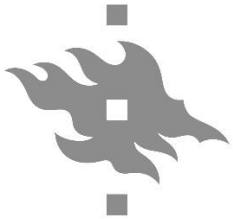


Master's thesis in Geography

Physical Geography

Mesoscale modelling of periglacial landforms in the circumpolar Arctic

Tommi Kukkonen

2020

Supervisor: Miska Luoto

Master's Programme in Geography

University of Helsinki

Faculty of Science

Department of geosciences and geography

P. O. Box 64 (Gustaf Hällströmin katu 2) 00014 University of Helsinki

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI
MATEMAATTIS-LUONNONTIETEELLINEN TIEDEKUNTA
MATEMATISK-NATURVETENSKAPLIGA FAKULTETEN
FACULTY OF SCIENCE

Tiedekunta – Fakultet – Faculty	Koulutusohjelma – Utbildningsprogram – Degree programme
Faculty of Science	Master's Programme in Geography

Opintosuunta – Studieinrikting – Study track
Physical Geography

Tekijä – Författare – Author
Tommi Kukkonen

Työn nimi – Arbetets titel – Title
Mesoscale modelling of periglacial landforms in the circumpolar Arctic

Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Master's thesis	May 2020	86 pages

Tiivistelmä – Referat – Abstract

The Arctic is warming with an increased pace, and it can affect ecosystems, infrastructure and communities. By studying periglacial landforms and processes, and using improved methods, more knowledge on these changing environmental conditions and their impacts can be obtained. The aim of this thesis is to map studied landforms and predict their probability of occurrence in the circumpolar region utilizing different modelling methods.

Periglacial environments occur in high latitudes and other cold regions. These environments host permafrost, which is frozen ground and responds effectively to climate warming, and underlays areas that host many landform types. Therefore, landform monitoring and modelling in permafrost regions under changing climate can provide information about the ongoing changes in the Arctic and landform distributions. Here four landform/process types were mapped and studied: patterned ground, pingos, thermokarst activity and solifluction.

The study consisted of 10 study areas across the circumpolar arctic that were mapped for their landforms. The study utilized GLM, GAM and GBM analyses in determining landform occurrences in the arctic based on environmental variables. Model calibration utilized logit link function, and evaluation explained the deviance value. Data was sampled to evaluation and calibration sets to assess prediction abilities. The predictive accuracy of the models was assessed using ROC/AUC values.

Thermokarst activity proved to be most abundant in studied areas, whereas solifluction activity was most scarce. Pingos were discovered evenly throughout studied areas, and patterned ground activity was absent in some areas but rich in others. Climate variables and mean annual ground temperature had the biggest influence in explaining landform occurrence throughout the circumpolar region. GBM proved to be the most accurate and had the best predictive performance. The results show that mapping and modelling in mesoscale is possible, and in the future, similar studies could be utilized in monitoring efforts regarding global change and in studying environmental and periglacial landform/process interactions.

Avainsanat – Nyckelord – Keywords
periglacial landforms, logistic regression, Arctic, modelling

Säilytyspaikka – Förvaringställe – Where deposited
HELDA, an open access repository of the University of Helsinki

Muita tietoja – Övriga uppgifter – Additional information

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI
MATEMAATTIS-LUONNONTIETEELLINEN TIEDEKUNTA
MATEMATISK-NATURVETENSKAPLIGA FAKULTETEN
FACULTY OF SCIENCE

Tiedekunta – Fakultet – Faculty	Koulutusohjelma – Utbildningsprogram – Degree programme
Matemaattis-Luonnontieteellinen tiedekunta	Maantieteen maisteriohjelma

Opintosuunta – Studieinrikting – Study track
Luonnonmaantiede

Tekijä – Författare – Author
Tommi Kukkonen

Työn nimi – Arbetets titel – Title
Periglasiaalisten muodostumien mesoskaalainen mallinnus sirkumpolaarisella Arktiksella

Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Kirjallinen tutkielma	Toukokuu 2020	86 sivua

Tiivistelmä – Referat – Abstract

Arktinen alue lämpenee kiihtyvällä tahdilla, ja sillä voi olla merkittäviä vaikutuksia ekosysteemeihin, infrastruktuuriin ja alueen yhteisöihin. Tutkimalla periglasiaalisia muodostumia ja prosesseja, sekä hyödyntämällä paranneltuja menetelmiä on mahdollista saada lisätietoa muuttuvista ympäristöolosuhteista ja niiden vaikutuksista. Tämän opinnäytetyön tarkoituksena on kartoittaa tutkimuksen kohteena olevia muodostumia/prosesseja ja pyrkiä ennustamaan niiden esiintymisen todennäköisyys sirkumpolaarisella vyöhykkeellä hyödyntäen eri mallinnusmenetelmiä.

Periglasiaalisia ympäristöjä esiintyy korkeilla leveysasteilla ja muilla kylmillä alueilla. Tavallisesti näissä ympäristöissä esiintyy myös ikiroutaa, joka on jäätynyttä maata ja joka reagoi tehokkaasti ilmaston lämpenemiseen. Ikirouta-alueilla esiintyy monia muodostumatyyppejä, joita tutkimalla ja mallintamalla voidaan saada tietoa muuttuvista olosuhteista Arktiksella sekä muodostumien ja prosessien levinneisyyksistä. Tässä tutkimuksessa tarkastelun kohteena oli neljä eri muodostumatyyppiä: kuviomaat, pingot, termokarstiaktiivisuus sekä solifluktio.

Tutkimus koostui kymmenestä tutkimusalueesta sirkumpolaarisella Arktiksella, joista kaikista kartoitettiin muodostumia. Tutkimuksessa hyödynnettiin GLM, GAM ja GBM analyyseja, joilla mallinnettiin muodostumien esiintymistä Arktisella alueella perustuen ympäristömuuttujiin. Mallien kalibroinnissa hyödynnettiin logit linkkifunktiota ja evaluoinnissa selitettyä poikkeavuutta. Data jaettiin evaluaatio- ja kalibraatio aineistoihin, jotta ennustekapasiteettia voitiin arvioida. Ennusteiden tarkkuuden määrityksessä hyödynnettiin ROC/AUC arvoja.

Termokarstiaktiivisuutta esiintyi eniten tutkituilla alueilla, ja solifluktiota kaikkein vähiten. Pingoja esiintyi tasaisesti kaikilla alueilla. Kuviomaata ei löytynyt tietyiltä alueilta, mutta jollain alueilla sitä esiintyi runsaasti. Ilmastomuuttujat ja maanpinnan keskilämpötila osoittautuivat merkittävimmiksi muuttujiksi selitettäessä muodostumien esiintyvyyttä sirkumpolaarisella alueella. GBM oli mallinnusmenetelmistä tarkin ja sillä oli paras ennustuskyky. Tulokset osoittavat, että kartoitus ja mallinnus mesoskaalassa on mahdollista. Tulevaisuudessa tuloksia voitaisiin hyödyntää Arktisen alueen monitorointi- ja globaalimuutostutkimuksiin, sekä mahdollisesti arvioitaessa periglasiaalisten muodostumien suhdetta alueellisiin ympäristömuuttujiin.

Avainsanat – Nyckelord – Keywords
periglasiaaliset muodostumat, logistinen regressio, Arktinen alue, mallinnus

Säilytyspaikka – Förvaringställe – Where deposited
HELDA, Helsingin yliopiston avoin digitaalinen arkisto

Muita tietoja – Övriga uppgifter – Additional information

Table of contents

1. Introduction

Geomorphology is a study of landforms, landscapes and land-surface features and their description, classification, origin, development, and history on planetary surfaces (Huggett 2011: 3; Pareta & Pareta 2015). The sub-discipline of geomorphology is periglacial geomorphology that focuses on landforms and processes occurring in cold, polar or high-altitude environments (Barsch 1993). Traditionally, these processes and landforms have been studied and mapped using photo interpretation methods (Brenning et al. 2007) and field surveys (Luoto et al. 2010), but nowadays these approaches together with Geographic Information System (GIS) and statistical methods provide a framework for analysing and displaying large datasets in even more complex systems and extensive areas beyond traditional mapping methods (Bishop et al. 2012). For regions of large extent and of little available information, these newer methods can be financially cheaper and less time consuming (Etzelmüller et al. 2006).

Determination of controlling environmental factors on earth surface processes, landforms and knowledge of their distribution in cold areas is a central concept in periglacial geomorphology (Luoto & Hjort 2005; Hjort et al. 2007). Modern techniques in spatial modelling can be used to create spatial predictions of geomorphological phenomena in unsurveyed areas and therefore contribute in e.g. land-use planning, slope failure research and predictive mapping of soil and landforms (Marmion et al. 2008). Although the study of geomorphological systems in cold regions has developed in the last decades, some processes are still much less studied (Etzelmüller et al. 2001a), and different spatial modelling methods are being constantly improved (Guisan & Zimmermann 2000). This is important because ongoing and predicted warming of high latitudes and the Arctic can severely impact infrastructure, ecosystems and communities in the region (Westermann et al. 2015). By using spatial models, it is possible to assess these impacts and determine what effects they have on geomorphological processes (Fronzek et al. 2006).

The aims of this thesis are **1)** to determine the spatial distribution and abundance of studied periglacial landforms and processes, **2)** determine which environmental variables affect the predictive variables most in the study areas using logistic regression, **3)** model the distribution of landforms and processes in the circumpolar Arctic region and to find out their probability of occurrence using GLM, GAM and GBM, and **4)** to assess the accuracy and prediction performance of the models.

2. Theoretical background

2.1 Northern circumpolar region

Although the definitions of the northern circumpolar region vary, geographically it can be defined as a complete region extending from the subarctic (50°N) environment beyond the Arctic Circle (66°33'N) (Hoberg et al. 2012). Typical to this entire region is the Arctic ocean, big Arctic states such as Canada, Russia, USA and Scandinavian countries, as well as vast regionwide resources. Lately, the areal focus has increasingly been on the changing environment and warming climate and the effects that these changing conditions have on the ecosystems of the circumpolar region (Dodds & Nuttall 2019: 3 – 11). In a biogeographical sense the high latitude environment can be separated into High- and Low Arctic environments (Figure 1) (French 2018: 41), of which Canada has the greatest geographical part of the High Arctic, and Russia of the Low Arctic (Barry & Hall-McKim 2018: 169).

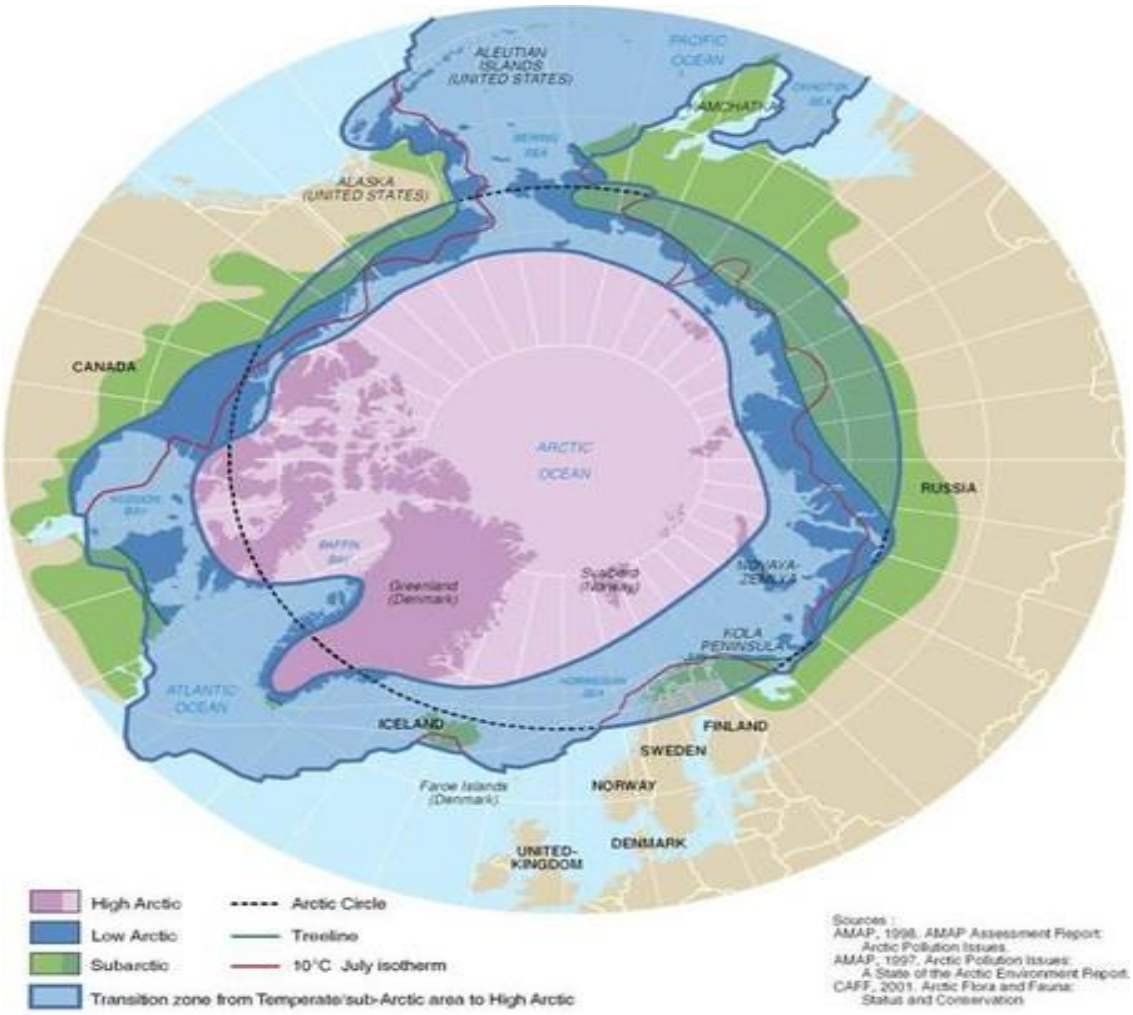


Figure 1. Boundary definitions of the high latitude circumpolar region (Rekacewicz, UNEP/GRID-Arendal 2005).

2.2 Periglacial environment

Periglacial environments are regions with year-round cold conditions and vast tundra areas and polar deserts. These environments are common in high latitudes and altitudes and other places where annual temperatures generally remain low. Characteristic factors for periglacial environments include soil material displacement, water substance migrations, unique terrain features and vegetation, with related freezing and thawing processes. The largest periglacial environments by areal extent are the Arctic and Subarctic regions of the high latitudes (Williams & Smith 1991: 2; French 2018: 5 – 6) that consist of characteristic periglacial processes and landforms related to cold arctic

climates involving freezing and thawing processes, frost action and ice- and wind bound physical weathering (Arbogast 2013: 473). The geographical location (i.e. longitude and latitude), as well as elevation, are important contributors in determining temperature conditions (Aalto et al. 2012), and high latitude environments especially are prone to quick changes due to effects caused by climate warming owing to the Arctic regions acting as reservoirs to strong greenhouse gases CO2 (carbon dioxide) and CH4 (methane) (French 2018: 39).

