l= ((n-1)/2)+1;

    %% Test if central value is an outlier and replace outliers with NaN
        if isnan(x(:,l)) == 1; % if central value NaN, do not interpolate
            y = NaN;
        % Test if value is an outlier
        elseif x(:,l) < (Median-(3*MADp)) && (NM > 4);
            y = NaN;
        elseif x(:,l) > (Median+(3*MADp)) && (NM > 4);
            y = NaN;
        else   % if yes then keep original value
            y = x(:,l);
        end
 end
```

Appendix E: Identify and remove mixed pixel profiles (R)

```
## IDENTIFY AND REMOVE MIXED LAND-WATER PIXELS, AVERAGE VALUES PER LAKE
# calculates the mean value of all pixels over summer period (July,
# August) when it is expected that the lakes will have no ice cover.
# Author: T.Smejkalova

# INPUTS:         MUST BE PLACED IN WORKING DIRECTORY!!!
# - OR.csv files
# - lakepoint.csv file for correct location
#   (contains Lake_ID, Point_ID, X, Y)
# - headers.csv file - headers for each year in format:
#   Year,'Lake_ID','Point_ID','X','Y', DOS1 for year...DOS365/6 for year

####################################################################
rm(list=ls(all=TRUE))
setwd('C:\\DATA\\')
Lake_ID <- read.csv('lakepoints_Y.csv')
headers <-read.csv('headers.csv', header = F)
Point_ID = Lake_ID[ ,2]
Lake_ID = Lake_ID[ , 1]
means = data.frame(Point_ID)

Year <- c(2000, 2001, 2002)
location <- 'Y'
```

```r
TR <- 390.47

###################################################################
## Calculate means for 62 days of summer for 5 years for all pixels
for (i in 1:5) {
  IN <- paste(Year[i], '_', location, '_OR.csv', sep='')
  V <-read.csv(IN, header = F)
  s = V[, (182+4):(243+4)]
  means[,(i+1)] = rowMeans(s, na.rm=T)
}
rm (IN, V, s)

###################################################################
## Identify mixed pixels as those that have mean summer value higher
# than threshold for 3 out of 5 years
M = data.frame(Lake_ID)
M$Point_ID = Point_ID
for (i in 1:5) {
  MM <- means[,(i+1)]
  MM[MM<= TR] <- 1
  MM[MM> TR ] <- 0
  M[,(i+2)] <- MM
  rm (MM)
}
M$Score = rowSums(M[,3:7])
M$Score[is.na(M$Score)==1] <- 0
Rem <- subset(M[,1:2], M$Score<3)
Keep <- subset(M[,1:2], M$Score>=3)
OUT <- paste(location, '_Mixed_Pixel_ID.csv', sep='')
write.table(Rem, file = OUT, append=F, sep= ",",
            row.names = FALSE, col.names=TRUE)
OUT <- paste(location, '_Water_Pixel_ID.csv', sep='')
write.table(Keep, file = OUT,append=F, sep= ",",
            row.names = FALSE, col.names=TRUE)
rm(TR, OUT, i, means, M)

##################################################################
## REMOVE MIXED LAND-WATER PIXELS AND AVERAGE VALUES PER LAKE
## Remove mixed pixels

for (i in 1:length(Year)){
  y <- Year[i]
  IN <- paste(y, '_',location,'_OR.csv', sep = '')
  V <-read.csv(IN, header=F)
  n <- ncol(V)
  c <- n+1
  h <- subset(headers[,2:c], headers$V1%in%y)
  h <- sapply(h, as.character)
  colnames(V) <- h
  W <- subset(V, Point_ID%in%Keep$Point_ID)

  ##################################
  ## Average values per day per lake
  L <- aggregate(W[, 5:n], by = list(Lake_ID = W$Lake_ID),
                 FUN = 'mean', na.rm = 'TRUE')
  n <- ncol(L)
  L <- L[,2:n]
  pixel_count <-aggregate(W[, 3], by = list(Lake_ID = W$Lake_ID),
                          FUN = 'length')
  Lake <- cbind(pixel_count, L)
  OUT <- paste(y,'_', location, '_L.csv', sep = '')
```

```
  write.table(Lake,file= OUT, append=F, sep= ",", row.names = FALSE,
              col.names=TRUE, na="NaN")
}
 # End of Script
```

Appendix F: Dark pixel value removal (MATLAB function)

```
function [ Y ] = DarkPixRem( X, Leap)
%DarkPixRem removes lake values affected by polar darkness
%   compares values in affected period with maximum value in 7 days
%   preceeding the start of the period. The 7 day maximum is the first
%   choice if it is available (if all values are not missing) if they are
%   the threshold is set to 390
% Author: T. Smejkalova
%
%   Syntax:
%     [Y] = DarkPixRem(X,Leap);
%
%   Input
%     X    -  timeseries for one year
%     Leap - choose if input data for perpetual (0) or leap (1) year