The characteristic factors for the periglacial environment are periglacial conditions, which in turn are defined by active frost processes, such as permafrost (French 2018: 65). It develops in environments that experience periglacial conditions in the form of seasonally- or perennially frozen ground (Barry & Hall-McKim 2018: 168) and is therefore thermally defined. The cold climate and local factors such as depth of snow cover, wind intensity, precipitation, vegetation, heat transfer through active layer, energy exchange on the ground level and local geological and hydrological conditions ultimately define the presence of permafrost (Etzelmüller et al. 2001b). The periglacial environment is spatially covered in the permafrost of different extent and is widespread. Permafrost covers 17 million km^2 of the terrestrial Northern Hemisphere (approximately 1/4) covering large areas of the Arctic and Subarctic and can be considered to be continuous, discontinuous, sporadic or isolated permafrost (Figure 2) (Yang et al. 2010; Huggett 2011: 291; Barry & Hall-McKim 2018: 14).

In general, frozen ground is considered to be permafrost if it occurs below 0 °C of mean ground temperature and for more than two back to back years. It is common for periglacial environments in both High- and Low Arctic. In both regions permafrost extends 10-30 metres in depth, but in the Low Arctic permafrost temperature is ranging between -3 and -4 °C, and in the High Arctic between -10 and -14 °C (Williams & Smith 1991: 1; French 2018: 70). Permafrost typically develops where the depth of winter freezing is greater than the depth of thawing in summer (Huggett 2011: 291). Permafrosts active layer is located at the top part of permafrost near the surface, and it experiences freeze-thaw action on a seasonal or daily basis. The thickness of the active layer varies depending on latitude, and in the Higher Arctic it can be 10 cm in depth and on the Lower Arctic as much as 2-3 metres (Huggett 2011: 291; Arbogast 2013: 474).

Permafrost responds effectively to climate warming, and as an important factor of the atmosphere-ocean-land system, it has a strong effect on the outcome of climate-induced changes (Jorgenson & Grosse 2016). In order to observe past and future changes in permafrost and in the Arctic region, it is possible to examine changes in different environmental indicators in high latitudes, such as periglacially bound landforms. These landforms can consist of e.g. frost mounds, thermokarst activity and mass wasting processes, and by studying these different cold environmental landforms, it is possible to acquire information on geological, cryological and hydrological changes in delicate ecosystems (Grosse & Jones 2011).

Figure 2. Permafrost distribution pictured in the Northern circumpolar region and showing the spatial permafrost extent in different latitudes (Rekacewicz 2005).

2.3 Global change

High latitude polar regions are facing major changes in the future, as the effects of climate change and their influence are expected to strengthen in these regions. The Earth has experienced 0.8 °C increase in global temperatures measured from the late 1800s to this day, with overall temperature increase in the Arctic being approximately 2-3 °C in total over the same time period (Post et al. 2019). The atmospheric CO2 concentrations have risen from 316 ppm in 1959 to 400.8 ppm in 2015, and in the future temperatures are expected to increase between 1.0 °C (RCP 2.6) and 3.7 °C (RCP 8.5) by the end of 2100s (based on IPCC models). For high latitude periglacial environments this would mean shrinking of Arctic sea ice, glaciers, ice caps, permafrost thawing, changes in tundra vegetation and hydrological changes throughout the polar region (Barry & Hall-McKim 2018: 386 – 389).

The duration of the arctic ice sheet is expected to shorten, with some models predicting completely ice-free summers in the coming decades (Barry & Hall-McKim 2018: 392). The volume of sea ice loss is expected to encompass all calendar months, with summer months sea ice being affected the most. The record for minimum sea ice extents have been measured in the last decade, and some models suggest that the extent of sea ice will drop below 1 million km^2 (current extent 4.2 million km^2) in September months, and an additional 800 Gt of CO2 being released in the next 20 – 25 year period (ECMWF 2019; Post et al. 2019). The disappearance of sea ice is also expected to lessen albedo feedback, and thus strengthen the ongoing positive feedback mechanism on polar ecosystems (Barry & Hall-McKim 2018: 385).

Precipitation is expected to increase due to airs ability to hold more water. This would increase year-round wetness around the Arctic region. Because of this, evaporation is also expected to increase (Barry & Hall-McKim 2018: 393). Precipitation in winter is expected to change compared to summer precipitation. This is especially seen as even wetter and purer rain-like precipitation in winter months. These changes in precipitation could indicate greater evapotranspiration intensity and/or greater subpermafrost groundwater infiltration and can be considered directly or indirectly global climate change driven effects (Hinzman et al. 2005).

A warming climate will most likely expand the active layer in permafrost, causing changes in ground hydrology and runoff. This would lessen the areal extent of continuous permafrost and trigger climatic feedbacks locally and globally (Westermann et al. 2015; French 2018: 102) as well as creating changes in hydrological and geoecological landscapes. Since arctic soil and especially northern wetlands contain over 50% of soils organic carbon, thawing of permafrost would release much of it in the atmosphere in the form of CH4, acting as a positive feedback mechanism to global climate warming (Arbogast 2013: 230 – 231; Karjalainen et al. 2019; Post et al. 2019). According to Jorgenson (2013), permafrost has already warmed by up to 4 °C in the last 30-year time period.

Regardless of RCP scenarios Arctic warming can reach 4 °C mean annual warming and up to 7 °C in the late boreal autumn period. However, under RCP 8.5 scenario the Arctic warming is expected to reach as high as 13 °C in months of late boreal autumn by the end of the 21st century (Post et al. 2019). Thus, increased knowledge in permafrost dynamics on hemisphere-scale are important in assessing feedback systems and climate change mechanics in higher latitudes in order to gain more information about changing conditions as well as infrastructure and maintenance threats (Karjalainen et al. 2019).

2.4 Studied periglacial formations

2.4.1 Patterned ground

Freezing and thawing processes can affect ground in many ways in the periglacial domain. One significant landform that this type of action creates is called *patterned ground* (Figure 3), which is a significant feature in the Arctic and Subarctic environments (French 2018: 240). Although the origins of the patterned ground are not fully clear, three major formative components are understood to exist (Huggett 2011: 296). Patterned ground forms as an outcome of cryoturbative action and involve frost heave, thaw settlement, and particle movement near the ground. This is all due to changes in ice bodies which occur in temperature change scenarios (Hjort & Luoto 2006; French 2018: 242). The patterned ground can be discovered in a variety of environments, and

thus a universal formative process is yet to be found, although several hypotheses have been presented. Even though similarly looking patterned ground in different environments may not share common formative processes, it has been agreed upon that many of them have similar, polygenetic origins (Haugland 2006; Hjort & Luoto 2006).

Figure 3. Patterned ground landforms Near Yukon River, Alaska, U.S (modified after Cowals 1973)

Patterned ground occurs typically in the active, top layer of the ground but also to some degree in seasonally frozen ground layers. The patterned ground can emerge as various types of geometries (Rowley et al. 2015) but usually occurs as circles, stripes, polygons and hummocky grounds and can be visible even from the air. Depending on the forming material, patterned ground features can be described as sorted or non-sorted and their shape can be highly dependable on the slope angle in which they appear (Huggett 2011: 301; French 2018: 240). Circular patterned ground shapes are most common, and they can appear as either flat or hummocky surfaces. Non-sorted circles can be 0.5-3.0 metres in diameter and 0.5 metres in height. Non-sorted material is usually fine-grained and frost-susceptible. Formation of circular patterned ground develops at the bottom

and top sections of the active layer as the ice lenses experience freeze and thaw action. This causes gravity-oriented cell-like circulation within the soil, which raises soil material and expands it, forming circular patterns on the ground. In areas with small, less than 1.0 m diameter polygonal patterned ground areas the temperature variations and patterned ground forming processes can constitute in forming nets in which several polygons have united. They are commonly smaller than circles and occur mostly in the High Arctic. Stripes occur in slopes and are usually 0.3-1.0 m wide. The same cryoturbative processes to create patterns act on a hillslope but appear in striped forms due to surface wash action (French 2018: 240 – 242).

2.4.2 Pingos

A class of different intrapermafrost features include frost-, ice- and earth-mounds as well as palsas and various hummocky landforms (Flemal 1976). *Pingos* (Figure 4) are a type of frost-mounds that only appear in areas underlain by permafrost activity (Mackay 1998). They are perennial in nature but still relatively rare in permafrost dominated landscapes. They stand out from regular cold environment frost mounds in that they can be relatively large in size. They can be up to 300 m in diameter and up to 60 m in height. Pingos usually appear as round or in conical form, and they consist of ice- or icy sediment body within them that may be covered by a vegetational mat (French 2018: 159).

Figure 4. Pingo pictured in Northwest Territories in Canada, just outside Tuktoyaktuk Peninsula (Jones 2013).

Pingos can be either hydraulic (open) system pingos or hydrostatic (closed) system pingos (Figure 5). Open system pingos develop under artesian pressure at places of intra- or sub permafrost and where groundwater reaches the ground surface. In the frozen ground, water can have a hard time finding these gaps, but one typical place for pingos to form are the *taliks*. They are unfrozen ground between frozen soil in which the groundwater can then seep through to the surface. Places where this kind of groundwater seepage is common and taliks commonly occur are at the base of hillslopes, alluvial fans or valley bottoms (Rowley et al. 2015; French 2018: 159 – 161), and usually in discontinuous permafrost zone for open system pingos (Burr et al. 2009). In these environments, open system pingos appear usually isolated or in small groups and can grow to be 20-50 m in height. Hydrostatic (closed) system pingos, unlike open system pingos, typically develop in continuous permafrost zones within shallow lakes or dried out lake beds where permafrost begins expanding in previously unfrozen saturated sediment, thus raising the soil upwards. Commonly these lake beds comprise a sub-talik where pore water can flow freely under its own hydrostatic (gravitational) mass forming an ice lense under the lake (Burr et al. 2009; French, 2018: 161). In terms of pingo growth, simple water to ice expansion is not enough for the pingos to grow so distinguishably large. For both types of pingos, relatively constant groundwater flow to pingo ice lense is required to maintain pingo growth which can be seen as periods of uplift or pulses (Burr et al. 2009). Nevertheless, pingos are very individualistic and vary widely depending on height, diameter, slope gradient and spacing between them (Rowley et al. 2015).

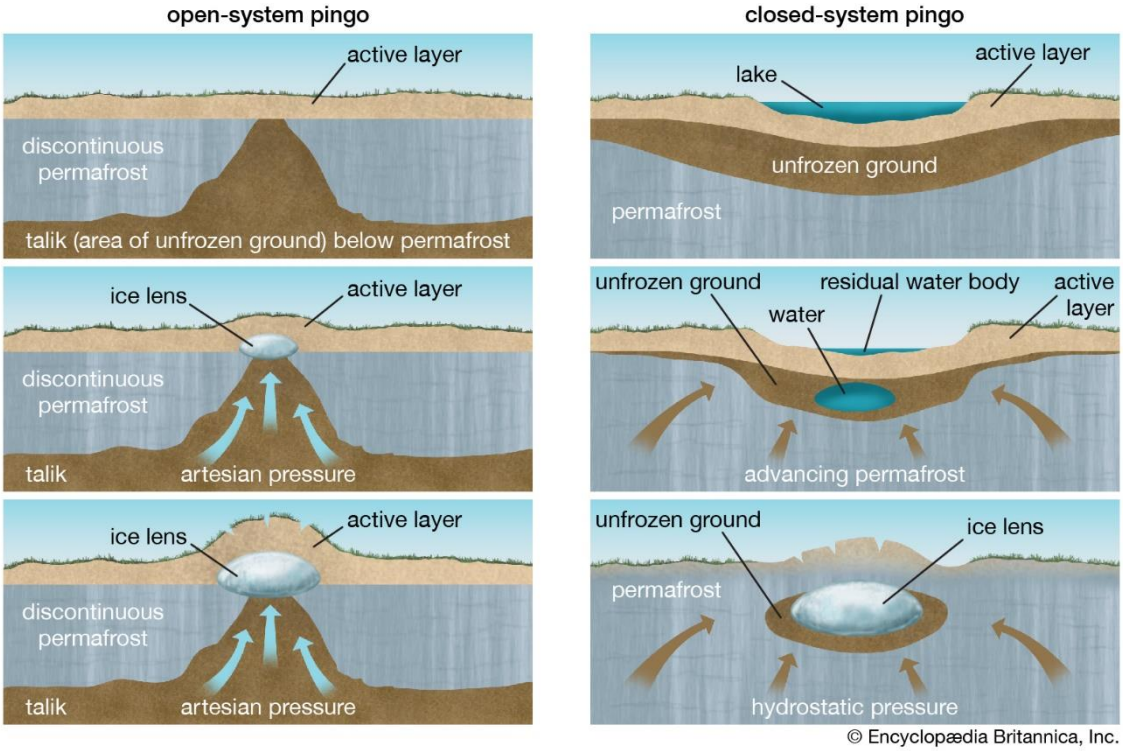


Figure 5. A visual explanation of a formation of open system and closed system pingo (Encyclopædia Britannica n.d.)

As the age of pingos grows, their rate of growth decreases. At some stage, pingos can also collapse. This process is first seen on the summit of pingos as cracking and ruptures. Next, the summit can be open to erosion and temperature changes that erode and break the ice lense even further causing it to melt faster. Eventually, pingos collapse, leaving a shallow and rimmed crater-like depression in their place with a possible lake in the middle (French 2018: 162). These depressions can also be referred to as *pingo scars* outlining the position of the original pingo (Flemal 1976).

The distribution of pingos covers pretty much the same extent as the permafrost region. Main areas of pingo distribution are known to be Northwestern Canada, Siberia (Figure 6) and Alaska. In smaller numbers pingos have been observed in Greenland, Svalbard, Scandinavia, China, Mongolia and Tibetan Plateau. In Tuktoyaktuk Peninsula (Canada), which has the greatest pingo concentration in the world, approximately 1350 closed-system pingos have been observed, although only 50 of these are known to be actively growing (Mackay 1998; Burr et al. 2009; Grosse & Jones 2011).

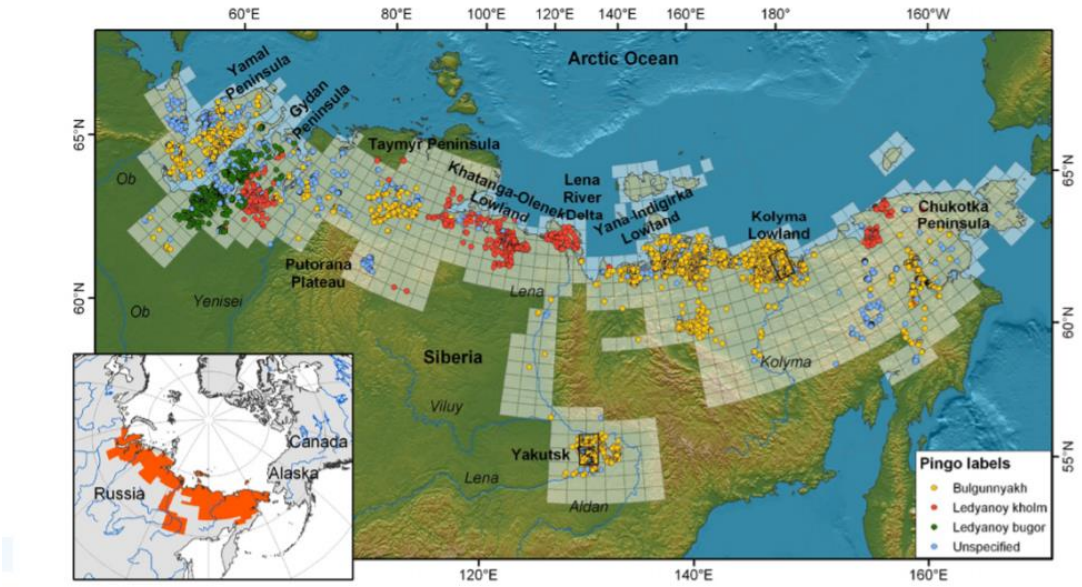


Figure 6. Example of pingo distribution in Northern Asia based on research conducted by Grosse & Jones (2011). Marked pingo labels are Russian names describing different types of pingos (Grosse & Jones 2011)

2.4.3 Thermokarst activity

Thermokarst is related to a variety of processes or features seen in periglacial environments as geomorphological effects or landforms that are affected by thaw related processes. Thermokarst is a common phenomenon occurring in areas underlain by permafrost (French 2018: 169) and has a major impact on permafrost landscape and biogeochemical cycle in the Arctic and boreal lowlands (Farquharson et al. 2016; Jorgenson & Grosse 2016). Although over twenty different thermokarst landforms can be identified based on their characteristics and topography (Jorgenson 2013), this chapter focuses mainly on large thermokarst ponds and lakes as well as freeze-thaw processes linked to them (Figure 7).

Figure 7. Thermokarst lakes pictured in the central part of Nadym-Pur watershed, Russia (Kirpotin 1999, in *Hydrological Changes: Historical Analysis, Contemporary Status, and Future Projections 2012)*

The formation of thermokarst can occur in at least two ways: thawing action of permafrost in the near-surface layer, which is followed by the accumulation of water at the surface or increase in active layer thickness. Second is the enlargement of lakes due to thermal erosion or subsidence. The major causes for thermokarst development, however, are understood to be the long- and short-term changes in climate, although the expected effects of climate variations to thermokarst formation, especially under the influence of global warming, can be problematic (French 2018: 170). Other major factors controlling thermokarst development are hydrology and characteristic substrate properties, which affect the formation of thermokarst locally (Bouchard et al. 2014). Fine-grained sediments often contain a lot of ice, and with warm near-surface conditions, can create thermokarst landscapes easier than in other regions. Thermokarst landscapes can occur in different ways such as lakes, basins or mounds. For example, in Siberia, thermokarst is often seen as groups of ponds and depressions that mark the extent of near surface thaw processes (French 2018: 189). In Alaska, thermokarst appears as drained thermokarst-lake basins and lakes, with less occurrence of shallow

pits and throughs (Farquharson et al. 2016). Thermokarst ponds and lakes can be 1-2 km in length, but usually, they are much smaller. The lakes are usually shallow, only up to few meters deep (French 2018: 189), and typically freeze to bottom (Jorgenson 2013). In many occasions they appear as elongated bodies, and although the process of growth and drainage of thermokarst lakes is not fully understood it is clear that this is an ongoing and dynamic process in the tundra (French, 2018: 189). Thermokarst landscapes (lakes and ponds) occur abundantly in Russia, Alaska and northern parts of Yukon and Northwest Territories of Canada (Jorgenson 2013).

As cloud cover and the overall amount of precipitation is expected to increase due to climate warming, it may be difficult to interpret the regional effects on thermokarst development (French 2018: 174). In Arctic regions, thermokarst activity and thus the number of thaw lakes has increased due to increased ground temperatures generated by warming climate (Olthof et al. 2015). According to Wetterich et al. (2018), thermokarst lake formation in the early Holocene was vast due to global warming, and late Holocene cooling led to permafrost aggradation due to different environmental changes. As a result, the landscape mosaic of periglacial landforms was intensified. Recent climate warming has increased the number of thermokarst lakes, thermokarst activity and lake expansion in Arctic regions as shown by e.g. Olthof et al. (2015) and Farquharson et al. (2016).

2.4.4 Solifluction

Mass wasting is used to describe the movement of soil mass downhill due to the effect of gravity and can be described with flow, slip and fall movements. *Solifluction* (Figure 8) is a form of mass wasting used to describe saturated soil moving from higher to lower ground (French 2018: 219) and are common features in Low Arctic environments (Huggett 2011: 295). Solifluction rates and processes are dependable on different environmental factors such as climate, topography, hydrology and geology (Matsuoka 2001). Intense frost action may cause horizontal or vertical movement of soil mass in periglacial environments (Huggett 2011: 294) and is often referred to as *gelifluction.* It describes solifluctive action in areas where the frozen ground is common and where

frost heaving affects ground surface, and thawing reduces internal friction initiating movement and cohesion in the soil, thus creating more effective creep (Millar 2013; French 2018: 223). When frozen ground thaws and consolidates to the soil layer, the increased moisture content can induce the deformation of the soil and increase the soils stress level to pre-failure stage. Sometimes additional water from precipitation may add to soils moisture content, thus increasing the possibility of soil movement (Matsuoka 2001). Gelifluction only occurs during the warm, thawing seasons and affects only in the active layer, which is approximately the top 50 cm of the ground. Solifluction material is normally matrix-supported, unstratified and fine-grained sediment, and in long-term produces features such as tongue-shaped lobes and uniform sheets of locally derived material, as well as stripes and hummocks. In permafrost regions, the solifluction lobes can be 60-75 cm in thickness (Matsuoka 2001; French 2018: 223). Key factors affecting the process and formation of solifluction features are most likely mean annual air temperature and mean of maximum snow depth. Also, it has been observed that where solifluction activity is relatively high, the sites are highly affected by winter snow distribution, a high degree of slope angle and amount of vegetation covering the ground (Rowley et al. 2015; French 2018: 225).

Figure 8. Saturated conditions on a solifluction slope. Unknown location (Giardino 2008, in *Periglacial Processes and Landforms in the Critical Zone 2015)*

In contrast to solifluction, mass wasting can be more rapid and devastating due to its mass and speed of detachment. Although mass wasting is less studied than solifluction, they too are a common feature in periglacial environments, and especially in North America and Northern Siberia in the Arctic- and Subarctic lowlands, where the ground is underlain by ice-rich unconsolidated sediments. The factors making mass wasting processes so high volume in cold environments are rapid thawing of ice and high precipitation in the summer seasons. These factors with possible disturbance to active layer due to surface vegetation loss and human-induced terrain disturbance can cause large mass wasting processes in the periglacial environment (French 2018: 226). Soil movement processes in the periglacial region provide loose material and sediments for transport and limit the infiltration of water to the ground. With higher environmental variations, such as precipitation changes and more intense freeze-thaw cycles, it is possible that mass wasting processes and debris flows happen more suddenly and rapidly (Millar 2013). Mass wasting can thus be a problem, for example, to

geoengineering and construction in places where these processes occur, and land movement is common (French 2018: 220).

2.5 Modelling in geomorphological research

Application of statistical methods in predicting the distribution of geomorphological processes and landforms has increased in recent decades and especially in the 21st century. This allows researchers to create predictive maps describing spatial relationships between environmental variables and geomorphologic processes or landforms utilizing different statistical techniques (Luoto & Seppälä 2002; Luoto & Hjort 2005; Luoto et al. 2010). In the field of geoscience, geomorphic distribution models (GDMs) have become widely popular empirical models to predict field observations related to explanatory variables, and especially in terms of changing climate, the role of statistical modelling has become more crucial in understanding the effects of climate change in Earth surface systems (Hjort & Luoto 2013; Karjalainen et al. 2019). Especially frost-related features and systems have their own complexity in terms of linear relationships, and thus multivariate approach is needed to explore the correlation between variables (Hjort et al. 2007). Predictive geomorphological modelling can be divided into three different stages: 1) model development and calibration based on the relationship between environmental variables and studied geomorphological phenomenon; (2) model evaluation with a test data set or other validating techniques; and (3) applying the model to a geographical database to create a predictive map (Marmion et al. 2008). The statistical methods for GDMs and their effectiveness and predictive accuracy have been assessed in the past by e.g Luoto & Hjort (2005); Luoto et al. (2010) and Hjort et al. (2010).

In the past geomorphological information and mapping has been based on field data, remote sensing and cartographic mapping. On a regional-scale physiographic analysis, geomorphology and mapping used to be based on small-scale maps and photography interpretation in order to classify types of terrain and regional scale features. These approaches were qualitative interpretations in nature, but nowadays, the analysis of data has shifted more towards quantitative approach (Bishop et al. 2012). The

advancements in the fields of computer science, photogrammetry, remote sensing, geographic information technology (GIT), statistics, geophysics, numerical modelling and geodesy are allowing geoscientists to utilize new spatio-temporal data, processing approaches, models and geocomputational algorithms (Bishop 2013).

Geomorphometry refers to the study of surface processes and geomorphological mapping using different methodological approaches, such as sampling land surface attributes, geodesy and digital terrain modelling (DTM), processing of digital elevation models (DEM), and producing geomorphic information. This approach allows geoscientists to model and analyze geomorphic systems in new ways (Bishop 2013). Remote sensing (e.g satellite data) can be utilized in gathering spatially continuous and versatile information of environmental determinants of the phenomena studied. Together with GIS-data and variables, it is possible to gather a variety of environmental surrogates in order to model complex spatial processes and landforms (Hjort & Luoto 2006). Utilization of satellite data has increased in recent years in the Arctic for various geomorphological purposes, which include different mapping, monitoring and analyzing purposes (Perreault et al. 2017).

2.5.1 Nonlinearity

In the field of geomorphology, the systems are typically nonlinear in nature due to the fact that they are mostly threshold-dominated. Nonlinear systems can experience behavioural complexity that is not common in linear geomorphic systems (Phillips 2006). In linear systems, the observed response variable and the values that are predicted by the linear model are linearly linked, i.e. they match one to one (Atkinson et al. 1998). In case of nonlinear systems if the outputs are not proportional to inputs along with the entire input range, and although they may be simple and fairly predictable, in geomorphic systems they are almost undoubtedly nonlinear due to their nature, meaning that they do not match perfectly. The chaotic nature of nonlinearity can be reduced with a greater amount of components, variables or processes involved in describing the geomorphic system, thus making the results more singular. This is

achieved by not reducing but increasing the number of factors or variables involved in the model (Phillips 2003, 2006).

3. Study area

In the High Arctic, the temperatures drop below zero for 5-6 months annually, and polar deserts are common in these regions (French 2018: 24), with the typical seasonal maximum for ice thickness reaching up to 2 metres. Regionally, however, ice thickness can vary significantly (Barry & Hall-McKim 2018: 188). Regions experience very cold winters, cold summers and low precipitation levels, with growing season extending between 1-2 months (French 2018: 43). Vegetation in the High Arctic is mostly tundra with polar deserts of different sizes while the Low Arctic domain consists mostly of low-shrub tundra vegetation (Barry & Hall-McKim 2018: 169; French 2018: 43). In the Low Arctic, the winters are also considered to be very cold, but temperature variations are more extensive, and vegetation can cover 80-100% of the ground. The precipitation is also low and having annually longer precipitation periods than in the High Arctic, with growing season extending between 3.5-5 months (French 2018: 43).

High- and Low Arctic form the periglacial domain, with the boreal forest region extending south (French 2018: 6). The boreal forest is located between latitudes 50° and 60° covering large parts of North America and Eurasia. In the transition zone (or treeline) between tundra and boreal forest trees are typically sparse and short (Dodds & Nuttall 2019: 49). The transition zone is approximately 30-150 km wide zone, and south of the tree line is defined by coniferous forest reaching through several regions and countries along with the circumpolar such as Russia, Scandinavia and North America. Of the global forest-canopy cover the extent of the boreal forest is about a quarter, and large parts of the boreal are covered in wetlands and treeless bogs across the region aswell (French 2018: 56).

The study consisted of 10 areas around the Arctic/Subarctic circumpolar region (Figure 9). The size of each area was approximately 50 x 250 km (west-east, north-south) in

extent and were chosen based on the geographical extent of the circumpolar region (i.e. so that the study areas chosen would be spread out quite evenly), available literature, and that the areas would be visually interpretable. In most cases, the latter meant above the circumpolar/hemiarctic treeline. The study areas needed to host landforms and processes that were of interest in this thesis, preferably so that each landform would be located in each area to at least some extent and that they could be visually mappable. The studied regions were mapped in order from western North Asia to eastern North America and numbered from 1 to 10 (Figure 9). Exact study area coordinates are presented in Appendix 1. Study areas are presented more precisely in Appendices 2 and 3.

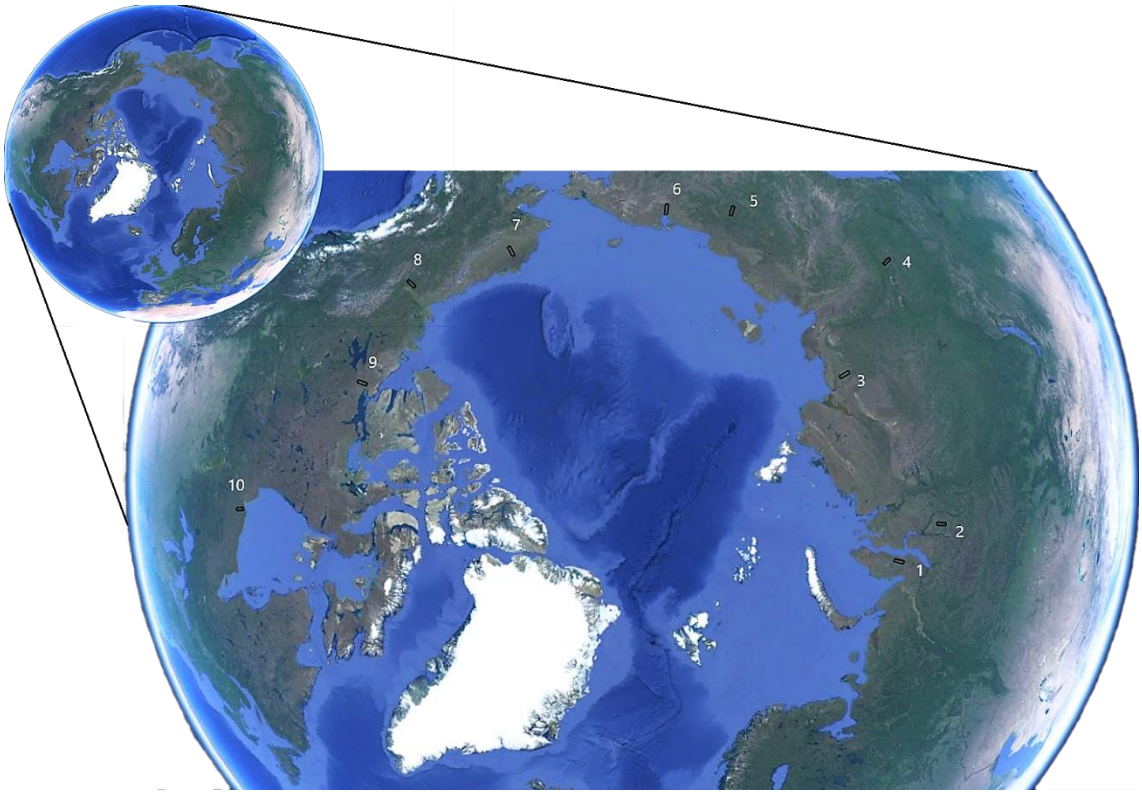


Figure 9. Selected study areas marked as black rectangles in the circumpolar region and numbered from 1 to 10. Eye altitude 7602 km measured from space (Google Earth Pro 2015)

3.1 Russia and Northern Asia

Study area 1 and *study area 2* are located in The Yamalo-Nenets district in Northwestern Siberia in central Russia. This region is covered in swamps and barren tundra, with stunted forest covering the southern parts of the district and extreme weather conditions covering the entire district. Study area 1 is specifically located in Yamal Peninsula, which is an Arctic lowland area. The area is surrounded by the Kara Sea to the west, Malygina Strait to the north and Gulf of Ob to the east. The Peninsula covers approximately 122 000 km^2 and has an irregular surface due to glacial and marine deposition. The coasts are low lying, but the topography rises towards the mainland in the south. Study area 2 is located further South between Purovsky district and Krasnoselkupsky district (Augustyn et al. 2019).