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Error checking 1
clc;
format long g
if ~nargin
  error('DarkPixRem:Inputs','There are no inputs.')
elseif nargin<2 || isempty(Leap)
  error('DarkPixRem:Inputs', 'Perpetual or leap year must be set')
end
ndim = ndims(X);
if (ndim ~= 2)
 error('DarkPixRem:Inputs','Input is not a vector or matrix.')
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Perpetual or leap year
if Leap == 0
    n = 365;
elseif Leap == 1
    n = 366;
else
    error('DarkPixRem:Inputs','Leap must be either 0 for perpetual \n'...
    'or 1 for leap year')
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Error checking 2
N = size(X,2);
if (Leap == 0 && N~=365)
 error('DarkPixRem:Inputs',...
  'Input data must have lenght or 365 for perpetual or 366 for leap
\n'...
  'year along the 2 dim.')
elseif (Leap == 1 && N~=366)
    error('DarkPixRem:Inputs',...
```

```
  'Input data must have lenght or 365 for perpetual or 366 for leap
\n'...
  'year along the 2 dim.')
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Calculate maximum for the 5 days before dark period starts (Europe)
ce = n-58;
cs = n-60; %5 val
M = max(X(:,cs:ce),[], 2);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Replace affected values with NaN
r = size (X, 1);
ds = n-58;
Y = X;
for i = 1:r
    if isnan(M(i,1))==0
           m = M(i,1);
        elseif isnan(M(i,1))==1
           m = 390;
    end
    for j = ds:n
        y = Y(i,j);
        if y < m
           Y(i,j)= NaN;
        else
           Y(i,j)= Y(i,j);
        end
    end
end

end
```

Appendix G: TIMESAT input preparation (MATLAB)

```
%% TIMESAT DATA PREPARATION
% preparation of the specific input file for TIMESAT
% Input: .mat file containing timeseries in rows(Y), days of study (X)
%        list of lake IDs (Lakes_ID)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
clc
clear all
location = 'Y';
IN = strcat(location, '_data.mat');
load(IN)
V = Y;       % load values
% for Timesat all years must have equal number of values
%   - therefore last day of each leap year is removed
l = [366 1827 3288 4749];
V(:,l)= [];

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Creating two 7 year parts for processing in Fortran
a = 7* 365;
V1 = V(:,1:a);
V2 = V(:, (a+1): 5110);
clear a
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Creating dummy year before and after the original data
% Timesat only extracts values for n-1 centermost years
V1 = padarray(V1, [0 365], 'symmetric', 'pre');
V1 = padarray(V1, [0 365], 'symmetric', 'post');
r1 = size(V1,1);
c1 = size(V1,2);
I1 = nan(r1,c1);
X1 = 1:1:c1;

V2 = padarray(V2, [0 365], 'symmetric', 'pre');
V2 = padarray(V2, [0 365], 'symmetric', 'post');
r2 = size(V2,1);
c2 = size(V2,2);
I2 = nan(r2,c2);
X2 = 1:1:c2;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Linear interpolation of the missing data
for i=1:r1
    y = V1(i,:);
    Sy = y(isnan(y)==0);
    Sx = X1(isnan(y)== 0);
    I1(i,:) = interp1(Sx, Sy, X1, 'linear');
end

for i=1:r2
    y = V2(i,:);
    Sy = y(isnan(y)==0);
    Sx = X2(isnan(y)== 0);
    I2(i,:) = interp1(Sx, Sy, X2, 'linear');
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Create input file for Timesat
T1 = I1;
c1 = size(T1,2);
n1 = c1/365;
if rem(n1,1)==0;
    nyr = n1;
else
    disp('You chose to cancel. Quit the program now?');
    disp('<a href="MATLAB: dbquit;">Yes</a> / \n' ...
        '<a href="MATLAB: dbcont;">No</a>');
    keyboard;
end
T1(isnan(T1)==1)= 8000;
header = [nyr 365 r1];
OUT1 = strcat('data_', location,'1.txt');
dlmwrite(OUT1, header, 'delimiter','\t');
dlmwrite(OUT1, T1, 'delimiter', '\t', '-append')

T2 = I2;
c2 = size(T2,2);
n2 = c2/365;
if rem(n2,1)==0;
    nyr = n2;
else
    disp('You chose to cancel. Quit the program now?');
    disp('<a href="MATLAB: dbquit;">Yes</a> / \n' ...
```