Study area 3 is located in Anabarsky district in Northern Yakutia near the Arctic coast. The climate in the region is one of the most hostile in the world with winter temperatures dropping close to -50 °C. Precipitation is typically also low with tundra vegetation being the most common type of vegetation (Augustyn et al. 2013). The area is majorly lowland (Khatanga-Olenek or Khatanga-Anabar lowland) consisting of medium to coarse-grained sandy lithologies and intensive water migration and ice growth. The common deposit type is the yedoma type, and seasonal thawing depth of the active layer in the region varies between years. The Lena River delta is located to the east of the study area (Grosse & Jones 2011; Lupachev & Gubin 2012).

Study area 4 is located in Yakutian lowland in Yakutsk between Koblain district and Namskiy district along Lena River. The annual air temperatures in the Yakutsk region vary from low to very low (Grosse & Jones 2011), for the regional climate is continental and mean temperatures are ranging between 19.5 and −38.6 °C for July and January months. The average annual precipitation is about 230 mm occurring mostly during summer seasons. During wintertime, the average regional snow depth rarely exceeds 60 cm, and typically ranges between 30–40 cm. The soil consists of ice-rich fine-grained loess-like deposits and remnants of previous glaciations (Kim et al. 2019). The vegetation is mostly taiga forest underlain by thick permafrost reaching down to few hundred metres (Iwahana et al. 2005).

Study area 5 is located in Srednekolymskiy district in Kolyma lowland, in Northeastern Siberia, next to Kolyma River. The area is covered by swamps and tundra or forest-tundra vegetation, and it has frequent amounts of low-lying ridges and hills (Augustyn et al. 1998). The landscape has developed under the impact of Holocene thermokarst processes and is filled with thaw-lake depressions (alases) and surface remnants of the Late Pleistocene landscape that is underlain by ice-rich deposits and systems. These types of sediment deposits are commonly referred to as *yedoma.* Nowadays Kolyma is known for having active thermokarst and ongoing thermoerosional transformation processes in its landscape (Veremeeva & Gubin 2009) while also being located 50-70 metres above sea level and having a 3000 km east-west extent (Augustyn et al. 1998).

Study area 6 is located at Chaunsky district in Chukotkas Autonomous State in Eastern Siberia. The study area is located south of Chaunskaya Bay. Typically, the Chukotka region is hilly or mountainous by topography (Grosse & Jones 2011), but the study region in question is located on a more lowlandish area. The climate in the area is severe with coastal characteristics. In the Chukotka region, January mean temperatures range between −15°C and −39°C and July mean temperatures between 5°C and 10°C, but in the study area the oceans influence may affect temperature variations greatly (NaVostok 2020). The region is mostly covered by loess-like sediments with larch forest type of vegetation (Kuzmina et al. 2011).

3.2 North America

Study area 7 is located in North Slope County in Western Alaska, south of Barrow. The area is underlain by permafrost and covered by tundra vegetation, which, according to Rupp et al. (2000), is mostly either treeless tundra or tundra with spruce canopy. Geologically the region is complex, and the study area is surrounded by Brooks mountain range in the south and the Arctic Ocean to the north (Schmidt & Miller 2016). Climate is characterized by long and cold winters with typical growing season expanding from May to September. Precipitation during the growing season counts for 60% of the total annual precipitation. Annually precipitation average is 405 mm and average snowfall being 150 mm, of which 95% falls in May (Rupp et al. 2000).

Study area 8 is located in Yukon Territory, Alaska. The area is located just north of Yukon mountain range and south of the Beaufort Sea. The region is largely mountainous with vast drainage network. Temperatures in the region vary, and summers are relatively short, with small amount of annual average precipitation (250 mm). The climate is relatively cool, and although the area is located mostly above treeline, some areas in the southern valleys are heavily forested. Lakes and wetlands are also common in some regions near the study area (Rea 2019). Permafrost extends to almost 100 metres below ground and is most likely discontinuous in nature (Lininger et al. 2018).

Study area 9 is located in Kitikmeot region in Nunavut, Canada, just to the west of the Kugluktuk village. The temperature range in the area in January average daily up to -30 °C and in July up to 10 °C. Typically, precipitation is low, being annually less than 200 mm. The area is above the tree line and covered in tundra vegetation. Land surface in the region is heavily glaciated, and glacial landforms are common (Rea 2019). The overall climate can be considered to be semi-arid, and the region falls within the subarctic ecoclimatic zone. The study area is located in the rocky outcrop of the ancient Canadian Shield (Ernst & Buddle 2013).

Study area 10 is located in the State of Manitoba in Canada, south coast of Hudson Bay, to the east of Port Nelson in the lowlands. The area is located on the ancient Canadian Shield and covered by tundra vegetation and occasional bogs. The temperatures in Manitoba vary from -40 °C in winters to 38 °C in summers. In the southeast region of Manitoba annual precipitation varies from 360 mm to 560 mm, with about 2/3 of the precipitation coming down between May and September. Typically, snow cover lasts from November to April in the southern areas and even longer in the northern parts where the study area is located (McLintock et al. 2019). The soil consists mostly of glaciolacustrine sediments underlain by discontinuous permafrost. The area is also highly and vastly connected by wetlands containing typic and humid mesisol soils (Metcalfe & Buttle 2001).

4. Methods

4.1 Methodological background

According to several studies, spatial modelling in meso-scale has great potential in periglacial studies because time-consuming data gathering is not necessarily needed, and it is possible to utilize already existing data sources. Research shows that GIS-techniques combined with linear regression models, e.g. explanatory variables derived from DEMs can be useful tools in predicting patterns in periglacial environments using meso-scale approach (Luoto & Hjort 2004, 2005) because it can potentially be more cost-efficient and overlapping, thus making it more accurate. Periglacial process models should be based on variables of causal nature, such as parameters regarding climate and ground properties that directly control periglacial processes, although including them in models might prove problematic (Hjort et al. 2010). Presence-absence models are useful tools in predicting probabilities as indirect indicators of abundance, especially in meso-scale. Even though they are less used their advances as surrogates for geomorphological feature abundances are known (Hjort & Luoto 2008).

4.1.1 Generalized linear models

Generalized linear models or GLMs are mathematical extensions based on linear models (McCullagh & Nelder 1989). In recent years their popularity has increased due to their ability to take nonlinear relationships and statistical distributions into account. Technically, they are related to practices of traditional manner and utilized in modelling of linear relationships and analysis of variance (ANOVA) (Guisan et al. 2002). GLMs have a linear predictor that is related to the mean of the response variable via a link function, which in turn is capable of transforming linearity and maintenance of the predictions within a coherent value range for the response variable. Thus, with GLMs, it is possible to handle Gaussian, Poisson, Binomial or Gamma distributions (Guisan & Zimmermann 2000). Logistic regression is a form of GLM which models probability surfaces of the variable relationships (Luoto & Hjort 2005). It utilizes binomial error structure and a logit

link function, which means that acquiring a positive responses probability is a logistic, s-shaped function when the linear predictor is returned by the model being a first-order polynomial. For second-order polynomials, a bell-shaped function will be approximated. Binary data, usually in the form of presence/absence, can be assessed using logistic regression (Crawley 1993).

4.1.2 Generalized additive models

Generalized additive models or GAMs are semi-parametrical extensions of GLMs (Hastie & Tibshirani 1990). Their major difference to GLMs is that they do not have the problematic phases of suggesting a shape of a response curve or specific function of parametric response and yet are able to maintain the working features of GLMs (Hastie & Tibshirani 1990; Wood & Augustin 2002). GAM too uses a link function similar to that of a GLM with a ”smoothing” function to explanatory variables. Smoothing function does not assume a parametric relationship but rather determines the relationship of the explanatory and response variables (Yee & Mitchell 1991; Guisan et al. 2002), thus minimizing the estimation residual (Marmion et al. 2008). Apart from GLMs, in GAMs some predictors can be modelled nonparametrically. Also, in GAMs it is important to choose a ”smoother” level for the predictor in order to maintain a balance between observations and number of degrees of freedom while fitting the model (Guisan et al. 2002).

4.1.3 Generalized boosting methods

Generalized boosting methods or GBMs are machine learning methods that are especially efficient in data fitting and combining the strengths of different statistical modelling techniques (Mathieu et al. 2008). They are an improvement among predictive modelling techniques and by the ability to combine simple models, they are considered to be more accurate (Ridgeway 1999; Friedman et al. 2000). Generally, they combine decision tree algorithms and boosting methods in order to improve the model accuracy by utilizing a tree classification technique in which estimate residual can be used as an

input to improve the prior classification. Repeating this sequence ultimately decreases estimate residual stepwise (Ridgeway 1999; Thuiller et al. 2006). In modelling, the interaction depth and shrinkage parameters must be specified in order to achieve the optimal predictive ability. GBMs are also able to handle skewed or multi-modal data as well as ordinal or non-ordinal categorical data, which adds to their predictive flexibility (Biodiversity and Climate Change Virtual Laboratory 2015). Boosting methods can be especially useful in predicting periglacial phenomena as shown by e.g. Hjort & Marmion (2009). They concluded that boosting methods showed more reliable results compared to the more traditional statistical modelling techniques and can be especially useful in explorative research.

4.2 Modelling material

The mapping part of the project was conducted using Google Earth Pro version 7.3, where the 10 study areas were chosen and defined using Google Earths *add polygon*-function. The areas were measured similarly, and the areal extents were approximated to be as close to each other as possible. From each study area, the four landforms or processes were visually mapped. These were patterned ground, pingos, thermokarst activity and mass wasting (namely solifluction) processes. Some areas contained high amounts of landforms, for example, thermokarst. Therefore, for time-saving purposes, the areal extents were limited to parts with smaller spatial occurrence of these landforms. The x/y coordinates of the visually mapped landforms were then transferred to an Excel sheet based on landform type, study area number and coordinates of the areas. As pingos were relatively difficult to visually interpret from satellite images, ASTER digital elevation models were used to assist in pingo identification. According to Grosse & Jones (2011), the ASTER imagery shows relief features that are normally hard to spot from aerial/satellite images, and at best, the images could resemble a shaded relief map. In this case, they were used merely as tools in landform identification and not so much as maps *per se*. The ASTER images were obtained from Japans ASTER GDEM Project (Japan Space Systems 2012).

Data modelling included response data (periglacial landforms) and explanatory variables (environmental predictors). The environmental predictors were obtained from different sources based on their availability and hypothesized importance, and before calculating the variable values, all DEMs were projected into the metric coordinate system (Equidistant cylindrical WGS84). ArcMap 10.5 was used to derive data from each study area. Altitude was derived by using Google Earths *add path*- function and transferring data points (as metres) to ArcMap. The elevation points were then transformed to DEM using Spatial Analyst Tools *Topo to Raster*- function. The slope in degrees data was derived using Spatial Analyst tools *slope*- function while input raster being the previously created DEM. Radiation was determined using Spatial Analyst Tools *Area Solar Radiation*- function while using DEM as input raster. Latitude was rounded to closest coordination point, with time configuration chosen as the whole year with monthly interval. Radiation values were displayed as Wh/m^2. Topographic wetness index (TWI) was calculated using a python script (Appendix 4) originally created by Prasad Pathak but later modified by Wolf and Fricker (2013). The modified script converts the terrain slope in degrees to radians and uses the default settings for the flow accumulation raster. As for denominator, a small constant was added to avoid dividing by zero. For all ArcMap variables, Z-factor values were determined based on ESRIs 60 (0.00001792) or 70 (0.00002619) Z-factor values depending on study areas latitudinal location (rounded up value) (Frye 2007). Climate data was derived from WorldClim dataset (Fick & Hijmans 2017). February and July temperature averages (°C) and precipitation averages (mm) were obtained for said months, all in 30-arc second resolution. Mean annual ground temperature (MAGT) data was based on Global Terrestrial Network for Permafrost (GTN-P 2018) dataset and downloaded in a 30-arc second resolution. The MAGT data is derived between 2007-2016. Soil data was derived from ISRIC Soil Data HUB, and it included Absolute depth to bedrock (or soil thickness in cm) and Soil organic carbon content or SOCC (g/kg), and was measured in 200 cm depth and was in 250 m resolution (Table 1). Finally, all DEMs were made to match in cell size using *Resample*- function for coarser 1 x 1 km resolution (30-arc seconds) and then projected back to WGS1984 coordinate system.

Table 1. Geospatial datasets, their description and original resolution.

Data Source	Variable description	Original resolution
WorldClim (1970-2000)	°C average, February	30 arc second
	°C average, July	30 arc second
	Precipitation average mm, February	30 arc second
	Precipitation average mm, July	30 arc second
ISRIC Soil Data Hub (200 cm depth)	Absolute depth to bedrock (in cm)	250 m
	Soil organic carbon content, g/kg	250 m
Arcmap 10.5	Slope gradient in °	30 arc second
	Area solar radiation, (WH/m2)	30 arc second
	Altitude (in metres)	30 arc second
	Topographic wetness index (TWI)	30 arc second
GTN-P (2007-2016)	Mean annual ground temperature, °C	30 arc second

Finally, the observations were combined in an Excel sheet. This was done due to the low number of observations in some study areas and for a better statistical capacity in the modelling phase. The study area observations would then be projected into a bigger geographical region consisting of the circumpolar Arctic and Subarctic (above latitude 55). In order to plot the projected results, the variables were stacked as rasters for the entire circumpolar region similarly utilizing previously mentioned methods with R version 3.5.1 *raster* function. In addition, all false or unmeaningful values were removed from the dataset to achieve more accurate results with the remaining observable values.

4.2.1 Model calibration and evaluation

Calibrating a model is usually based on *p-values* or *Akaike's information criterion (AIC)* (Burnham & Anderson 2002). The link function for calibration was chosen to be logit because it is the most commonly used in binary responses (McCullagh & Nelder 1989). To assess the models´ success of fitting deviance is commonly calculated as:

Explained deviance (D-squared) = (Null deviance – Residual deviance) / Null deviance

In most cases, cross-validation (CV) is used to compare models value prediction abilities. A split-sample approach is a CV strategy, and it is utilized in evaluating and calibrating datasets. A calibration set (70%) was used to build the model and evaluation set (30%) was used to assess the models´ prediction quality (Guisan & Zimmermann 2000). The

evaluation set, however, can not be considered to be fully independent because it was taken from the same dataset as the calibration set (Guisan & Hofer 2003). Also, Spearman correlation was calculated for evaluation and calibration sets for better numerical interpretation (Guisan & Zimmermann 2000). The model calibrations were performed using statistical package R version 3.5.1 with standard *glm, gam* and *gbm* functions. Variable selection was conducted using p-values as criterion. The variables were selected utilizing backward elimination approach, with first selecting all variables and eliminating the less suitable variable from the model. The elimination was based on criterion $p<0.05$, where variables with greater value would be excluded. The p-value was chosen due to low sample sizes in the study areas. The correlation between variables was determined using Spearmans correlation coefficient value of <0.7 to avoid multicollinearity. If the value was higher between two environmental variables other would be removed. For GAMs, the models were build using *mgcv*-package in R, and the same principles were utilized using ”binomial” as a family function, and *k=3* as smoothing line for plots. For GBMs, all available explanatory variables were included in the model because according to Elith et al. (2008), in boosted regression trees (BRT) unimportant variables have minimalistic effect on overall prediction ability. Used distribution factor was ”bernoulli”, and the number of trees (n.trees) was chosen to be 3000.

Evaluation of the generated model is important while building a model (Oreskes et al. 1994). The process concludes fitted response functions realism and explanatory variables, models fit to data, residual characteristics and predictive performance (Ott & Longnecker 2010). For measuring the performance of binary models, there is no one unified method (Guisan & Zimmermann 2000). However, an increasingly popular method for binary classification accuracy is the receiver operating characteristic (ROC) plot and area under the curve (AUC), which was used for model assessment and predictions success in this project. AUC relates relative proportions of cells that are correctly and incorrectly classified over a various range of threshold levels. This makes measuring thresholds independent (Pearce & Ferrier 2000; Guisan & Hofer 2003). According to Swets (1988), AUC measures can be classified as models with 0.50 - 0.70 = low model accuracy, 0.71 - 90 good model accuracy and >0.90 = high model accuracy.

5. Results

5.1 The occurrence of periglacial formations

The total combined number of 906 formations was observed between studied areas, and a total of 205 patterned ground, 220 pingo, 392 thermokarst and 89 solifluction observations were made with all areal observations combined. Study area 1 consisted of 43 patterned ground, 42 pingo, 124 thermokarst (Figure 10) and 13 solifluction observations, of which areally thermokarst proved most common in occurrence out of all areas. Study area 2 consisted of 7 patterned ground, 21 pingo, 5 thermokarst and 7 solifluction observations (Figure 11) in total. Observations in study area 3 consisted of 31 patterned ground, 51 pingo, 28 thermokarst and 19 solifluction observations. Also, study area 3 hosted most of pingo and solifluction observations out of all the studied areas, even though the latter was spread out quite evenly and differences were not so great between single areas. Study area 4 consisted of 10 patterned ground, 13 pingo, 49 thermokarst and 2 solifluction observations. In study area 5, no patterned ground activity was observed. However, it consisted of 10 pingo, 4 thermokarst and 8 solifluction observations.

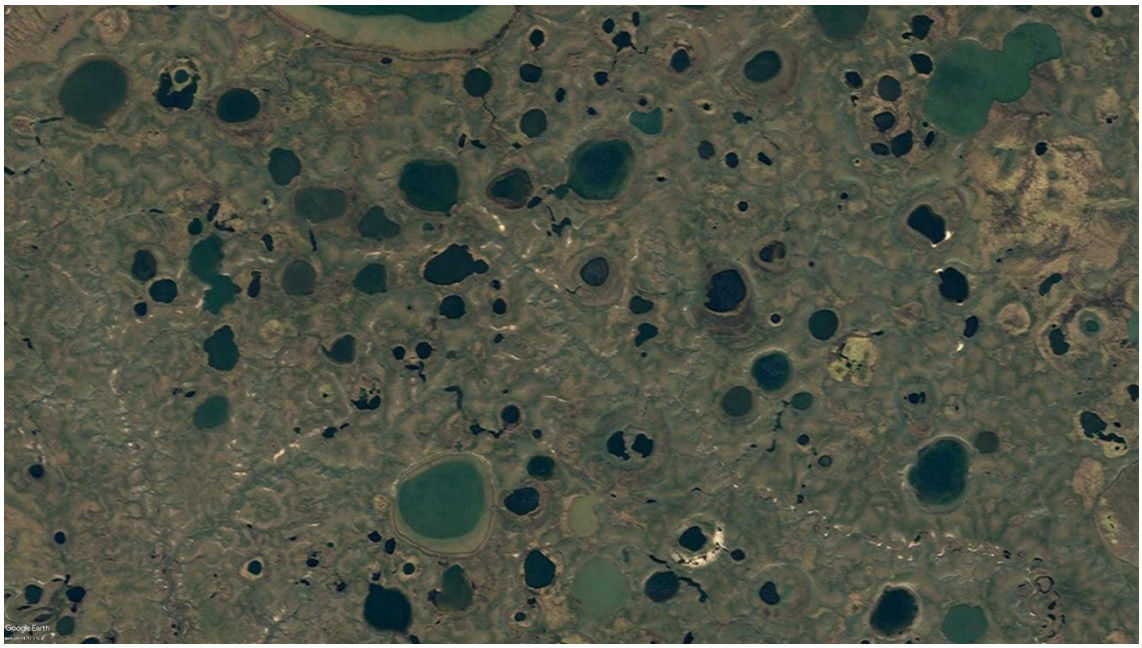


Figure 10. Example of thermokarst landscape observed in study area 1 in Yamalo- Nenets region, Russia. Coord. 69°62´16.40”N, 71°53´14.76”W. Eye altitude 12.65 km (Google Earth Pro n.d.)

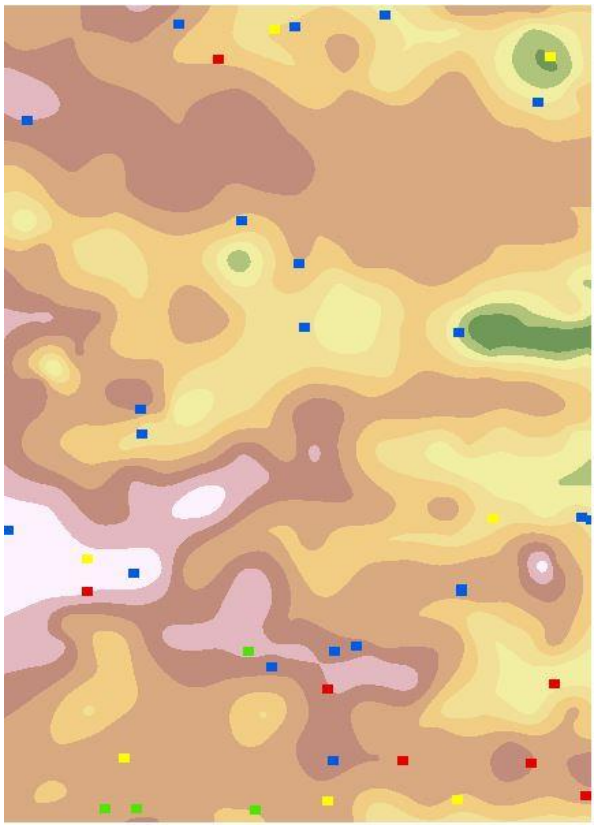


Figure 11. Example of observed landform distribution in study area 2 shown on a contour DEM. Square colors indicate landform type: red = patterned ground, blue= pingo, green = thermokarst, yellow = solifluction (scale 50 x 250 km).

Again, study area 6 had no visual observations of patterned ground activity. It had, however, 14 pingo, 14 thermokarst and 9 solifluction observations. In study area 7, 84 patterned ground observations were made (Figure 11), which was the highest number of patterned ground observations made out of all studied areas. In addition, 19 pingo, 15 thermokarst and 10 solifluction observations were also made. Study area 8 had no observable patterned ground activity but had 34 pingo, 31 thermokarst and 14 solifluction formations observed. Study area 9 had 21 patterned ground, 6 pingo, 67 thermokarst and 3 solifluction observations. Finally, study area 10 had 9 patterned ground, 10 pingo, 55 thermokarst and 4 solifluction observations concluded. The total number of observations are presented fully in Table 2. More landform distribution images (DEMs) are presented in Appendix 5.

Figure 12. Example of observed patterned ground phenomena in the East side of study area 7 in North Slope County, Alaska. Coord. 69°84´16.68”N, -157°17´56.11”W. Eye altitude 776 m (Google Earth Pro 2002)

Figure 13. Example of possible solifluction activity occurring in study area 8 in Yukon, Canada. Coord. 66°03´73.55”N, -137°00´47.72”W. Eye altitude 10.34 km (Google Earth Pro n.d.)

Table 2. Visual observations made in all study areas and their combined total number.

Study Area	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**n =**
Patterned Ground	*43*	*7*	*31*	*10*	*0*	*0*	*84*	*0*	*21*	*9*	***205***
Pingo	*42*	*21*	*51*	*13*	*10*	*14*	*19*	*34*	*6*	*10*	***220***
Thermokarst	*124*	*5*	*28*	*49*	*4*	*14*	*15*	*31*	*67*	*55*	***392***
Solifluction	*13*	*7*	*19*	*2*	*8*	*9*	*10*	*14*	*3*	*4*	***89***

5.2 Patterned ground

After combining landform observation data from each study area, a stepwise backward elimination approach for the variables was used, with criterion $p<0.05$ for response variables using Rs Analysis of variance (ANOVA) that is based on *Chisq*-test for the built GLM models. For GAMs and GBMs, a summary function was utilized. For patterned ground, the most significant variables in GLM models were determined to be February average precipitation, February average temperature, July average precipitation and altitude (Table 3).

Table 3. Environmental variables with most significance to explaining patterned ground occurrence in GLM. Statistical significance is based on Chisq-test (*** p<0.001, ** p<0.01, * p<0.05)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			863	932.55	
precip_2	1	54.845	862	877.70	1.304e-13 ***
temp_2	1	32.645	861	845.06	1.106e-08 ***
precip_7	1	60.311	860	784.75	8.100e-15 ***
altitude	1	106.382	859	678.36	< 2.2e-16 ***

Presented in Table 3 are the environmental variables chosen for model building. All variables were significant. Explained deviance was calculated of the null and residual deviance. ANOVA Chisq-test results are provided in Table 4. Altitude had the lowest residual deviance (678.36), and February average precipitation had the highest (877.70). Null deviance was 932.55. According to ANOVA test, all variables were still significant with values of p<0.001. As a result, the explained deviance was calculated to be 0.272575197, so the variables predicted 27.3% of the variance. All other variables proved non-significant in terms of improving the model, with the calculated variance being highest or close to the highest possible value in terms of variables available. Spearman correlation value for the calibration set was 0.4944452 and for the evaluation set 0.5103888.

For GAM models, the most meaningful variables were February average temperature and average precipitation, July average temperature, soil thickness and altitude (Table 4). Adjusted R-squared was 0.398, and so the model explained 39% of the patterned ground variance. No added or excluded variables would improve the deviance or models performance ability. For the five variables in the model, GAM curves were plotted with a smoothing function of *k=3* and are presented in Appendix 6. Spearman correlation for calibration set was 0.5639558 and for evaluation set 0.6340251.

Table 4. Environmental variables with most significance to explaining patterned ground occurrence in GAM model. Statistical significance is based on backwards elimination approach (*** p<0.001, ** p<0.01, * p<0.05) (n=702).

Approximate significance of smooth terms:				
	edf	Ref.df	Chi.sq	p-value
s(precip_2)	1.942	1.991	78.90	< 2e-16 ***
s(temp_2)	1.423	1.661	66.93	7.99e-09 ***
s(temp_7)	2.000	2.000	50.07	1.38e-11 ***
s(soil_thickness)	1.875	1.984	26.61	4.19e-06 ***
s(altitude)	1.000	1.000	33.02	9.11e-09 ***

For GBM models, all variables were included and the most meaningful ones in order were mean annual ground temperature (MAGT) and July average precipitation and temperature. The least relevant variables were slope, topographic wetness index and altitude having the least influence (Table 5). Variables were plotted into response curves for better interpretation abilities (Appendix 10). Based on the curves' precipitation and temperature, as well as altitude and MAGT seem to have most threshold changes as well as variation in curve shape, whereas other curves seem to be more straightforward. Spearman correlation for calibration set was 0.720738 and for evaluation set 0.7306983.

Table 5. All explanatory variables were chosen for GBM model, where the relative influence of the variable is described.

var	rel.inf
magt	20.207999
precip_7	19.387704
temp_7	16.190698
temp_2	9.026906
precip_2	6.856578
soil_thickness	6.684382
radiation	6.577589
carbon_content	4.639786
altitude	3.703055
twi	3.684994
slope	3.040306

Patterned ground occurrence prediction for GLM, GAM and GBM models were plotted based on chosen environmental variables. For prediction plots, Rs colour package *viridis* was utilized for visual interpretation. GBM plot is depicted in Figure 14. Patterned ground GLM and GAM plots are presented in Appendices 14 and 15.

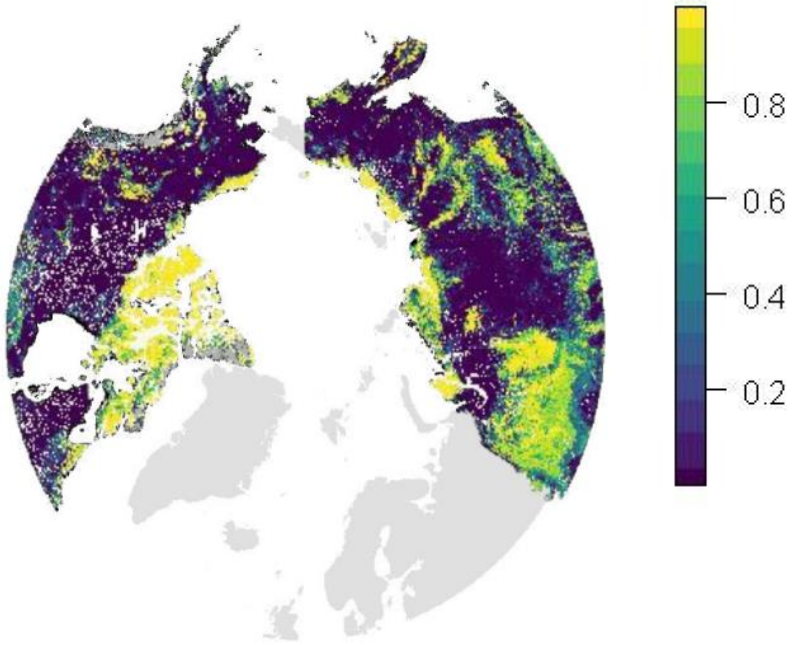


Figure 14. The predicted occurrence of patterned ground in the circumpolar region using GBM plot. The scale indicates the probability of occurrence where 1 = highest probability and 0 = lowest probability.

The model performance was measured using plot ROC curves and AUC values (Figure 15). For patterned ground, GLM models predictive performance AUC was calculated to be 0.7597022, for GAM 0.5156111 and for GBM 0.9398003. Of the three GBM had the best accuracy in predicting patterned ground occurrence with high model accuracy. GAM had low model accuracy, and GLM had relatively good model accuracy.