```
        '<a href="MATLAB: dbcont;">No</a>');
    keyboard;
end
T2(isnan(T2)==1)= 8000;
header = [nyr 365 r2];
OUT2 = strcat('data_', location,'2.txt');
dlmwrite(OUT2, header, 'delimiter','\t');
dlmwrite(OUT2, T2, 'delimiter', '\t', '-append')

%% End of Script
```

Appendix H: Settings file for TIMESAT event extraction

Relevant settings are highlighted in red, other setting are either not applicable (for data or used method) or determined automatically.

- Extraction of BUE and FUS events

```
BUE_FUS_1       %Job_name (no blanks)
0                %Image /series mode (1/0)
0                %Trend (1/0)
0                %Use mask data (1/0)
D:\Timesat\timesat311\run\data1.txt  %Data file list/name
dummy           %Mask file list/name
1                %Image file type
0                %Byte order (1/0)
1 1             %File dimension (nrow ncol)
1 1809 1 1      %Processing window (start row stop row start col
                       stop col)
9 365            %No. years and no. points per year
0 10000          %Valid data range (lower upper)
-1e+06 1e+06 1          %Mask range 1 and weight
-1e+06 1e+06 1          %Mask range 2 and weight
-1e+06 1e+06 1          %Mask range 3 and weight
0                %Amplitude cutoff value
0                %Print functions and weights (1/0)
1 1 0            %Output files (1/0 1/0 1/0)
0                %Use land cover (1/0)
                 %Name of landcover file
2                %Spike method
30               %Spike value
1                %No. of landcover classes
************
1                %Land cover code for class  1
1                %Season parameter
3                %No. of envelope iterations (1-3)
0.1              %Adaptation strength (1-10)
0 0              %Force minimum (1/0) and value
3                %Fitting method (1-3)
1                %Weight update method
4                %Window size for Sav-Gol.
1                %Season start method
0.05 0.12        %Season start / stop values
```

- Extraction of BUE and FUE events

```
FUE_BUS_1        %Job_name (no blanks)
```

```
0                %Image /series mode (1/0)
0                %Trend (1/0)
0                %Use mask data (1/0)
D:\Timesat\timesat311\run\data_1.txt  %Data file list/name
dummy            %Mask file list/name
1                %Image file type
0                %Byte order (1/0)
1 1              %File dimension (nrow ncol)
1 1809 1 1       %Processing window (start row stop row start col
                        stop col)
9 365            %No. years and no. points per year
0 10000          %Valid data range (lower upper)
-1e+06 1e+06 1         %Mask range 1 and weight
-1e+06 1e+06 1          %Mask range 2 and weight
-1e+06 1e+06 1          %Mask range 3 and weight
0                %Amplitude cutoff value
0                %Print functions and weights (1/0)
1 1 0            %Output files (1/0 1/0 1/0)
0                %Use land cover (1/0)
                 %Name of landcover file
2                %Spike method
30               %Spike value
1                %No. of landcover classes
************
1                %Land cover code for class 1
1                %Season parameter
3                %No. of envelope iterations (1-3)
10               %Adaptation strength (1-10)
1 0              %Force minimum (1/0) and value
3                %Fitting method (1-3)
1                %Weight update method
4                %Window size for Sav-Gol.
1                %Season start method
0.4 0.8          %Season start / stop values
```

Appendix I: TIMESAT output seasonality processing (MATLAB)

- Example code for break-up events (options for freeze-up commented)

```
%% PROCESS THE OUTPUT SEASONALITY FILE FROM TIMESAT
% name: TS_output_process.m
% Read Timesat output seasonality data part 1
clc
clear all
E = 'BUS'; % BUE, FUE, FUE
F= 'FUE_BUS_'; % BUE_FUS
location = 'Y';
Part = '1';
IN = strcat(F, Part, '_', location,'_seasonality_last.csv');
V1 = csvread(IN,1,0);
% get rid of unnecessary columns
d = [2 4 6:14]; % BUE, BUS
% d = [2 5:14]; % FUE, FUS
V1(:,d) = [];
V1(:,4) = round(V1(:,3)-365);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Read Timesat output seasonality data part 2
Part = '2';
```

```
IN = strcat(F, Part, '_', location,'_seasonality_last.csv');
V2 = csvread(IN,1,0);
% get rid of unnecessary columns
d = [2 4 6:14]; % BUE, BUS
% d = [2 5:14]; % FUE, FUS
V2(:,d) = [];
V2(:,4) = round(V2(:,3)-365 +(7*365));
clear d IN Part
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Delete dummy season 8 - BUE, BUE
V1(V1(:,2)==8,:)=[];
V2(V2(:,2)==8,:)=[];
V2(:,2)=V2(:,2)+7;
V = [V1; V2];
V = sortrows(V, [1 2]);