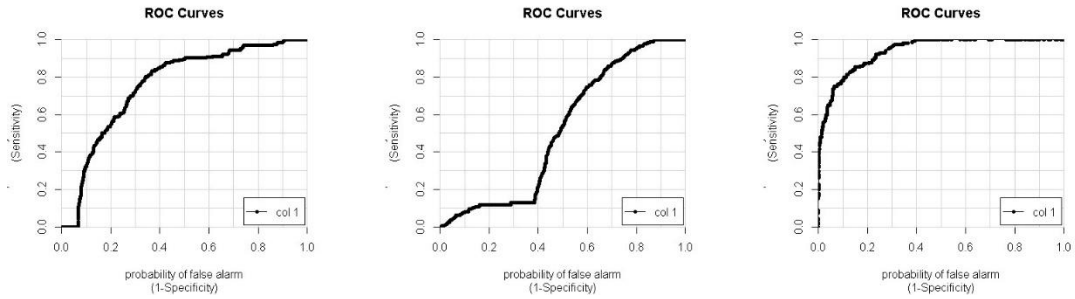


Figure 15. Plotted ROC curves for the three models. GLM (left), GAM (middle) and GBM (right). GBM had the best predictive performance, and GAM had the worst predictive performance.

5.3 Pingos

For pingos, the environmental variables with the most influence in the GLM model were mean annual ground temperature, radiation and altitude. According to Chisq-test, MAGT had the least amount of significance of the three variables in the model (Table 6).

Table 6. Environmental variables with most significance in explaining pingo occurrence in GLM. Statistical significance is based on Chisq-test (*** p<0.001, ** p<0.01, * p<0.05)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			870	966.75	
magt	1	4.268	869	962.48	0.03883 *
radiation	1	32.389	868	930.09	1.262e-08 ***
altitude	1	65.484	867	864.61	5.858e-16 ***

In Table 6 are the environmental variables chosen for model building. Altitude and radiation had values of p<0.001, but MAGT was less significant with p<0.05. Explained deviance was calculated of the Null and residual deviance. Residual deviance was 864.61. Null deviance was 966.75. Variance prediction was calculated based on Null and residual values and was 0.105652961. Thus, the variables predicted 10.6% of the variance. Other variables proved to be insignificant in improving the model or the deviance. Spearman correlation value for calibration set was 0.3235207 and for evaluation set 0.223907. For GAM models most influential variables were February average precipitation, February average temperature, radiation, altitude, July average

temperature and mean annual ground temperature, with July average temperature and MAGT being less significant than the other variables. Adjusted R-squared was 0.216, and the explained deviance was 20.7% of the model. Results are presented in Table 7 and Appendix 7. Spearman correlation for calibration set was 0. 4584252 and for evaluation set 0.348135.

Table 7. Environmental variables with most significance to explaining pingo occurrence for GAM (*** p<0.001, ** p<0.01, * p<0.05).

Approximate significance of smooth terms:				
	edf	Ref.df	Chi.sq	p-value
s(precip_2)	1.000	1.000	33.106	8.72e-09 ***
s(temp_2)	1.958	1.997	16.487	0.000214 ***
s(radiation)	1.939	1.995	55.798	4.89e-13 ***
s(altitude)	1.206	1.367	49.239	1.54e-09 ***
s(temp_7)	1.000	1.000	4.493	0.034032 *
s(magt)	1.000	1.000	6.008	0.014242 *

All variables were included in the GBM model with the most significant ones being July average temperature, mean annual ground temperature and February average temperature, while depth to bedrock, February average precipitation and slope having the least influence (Table 8). Variables were plotted into response curves for better interpretation abilities (Appendix 11). Based on the curves' precipitation averages, TWI and July temperature have rising curve values. Temperature averages and radiation seem to have the highest threshold values with the rest of the response curves having less or no observable pattern in the curve. The predicted occurrence (plots) for the GBM model is presented in Figure 16. GLM and GAM are presented in Appendices 16 and 17. Spearman correlation value for calibration set was 0.7508811 and for evaluation set 0.6967609.

Table 8. All variables for GBM model and their relative influence to the response variable.

var	rel.inf
temp_7	21.478279
magt	14.843308
temp_2	13.469934
twi	8.593495
radiation	8.468741
precip_7	8.406875
altitude	6.129018
carbon_content	5.085380
soil_thickness	4.909848
precip_2	4.755750
slope	3.859371

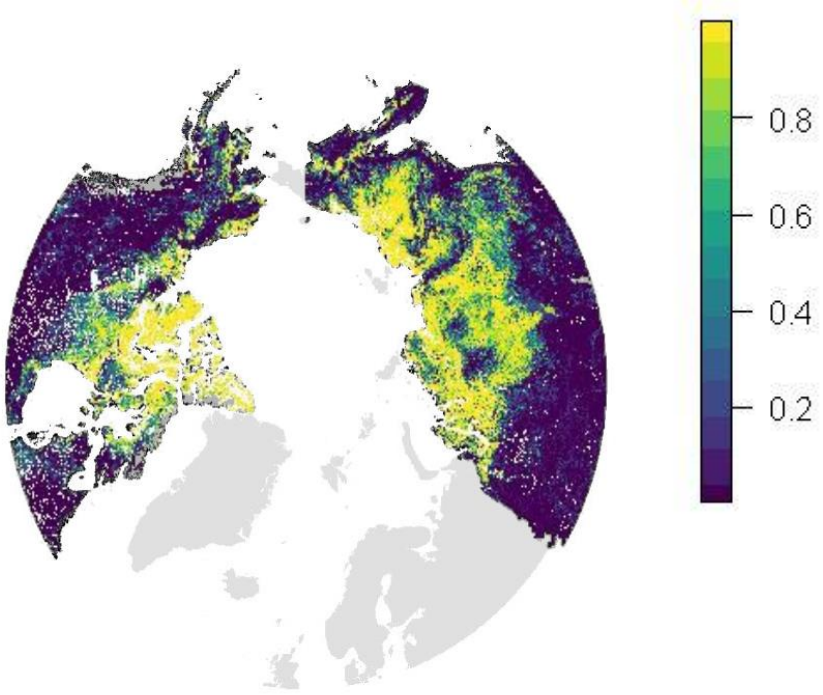


Figure 16. The predicted occurrence of pingo in the circumpolar region using GBM plot. The scale indicates the probability of occurrence where 1 = highest probability and 0 = lowest probability.

The model performance was measured using plot ROC curves and AUC values (Figure 17). For pingo models, AUC values calculated were for GLM 0.6272443, for GAM 0.7044626 and for GBM 0.8638153. GBM had the best value indicating very good accuracy, with GAM indicating good accuracy and GLM poor accuracy.

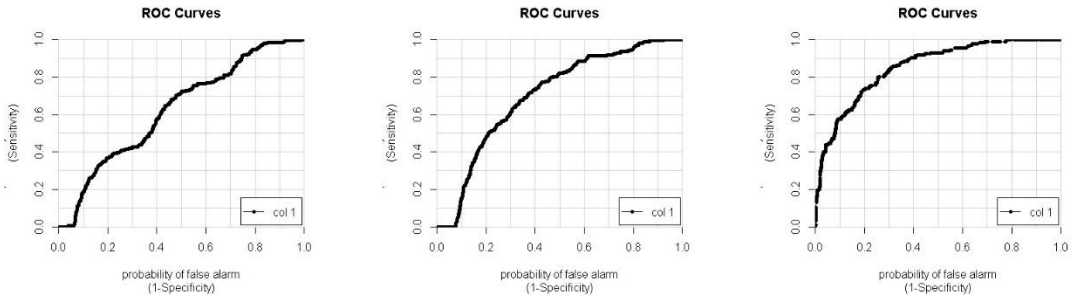


Figure 17. Plotted ROC curves for the three models. GLM (left), GAM (middle) and GBM (right). GBM had the best predictive performance, and GLM had the worst predictive performance.

5.4 Thermokarst activity

For observed thermokarst activity, the environmental variables with most significance for the GLM model were February average precipitation, February average temperature, July average precipitation, soil thickness, MAGT and altitude. February average temperature had the least significance of the variables. The null deviance was 960.46 and residual deviance 760.92. Calculated prediction of variance was 0.2077650293. The GLM model explained 20.8% of the variance (Table 9). Spearman correlation for calibration set was 0.5341432 and for evaluation set 0.5392896.

Table 9. Environmental variables with most significance to explaining thermokarst occurrence in GLM. Statistical significance is based on Chisq-test (*** p<0.001, ** p<0.01, * p<0.05)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			700	960.46	
precip_2	1	15.634	699	944.83	7.686e-05 ***
temp_2	1	8.742	698	936.09	0.00311 **
precip_7	1	33.769	697	902.32	6.205e-09 ***
soil_thickness	1	22.589	696	879.73	2.006e-06 ***
magt	1	55.308	695	824.42	1.030e-13 ***
altitude	1	63.505	694	760.91	1.600e-15 ***

For GAM models, most influential variables were July average precipitation, February average temperature, July average temperature, MAGT and altitude. Less meaningful

were radiation, February average precipitation, soil thickness and SOCC. Adjusted R-squared was 0.366, and the explained deviance was 30.8%. Thermokarst results for GAM are presented in Table 10 and Appendix 8. Spearman correlation for calibration set was 0.6130751 and for evaluation set 0.6075887

Table 10. GAM-values.

Approximate significance of smooth terms:				
	edf	Ref.df	Chi.sq	p-value
s(temp_2)	1.803	1.960	31.935	1.25e-06 ***
s(precip_7)	1.000	1.000	31.955	1.58e-08 ***
s(temp_7)	1.976	1.999	29.090	6.73e-07 ***
s(magt)	1.000	1.000	64.450	9.90e-16 ***
s(altitude)	2.000	2.000	83.195	< 2e-16 ***
s(radiation)	1.000	1.000	8.231	0.00412 **
s(precip_2)	1.618	1.853	11.880	0.00132 **
s(soil_thickness)	1.000	1.000	5.084	0.02414 *
s(carbon_content)	1.000	1.000	5.298	0.02135 *

All variables were included in the GBM model with the most significant ones being February average temperature, July average temperature and precipitation and MAGT. The least significant variables were soil thickness, February average precipitation and SOCC (Table 11). The plotted thermokarst response curves are presented in Appendix 12. The curves show rising threshold values for MAGT, altitude and February average temperature. July average temperature shows lower values with multiple threshold values, but then rising. Other curves are more or less rising, lowering or completely straight neutral curves. Spearman correlation for calibration set was 0.8530076 and for evaluation set 0.8569052. The predicted occurrence for the GBM model is presented in Figure 18. Appendices 18 and 19 show thermokarst GLM and GAM plot results.

Table 11. The relative influence of variables in GBM model.

var	rel.inf
temp_2	17.587950
magt	13.977052
temp_7	12.509902
precip_7	12.017436
radiation	9.203961
twi	8.338871
altitude	7.329460
slope	5.466989
soil_thickness	4.740769
precip_2	4.739541
carbon_content	4.088068

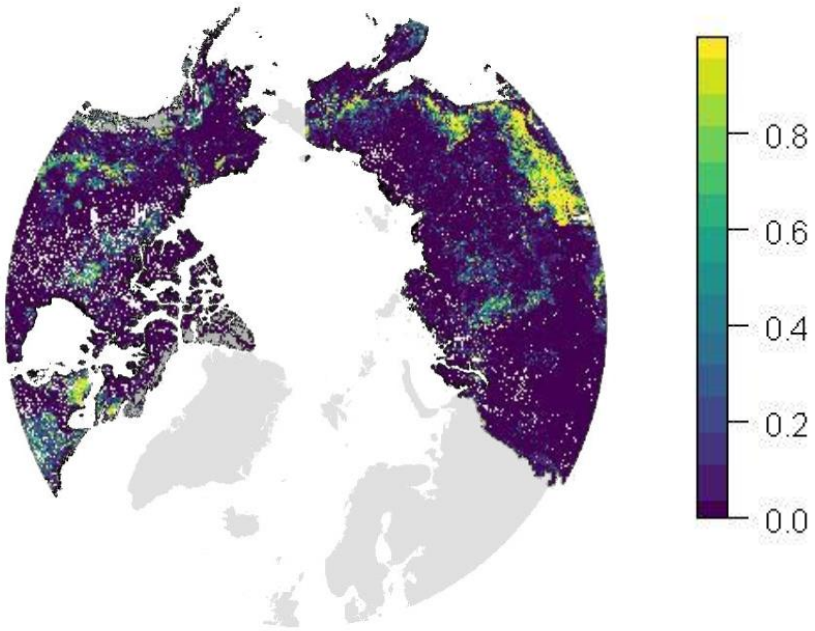


Figure 18. The predicted occurrence of thermokarst activity in the circumpolar region using GBM plot. The scale indicates the probability of occurrence where 1 = highest probability and 0 = lowest probability.

The model accuracy was measured using plot ROC curves and AUC values (Figure 19). For thermokarst models´ calculated AUC values were for GLM 0.5063206, for GAM 0.7044626 and for GBM 0.9200871. GBM had high accuracy, and GAM had good accuracy, although based on GAM value and ROC plot, the accuracy can be considered on the edge of poor. GLM had the lowest accuracy values based on AUC/ROC.

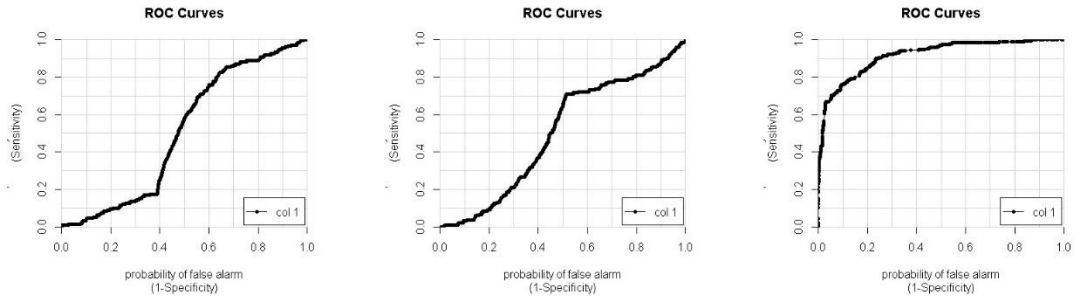


Figure 19. Plotted ROC curves for the three models. GLM (left), GAM (middle) and GBM (right). GBM had the best predictive performance, and GLM had the worst predictive performance.

5.5 Solifluction

For solifluction activity, the observed environmental variables with most significances for GLM model were MAGT, July average precipitation and February average precipitation (Table 12), with MAGT having the least significance according to the Chisq-test. Null deviance for solifluction model was 549.85, and the residual deviance was 515.59. The calculated variance was 0.0623079022, thus the model explaining 6.2% of the variance. Spearman correlation for calibration set was 0.2130171 and for evaluation set 0.1952117, which are both very low.

Table 12. Environmental variables with most significance to explaining thermokarst occurrence in GLM. Statistical significance is based on Chisq-test (*** p<0.001, ** p<0.01, * p<0.05)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			879	549.85	
magt	1	4.5484	878	545.30	0.0329497 *
precip_2	1	17.4385	877	527.87	2.968e-05 ***
precip_7	1	12.2729	876	515.59	0.0004596 ***

For GAM models most influential variables were July average precipitation, February average precipitation, February average temperature and MAGT. Adjusted R-squared was 0.14, and the explained deviance was 18.4%. GAM results are shown in Table 13 and Appendix 9. Spearman correlation for calibration set was 0.3039689 and for evaluation set 0.3143444.

Table 13. GAM results.