% %% Delete dummy season 1 - FUS, FUE
% V1(V1(:,2)==1,:)=[];
% V2(V2(:,2)==1,:)=[];
% V2(:,2)=V2(:,2)+7;
% V = [V1; V2];
% V = sortrows(V, [1 2]);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Read file containing lake IDs and row number
V(V(:,3)==0, :) = [];
IN = strcat('ID_',location,'.csv');
ID = csvread(IN, 1,0);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Remove incomplete time series
count = unique(V(:,1));
count(:,2) = histc(V(:,1),count);
rem = count(count(:,2)<11,:);
for i = 1:size(rem, 1)
    r = rem(i,1);
    V(V(:,1)==r,:)=[];
end
clear i r
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Fit Lake_ID to each record
[~,loc] = ismember(V(:,1),ID(:,1));
a = ID(loc(loc > 0), 2);
newCol = NaN(size(V,1),1);
newCol(loc > 0) = a;
V = [V, newCol];
clear newCol loc a
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Reorder columns
% #, Lake_ID, Season, Timesat_val, Timesat_val - 365
V = V(:, [1 5 2 3 4]);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Convert to DOY values - Column 6 - DOY - BUS, BUE
for i = 1:size(V,1)
    V(i,6)= round(V(i,5)-((V(i,3)-1)*365));
end
% %% Convert to DOY values - Column 6 - DOY - FUS, FUE
% for i = 1:size(V,1)
%     V(i,6)= round(V(i,5)-((V(i,3)-2)*365));
% end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Write into table
V(:,4)=[];
```

```
OUT = strcat(E, '_', location,'_dates_last.csv');
dlmwrite(OUT,V)
```

Appendix J: Create final seasonality file (MATLAB)

Example for European study area (E)

```
%% COMBINE EXTRACTED DATES INTO ONE DATASET
% removes outlying values such as freeze-up dates later than 400th day
% checks if BUE does not occur before BUS and FUE before FUS
% name: Combine.m
%% Load data - output of TS_output_process.m
clc
clear all
FUS = csvread('FUS_E_dates.csv',1,0);
FUE = csvread('FUE_E_dates.csv',1,0);
BUS = csvread('BUS_E_dates.csv',1,0);
BUE = csvread('BUE_E_dates.csv',1,0);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% 15 seasons for each lake (for first season only break-up and for last
%only freeze-up == 14 years)
r = 1809*15;
C = nan(r, 10);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
ID_lakes = csvread('ID_E.csv', 1,0);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
row = 1;
for i = 1:15:r
    ID = ID_lakes(row,2);
    A = ID_lakes(row,4);
    X = ID_lakes(row,5);
    Y = ID_lakes(row,6);
    S =0;
    for j = (i):(i+14)
        S = S+1;
        C(j, 1)= row;
        C(j, 2)= ID;
        C(j, 3) = A;
        C(j, 4) = X;
        C(j, 5) = Y;
        C(j, 6) = S;
%%%%%%%%%%%%%%%%% FUS
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        Fus = FUS(((FUS(:,1)==ID)&(FUS(:,2)==S)),4);
            if isempty(Fus)
                C(j, 7)= NaN;
            elseif Fus > 390
                C(j, 7)= NaN;
            else C(j,7)= Fus;
            end
%%%%%%%%%%%%%%%%% FUE
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        Fue = FUE(((FUE(:,2)==ID)&(FUE(:,3)==S)),5);
            if isempty(Fue)
                C(j, 8)= NaN;
            elseif (Fue > 400) | (Fue< Fus);
                C(j, 8)= NaN;
            else C(j,8)=Fue;
            end
```

```
%%%%%%%%%%%%%%%%% BUS
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
         Bus = BUS(((BUS(:,2)==ID)&(BUS(:,3)==S)),5);
                if isempty(Bus)
                    C(j,9)= NaN;
                elseif Bus < 0
                    C(j,9)= NaN;
                else C(j,9)=Bus;
                end
%%%%%%%%%%%%%%%%% BUE
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        Bue = BUE(((BUE(:,2)==ID)&(BUE(:,3)==S)),5);
                if isempty(Bue)
                    C(j, 10)= NaN;
                elseif (Bue <50) | (Bue < Bus);
                    C(j,10)= NaN;
                else
                    C(j,10)=Bue;
                end
    end
    row = row+1;
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
for c = 7:10
    for l = 1:size(ID_lakes,1)
        ID = ID_lakes(l,2);
        t = C(C(:,2)==ID, c);
        x = nanmean(t);
        sd = nanstd(t);
        for v = 1:15
            z = (t(v)-x)/sd;
            if abs(z)>2.572
                C((((l-1)*15)+v),(c))= NaN;
            end
        end
    end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
dlmwrite('E_Seasonality.csv', C(:,1:10));
```