Approximate significance of smooth terms:				
	edf	Ref.df	Chi.sq	p-value
s(precip_2)	1.000	1.000	39.98	2.57e-10 ***
s(temp_2)	1.996	2.000	51.22	7.03e-12 ***
s(precip_7)	1.880	1.986	26.16	1.23e-06 ***
s(magt)	2.000	2.000	46.15	9.28e-11 ***

The most significant out of all variables in the GBM model were average temperatures of July and February, with the least significant being July average precipitation and radiation (Table 14). GBM response curves (Appendix 13) were plotted for solifluction activity. The curves show no significant threshold variations, but the curve direction is divided rather evenly amongst variables indicating clear trends in curve shapes. Spearman correlation for calibration set was 0.4915072 and for evaluation set 0.5653801. The predicted occurrence for GBM model is presented in Figure 20. GLM and GAM plot results are presented in Appendices 20 and 21.

Table 14. GBM relative influence of variables for solifluction.

var	rel.inf
temp_2	25.449361
temp_7	15.787882
magt	9.980545
carbon_content	7.698097
twi	7.574244
soil_thickness	6.544351
slope	6.478027
altitude	5.872732
precip_2	5.209459
radiation	4.764869
precip_7	4.640433

.

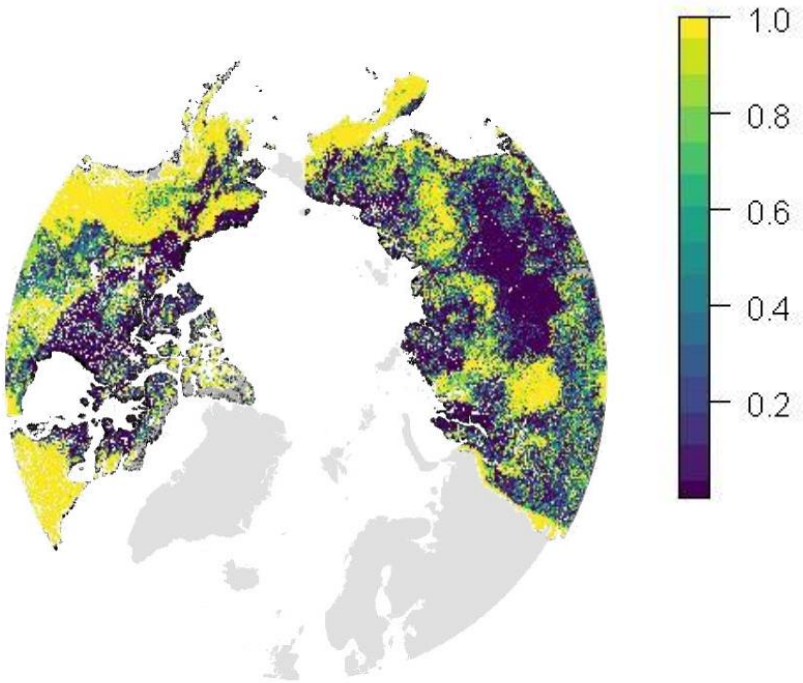


Figure 20. The predicted occurrence of solifluction in the circumpolar region using GBM plot. The scale indicates the probability of occurrence where 1 = highest probability and 0 = lowest probability.

The model performance was measured using plot ROC curves and AUC values (Figure 21). For solifluction models, AUC values calculated were for GLM 0.6347146, for GAM 0.7197613 and for GBM 0.9007605. Once again, GBM shows high AUC values, whereas GAM shows relatively good AUC values. GLM had poor AUC values, with basically random predictive performance.

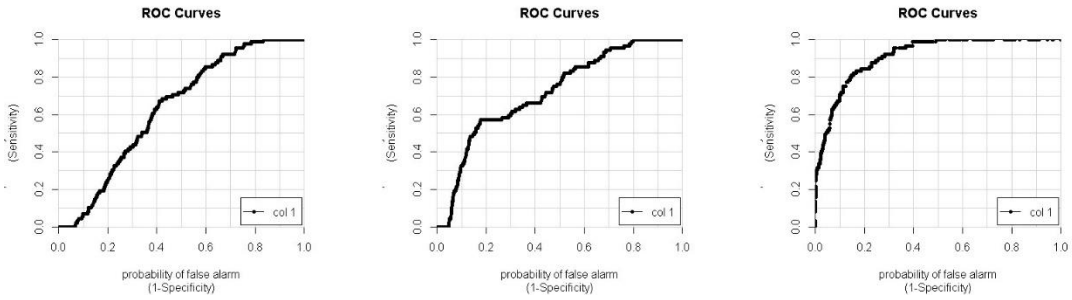


Figure 21. Plotted ROC curves for the three models. GLM (left), GAM (middle) and GBM (right).

6. Discussion

6.1 Mapping observations

Of the mapped landforms thermokarst activity was observed most in the studied areas. Pingos and patterned ground activity were also found plenty whereas solifluction observation occurrence was rarest. The landforms/process extent was vast in some study areas, and therefore some amount of consideration was used to determine the largest and most observable landform types. Therefore, the actual number of landforms might be higher in some areas where visual observations were made. The overall landform distribution and abundance was greatest in areas 1, 3, 4 and 7, and these areas could be considered hotspots for specific landform types.

Thermokarst activity occurred greatly in the northern parts of Northern Asia and especially close to the Arctic ocean. These findings concur with e.g. Olefeldt et al. (2016). Particularly areas 1, 4, 9 and 10 hosted plenty of thermokarst activity. The patterned ground activity was absent in study areas 5, 6 and 8; however, there was plenty in areas 1, 3 and 7. There also seemed to be some sort of relation in observations between thermokarst and patterned ground: study areas where thermokarst activity was greater seemed to host patterned ground activity as well. Pingos were scattered around the study areas and were difficult to identify from other similar landforms. The previous studies, however, suggest (e.g. Grosse & Jones, 2011) that the studied areas host a lot of pingos. This would testify for the relatively high number of pingos in some places, such as areas 1 and 3. As expected *a priori,* plenty of pingos were located in thermokarst areas, with characteristic hydrological conditions, and not many pingos are located close to each other (e.g. Grosse & Jones 2011; French 2018: 160 – 161).

Solifluction activity was not observed as much as other studied landforms, and the rate and size of solifluction activity varied greatly in different areas making the mapping part more challenging. However, higher observed solifluction activity was noted to take place further inland. Even though the differences in solifluction observations were small between areas, solifluction activity was observed to be greatest in areas 1 and 3. These areas also hosted plenty of thermokarst activity indicating uniform hydrological

conditions that would create them. The other important factor in solifluction occurrence is slope gradient, which was small in these areas. This might explain the relatively small number of observations compared to other landform types (e.g. Matsuoka, 2001).

Study areas 1 and 3 had the greatest landform distribution, and study area 1 had the greatest landform abundance. In relative terms, Northern Russia hosts quantitatively more landforms than North America. Especially thermokarst activity in Northern Russia was greater compared to Alaska and regions in Northern Canada (e.g. Jorgenson 2013; Farquharson et al. 2016). Study areas 7 and 8 had greater solifluction occurrence opposed to many regions in North Asia, which can be explained by Alaskan mountain ranges and higher slope gradient, for Northern Russia has relatively flatter surface compared to Alaskan areas. Study area 8 had multiple pingo observations. The area is located close to the famous Tuktoyaktuk Peninsula where pingos have been studied for decades (e.g. Mackay 1998; Olthof et al. 2015), which explains their high abundance. The mapping results are somewhat expected. Those study areas that have a more specific formational activity of landforms, such as greater temperature variations, can be considered to host more landforms and process activity in general.

6.2 Patterned ground

Of the variables used in this study, the most significant ones affecting the occurrence of the response variables were chosen by utilizing a backwards elimination approach. Most meaningful variables in patterned ground (PG) occurrence in GLM were July and February average precipitations, February average temperature and altitude. The patterned ground GLM model explained 27.3% of the variance, which can be considered low. This, however, can be explained by the nature of the study conducted, the binomial variables and the extent of the study area. Spearman correlation for evaluation and calibration sets were not too far from each other, but are rather low, indicating low prediction abilities. Most significant variables in GAM were altitude, soil thickness, February average precipitation and July and February average temperatures. All variables proved to be meaningful in the model. GAM explained 39% of the PG variance,

which is slightly better than with GLM. Spearman correlation values were also higher testifying GAMs better prediction ability.

All environmental variables were utilized in GBMs. For PG, MAGT proved the most meaningful, with July average precipitation and temperature having high significance aswell. Interestingly, the altitude had no notable significance in GBM anymore in contrast to GLM and GAM. It is possible that GBMs more complex modelling capacity and altitudes influence with all variables included is the reason. Spearman correlation for both sets was highest with GBM. AUC values for GBM showed high accuracy with GBM, and moderate with GLM and poor with GAM, also testifying GBMs better predictive performance.

Although some formational aspects of PG still have unanswered questions, it is known that temperature variations are a common nominator for PG occurrence (e.g. Luoto & Hjort 2004). Thus, the most significant variables that were discovered for PG were also expected *a priori*. Temperature and precipitation variables were a common nominator in pretty much all the models. This makes sense since PG formation is known to be highly dependable in the variations of annual temperatures and precipitations (Walker et al. 2008). Both GLM and GAM showed altitude as a meaningful variable, which may be a bit hard to explain in terms of what affects PG occurrence in mesoscale, but it could be concluded that PG occurs where the relative altitude is less than 200 m (as seen in e.g. Jones 2009: 16).

GBM showed MAGT as the most meaningful variable along with July average precipitation and temperature. PG formation is most likely dependent on these variables along with other cryoturbative processes, for it has been shown that PG formation is related to hydrological and vegetational changes that in turn are related to atmospheric variations (Matthews et al. 1998; Walker et al. 2008). Still, changes in MAGT probably had the most significant effect because this study only included temperature data for two months, but MAGT data covered the entire annual average. Also, MAGT is known to have a significant influence on soil properties and ground thermal regime on a larger scale (Karjalainen et al. 2019). Soil thickness was also a meaningful variable in GBMs prediction of PG. It can cause disturbances in water flow, thus enhancing the formative PG processes along with temperature variations, and the thickness of soil can thus affect

PG formation (Peterson & Krantz 2003; Guglielmin et al. 2008). TWI was expected to hold significance, but according to GBM, it was one of the least meaningful variables. Reason for this could be in the nature of the data, or alternatively, errors in data handling process.

Based on the GBM curves, the indication is that altitude variable has no significant effect on PG on higher altitudes. Other variables show significant threshold changes. For example, February average precipitation shows threshold curving downwards but increasing with higher precipitation values. This can be in relation to soils ability to hold water. February average temperature possibly also indicates higher PG occurrence between -30 to -25 °C, and MAGT shows a stepwise drop in lower ground temperatures. It would seem that significant temperature variations or latitude and precipitation changes are affecting PG occurrence the most (e.g. Warburton 2013; Uxa et al. 2017).

Visual interpretation of the plotted maps suggests that GLM barely predicts any occurrence of PG, whereas GAM predicts perhaps too much. Based on mapping experience and quantitative approaches, it is safe to say that of the three models GBMs prediction ability was the best. According to GBM, patterned ground distribution is located mostly within the range of continuous permafrost and in areas with lower topographical gradient. In North America, the occurrence of PG is located mostly in the northern islands. In Northern Russia, the occurrence is highest along the shoreline and in some regions further inland. Near the ocean, the occurrence suggests stronger temperature and precipitation variations that could be the reason for the high probability of occurrence. Higher probability of occurrence more inland could occur because of lower altitudes and variations in MAGT.

6.3 Pingos

Most significant variables explaining pingo occurrence in GLM were radiation and altitude. MAGT was also significant, however, less so than previously mentioned. GAM showed significance with similar variables but in addition with February and July temperatures and February precipitation. In terms of variance, GLM and GAM model

predictions were rather low, which might be the cause of the scarceness of pingo observations, thus affecting models predicting abilities. Spearman correlations for both models were also low, indicating poor prediction capabilities. GBM also showed that both temperature variables and MAGT were the most meaningful when all variables were included in the model.

It can be concluded that higher temperatures (including MAGT) and average precipitation levels improve pingo occurrence in the circumpolar region. Also, higher TWI can affect the occurrence of pingos positively because of its link to thermokarst terrain via formative processes (e.g. Flemal 1976; Grosse & Jones 2011; Wetterich et al. 2018). July average precipitation had the highest threshold spike at value 30, which indicates turnover in precipitation value for maximum pingo occurrence. Higher radiation and TWI, as well as lower altitudes, support pingo occurrence. SOCC and slope gradient have no significant effect on pingo occurrence probabilities although some studies suggest that some (underwater) pingos might have relation to permafrost gas emissions (Paull et al. 2007; Hodson et al. 2019). The graphs generally indicate more probable pingo occurrence in more moist conditions, higher temperatures and in relatively flat surfaces within the permafrost region, which support the theoretical background of important environmental variables affecting pingos and their distribution (e.g. Grosse & Jones 2011; French 2018: 159 – 161).

Plotted maps and AUC values state that GBM had good predictive accuracy being the best of the three, while GLM had the worst predictive accuracy. GLM explained very little of the variance, and with low accuracy, it can be seen in the plotted map as it predicts occurrence almost everywhere. GBM is the most accurate and visually the most reliable. It shows pingo occurrence in places where study observations were made and also other places with known pingo distribution, such as Tuktoyaktuk Peninsula in North America (e.g Mackay 1998), and thermokarst lowlands especially in Northern Russia (Grosse & Jones 2011; Jorgenson 2013). The distribution study conducted by Grosse & Jones (2011) correlates well with the results of this study of pingo occurrence, and also with the actual mapping results of Russia. This is promising and adds more reliability to the outcome of this study. Even though global pingo hotspots are well known the regional distribution and probability of occurrence is important, for under the effects of

Arctic warming decreasing permafrost distribution also means a decrease in pingo numbers overall (e.g. Demidov et al. 2019). These two can be used to monitor each other, and if pingo occurrences are known, more information about permafrosts areal extent there is available.

6.4 Thermokarst activity

Most meaningful variables in thermokarst GLM were February average precipitation, February average temperature, July average precipitation, soil thickness, MAGT and altitude. In GAM, additional meaningful variables were SOCC, radiation and July average temperature. GLM explained low values of thermokarst variance, with pretty poor Spearman correlation values, whereas GAM explained slightly more of the variance while having higher Spearman correlation values. Interestingly, GLM showed significance with soil thickness, but GAM showed only minor significance with soil thickness compared to other variables in the model. Also, the insignificance of TWI in GLM and GAM was surprising, as was SOCC, bearing little or no significance whatsoever, thus contradicting with the results of e.g. Jorgenson (2013) and Olefeldt et al. (2016).

For GBM Spearman correlation values indicated high prediction abilities, but also in GBM, soil thickness and SOCC were one of the least significant variables. It is possible that the increased number of explanatory variables in GAM and GBM reduce the significance of certain variables, and they became more meaningless and were overshadowed by the influence of climate variables. Despite this, the most meaningful variables, July and February average temperatures and MAGT are believable in explaining thermokarst occurrence the best out of all the environmental variables involved. These results would agree with e.g. Jones et al. (2011); Olefeldt et al. (2016) and Farquharson et al. (2016). From the GBM curves, it can be seen that SOCC, soil thickness and slope seem to show no meaningful variation whatsoever. However, average temperatures show interesting ”spikes” in their values which might indicate higher probabilities of occurrence. With these factors taken into consideration, thermokarst seems to have favourable values of occurrence, indicating variation

intensity in the prediction of the occurrence. AUC values show very high prediction accuracy with GBM, good with GAM and poor with GLM.

Visual interpretation of the predicted occurrence proved best with GBM plot. However, it too shows some degree of flaw. Thermokarst activity hotspots, such as Northern Russia, show no probable occurrence of thermokarst activity, but as can be seen from, for example, study area 1 the nearshore region of Northern Russia is regionally thermokarst hotspot. This flaw might occur because of the following reasons: time-saving during mapping part, the shear vastness of thermokarst activity in these regions and study areas mostly chosen further south. This excluded most of the thermokarst activity hotspots from the study areas, although it left thermokarst in northern parts of the region. It is also possible that the model fails to predict thermokarst activity in these areas due to the lack of observations and/or study areas altogether. With more study areas throughout the region, the predicting accuracy would most likely improve. It is also a possibility that current variables fail to predict thermokarst in these regions accurately. With vegetational data added to the mix, the visual prediction accuracy might increase. This being said, the occurrence that the models now predict might still prove to be accurate and precise, and only the extent of the predictions is flawed. There is also different kind of thermokarst activity, and it would seem that this particular model predicts mostly wetland-type thermokarst, as has been shown by e.g. Bouchard et al. (2014); Olefeldt et al. (2016) and Wetterich et al. (2018).

With better existential prediction abilities in mind, perhaps thermokarst types should be separated with different explanatory variables involved so that modelling would prove more productive. Regarding these results, it would perhaps be beneficial that while interpreting the maps visually, it would be in the best interests to keep in mind that it might lack certain ”extensive power” to its prediction ability.

6.5 Solifluction

In solifluction GLM the most significant variables were July and February precipitation averages and MAGT, with MAGT having less significance than the former. The GLM

model explained only 6.2% of the deviance and had very low Spearman correlation values. GAM explained a bit more (18.4%) of the variance with slightly higher Spearman correlation values. Also, February average temperature was added as an additional significant variable in the model, with MAGTs significance also being higher in GAM. GBM had the highest Spearman correlation values of the three models, with both average temperatures and MAGT being the most meaningful variables in the model. The results of thawing activity in terms of temperature variations and MAGT affect the solifluction occurred areas as expected (Matsuoka 2011), which might increase the significance of slope gradient, which in this case stayed rather low. Interestingly, in GBM, the significance of the precipitation averages decreased greatly in contrast to GAM and GLM, which generally contradicts with the results of Matsuoka (2011).

Temperature averages were the most meaningful in GBM while TWI, soil thickness and slope gradient had mediocre influence in the GBM model. Interestingly, the results partly agree and partly disagree with other studies: temperature variations and soil moisture were expected to explain solifluction occurrence (Walker 1986; Matsuoka 2001), but surprisingly slopes significance was much smaller than expected *a priori* (Hjort et al. 2014). This could be explained by the low topographical gradient in the study areas, thus decreasing the significance of slope gradient. The relative topographical gradient can be seen from e.g. Jones et al. (2009: 16). AUC values once again showed the best prediction accuracy for GBM, good for GAM and poor for GLM.

GBM curves indicate clear threshold values of variable behaviour for solifluction activity. In general, it could be argued that solifluction occurs more likely with higher precipitation levels and lower temperatures (e.g. Oliva et al. 2014; Kellerer-Pirklbauer 2017), including lower MAGT, even though GBM showed minor significance in precipitation averages. This is logical considering the nature of solifluction activity. The threshold curves and GAM curves indicate that with higher average precipitation and temperature values probability of occurrence for solifluction increases. It is also interesting to see that SOCC indicates no change but decreases the probability of occurrence after 300 quickly increasing the overall value, thus indicating more probable occurrence in higher SOCC values, perhaps in permafrost regions that hold more organic carbon in sediments (Knowles et al. 2019).

Based on the low number of observations for solifluction and visual interpretation of the plotted predictions as well as low explained deviance, it can be concluded that the models could use improvement. The low explained deviance in GLM and GAM can be explained by the low number of total observations (n=89) because according to Hjort & Marmion (2008), the models should include at least 200 observations for robust predictions. Therefore, a higher number of observations from the study areas would improve the models' prediction accuracy. Also, an addition of a vegetation variable to explain the occurrence would also most likely improve the predictive power of the models since vegetation is known to affect solifluction activity as shown by Matsuoka (2001) and Kinnard & Lewkowicz (2006). GBM plot predicts solifluction occurrence mostly in the regions of higher topography, for example, mountainous areas in Russia and Alaska, which seem fairly logical predictions. However, most likely because of the low number of observations, the predicted extent of occurrence appears rather wide. Therefore, the models might work best as a combination of visual aid with topographical data and more environmental variables and confirmed field observations.

Even though the overall predictive ability of solifluction occurrence might have more uncertainties than other landform predictions, it might provide useful in e.g. studies related to determining boundaries for solifluction activity, for current prediction plot (GBM) shows strong transition between presence-absence of solifluction between areas. GBMs proved to be the most accurate in predicting probabilities for landform occurrences based on studied response variables in the study area. Table 15 summarizes the most meaningful environmental variables for each landform/process type based on the GBM models.

Table 15. The significance of an environmental variable in explaining the occurrence of the response variables based on GBMs relative influence value between variables. Dark green = meaningful variable (> 10.0), lighter green = less meaningful variable (5.0 – 10.0), blank white = not very meaningful variable (< 5.0).

	Patterned ground	Pingos	Thermokarst	Solifluction
°C, February				
°C, July				
Precip, February				
Precip, July				
Soil thickness				
SOCC				
Slope				
Radiation				
Altitude				
TWI				
MAGT				

6.6 Uncertainties

There are some uncertainties that should be taken into account while interpreting or utilizing the results of this study. Firstly, due to the nature of the study, the visual interpretation can be biased, however, conducted as objectively and coherently as possible. Also, there is always some uncertainties when conducting aerial mapping, especially regarding some landforms that might be hard to identify even from ground level (Otto & Smith 2013). Secondly, the low number of observations, especially when projected onto a large study region, such as the circumpolar region, can be ambiguous (Hjort & Marmion 2008; Hjort et al. 2014). The original intention in this study was to model landforms in the ten areas in question, but due to observational deficiency, the nature of the study changed, therefore most likely affecting modelling performance.

Third, slight geographical differences in study area locations and explanatory variables can affect the models' prediction abilities. Because of this, the visual and theoretical interpretation is important (Aalto & Luoto 2014). Fourth, the different temporal and spatial scale of the environmental variables can skew the significance of the variables against the response variables. For more accurate results, the temporal scale of the data

could be altered to match each other so that more accurate predictions can be made (Karjalainen et al. 2019). Some variables were originally derived to match the entire circumpolar region, whereas others were calculated for each study area specifically. This might cause distortion in the explanatory power of the variables.

Lastly, the modelling results could possibly be improved by taking into account the previously mentioned uncertainties and a more ”area oriented” approach because modelling such large regions is computationally very heavy and time-consuming. Therefore, multiple unified smaller scale approaches could give more applicable results for occurrence modelling purposes. Now because of the geographically large study region extent and multiple different landform types studied the available explanatory variables are also very regional and subject-based rather than specified to explain a certain type of variable in a certain geographical area. According to Luoto et al. (2010), recognizing these types of uncertainties in geomorphological mapping should improve overall model accuracies and prediction capabilities as well as improve the planning and management of landscapes in different areas.

7. Conclusions

Out of all the studied periglacial formations thermokarst activity was richest, whereas solifluction abundance occurred the least. In the study areas, pingo and patterned ground activity were also common, with specific areas being more prone to hosting more formations than other. Patterned ground distribution is spread out quite evenly along the circumpolar Arctic, and thermokarst activity was distributed more closer to the sea and was especially rich in Northern Russia. Pingo distribution is located in close contact with already well-known pingo hotspots, and solifluction activity was focused more closely in areas of higher topographical gradient, such as mountainous Alaska.

Using logistic regression, it was determined that climate variables had the most considerable significance in explaining the occurrence of formations. Temperature and

precipitation variables had the greatest significance in explaining the occurrence of studied landforms and processes. Some other variables, such as radiation and MAGT, also had minor influence, but in general February and July average temperatures and precipitation values had the most explanative power out of all chosen variables in the study.

GBM proved to have the best performance and accuracy in predicting the occurrence of landforms and processes along the study region. Plotted maps showed the occurrence and distribution of formations as expected – however, they should not be considered absolutes because of the uncertainties that occurred. The patterned ground formations are most likely to occur in North American islands as well as in the shoreside in Northern Asia. Some areas further south also have a strong probability of PG occurrence, such as Western and Southeastern Russia. Pingo occurrence is most probable in northern parts of Canada and Alaska, with greater variation of occurrence in Russia and North Asia. Thermokarst activity is focused south, with little or no occurrence in the northern parts of Americas and Russia, which generally contradicts with known thermokarst occurrence; however, the current maps explaining its occurrence in other parts of the map (i.e. south of the circumpolar region and other parts of North America) have been known to host thermokarst activity in some form. Solifluction is predicted to occur in areas further from the Arctic Sea, i.e. in areas of a higher topographical gradient. The predicted occurrence, however, is vast, especially in North America, which may be because of a lower number of solifluction observations. The visual interpretation of solifluction occurrence might, therefore, be best used with all plotted maps taken into consideration and with possible further research in mind.

As Luoto & Hjort (2004) have shown, and to a certain degree this thesis has confirmed, mesoscale mapping and modelling of landforms is possible although there is room for improvements. The results of this study could be taken further by conducting a field work assesment of the observed landforms and to confirm the aerial observations. This might suit well for e.g. field research and other mapping projects. By doing so, it is also possible to improve model performances for better understanding of the extents of periglacial processes. The results could also be used to provide background for future studies e.g. studying temporal changes of landform distributions, climate change effects

or changes in permafrost extent. Recently, observation and monitoring projects have taken place in the Arctic, and extent of permafrost and other periglacial landforms have increasingly been studied. This will eventually lead to a better understanding of the consequences of climate warming in the Arctic. This study also gives insight into different environmental variables affecting the periglacial regime and landforms in the circumpolar region, thus providing information on specific factors of landform-variable relationships and extents of cold environment processes for potential mesoscale studies in the future.

References

Aalto, J., le Roux, P. C. & Luoto, M. (2012) The meso-scale drivers of temperature extremes in high-latitude Fennoscandia. *Climate Dynamics*, **42**(1-2), 237–252. doi:10.1007/s00382-012-1590-y

Aalto, J. & Luoto, M. (2014) Integrating climate and local factors for geomorphological distribution models. *Earth Surface Processes and Landforms*, **39**(13), 1729–1740. doi:10.1002/esp.3554

Arbogast, A. (2013) *Discovering physical geography.* Third Edition. 672 p. John Wiley & Sons, Hoboken, NJ.

Atkinson, P., Jiskoot, H., Massari, R. & Murray, T. (1998) Generalized linear modelling in geomorphology. *Earth Surface Processes and Landforms*, **23**(13), 1185–1195. doi:10.1002/(sici)1096-9837(199812)23:13<1185::aid-esp928>3.0.co;2-w

Augustyn, A., Bauer, P., Duignan, B., Eldridge, A., Gregersen, E., McKenna, A., Petruzzello, M., Rafferty, J.P., Ray, M., Rogers, K., Tikkanen, A., Wallenfeldt, J., Zeidan, A. & Zelazko, A. (1998) North Siberian Lowland. Encyclopaedia Britannica, Inc. 5.2.2020. <*https://www.britannica.com/place/North-Siberian-Lowland*>

Augustyn, A., Bauer, P., Duignan, B., Eldridge, A., Gregersen, E., McKenna, A., Petruzzello, M., Rafferty, J.P., Ray, M., Rogers, K., Tikkanen, A., Wallenfeldt, J., Zeidan, A. & Zelazko, A. (2013) Sakha. Encyclopaedia Britannica, Inc. 5.2.2020. <*https://www.britannica.com/place/Sakha-republic-Russia*>

Augustyn, A., Bauer, P., Duignan, B., Eldridge, A., Gregersen, E., McKenna, A., Petruzzello, M., Rafferty, J.P., Ray, M., Rogers, K., Tikkanen, A., Wallenfeldt, J., Zeidan, A. & Zelazko, A. (2019) Yamalo-Nenets. Encyclopaedia Britannica, Inc. 5.2.2020. <*https://www.britannica.com/place/Yamalo-Nenets*>

Barry, R. & Hall-McKim, E. (2018) *Polar environments and global change*. 442 p. Cambridge University Press. New York, NY.

Barsch, D. (1993) Periglacial geomorphology in the 21st century. *Geomorphology*, **7**(1-3), 141–163. doi:10.1016/0169-555x(93)90015-t

Biodiversity and climate change virtual laboratory. (2015) Generalized boosting model. 17.1.2020. <*https://support.bccvl.org.au/support/solutions/articles/6000083212-generalized-boosting-model#header-page0*>

Bishop, M. P., James, L. A., Shroder, J. F. & Walsh, S. J. (2012) Geospatial technologies and digital geomorphological mapping: Concepts, issues and research. *Geomorphology*, **137**(1), 5–26. doi:10.1016/j.geomorph.2011.06.027

Bishop, M. P. (2013) Remote sensing and GIScience in geomorphology: introduction and overview. In Shroder, J. (Editor in Chief), Bishop, M.P. (Ed.), *Treatise on Geomorphology,* pp. 1–24. Academic Press, San Diego, CA, vol. 3. Remote Sensing and GIScience in Geomorphology. doi:10.1016/b978-0-12-374739-6.00040-3

Bouchard, F., Francus, P., Pienitz, R., Laurion, I. & Feyte, S. (2014) Subarctic thermokarst ponds: investigating recent landscape evolution and sediment dynamics in thawed permafrost of Northern Québec (Canada). *Arctic, Antarctic, and Alpine Research*, **46**(1), 251–271. doi:10.1657/1938-4246-46.1.251

Brenning, A., Grasser, M. & Friend, D. A. (2007) Statistical estimation and generalized additive modeling of rock glacier distribution in the San Juan Mountains, Colorado, United States. *Journal of Geophysical Research*. **112**(F2). doi:10.1029/2006jf000528

Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-theoretic Approach*. Second Edition. 488 p. Springer, New York.

Burr, D. M., Tanaka, K. L. & Yoshikawa, K. (2009) Pingos on Earth and Mars. *Planetary and Space Science.* **57**(5-6), 541–555. doi:10.1016/j.pss.2008.11.003

Cowals, D. (1973) Alaska patterned ground 1973. Environmental Protection Agency. National Archives at College Park, U.S. National Archives. 31.1.2020. https//commons.wikimedia.org/wiki/File:Alaska_patterned_ground_1973.jpg

Crawley, M.J. (1993). *GLIM for ecologists*. 379 p. Blackwell Scientific Publications, Oxford.

Demidov, N., Wetterich, S., Verkulich, S., Ekaykin, A., Meyer, H., Anisimov, M., Schirrmeister, L., Demidov, V. & Hodson, A.J. (2019) Pingo development in Grøndalen, West Spitsbergen. *The Cryosphere Discussions*. **13**, 3155–3169, 2019. doi:10,5194/tc-2019-76

Dodds, K. & Nuttall, M. (2019) *The Arctic: what everyone needs to know.* Oxford University Press, New York, NY.

ECMWF. (2019) Sea ice cover for September 2019. Climate Change Service. 18.1.2020. *<https://climate.copernicus.eu/sea-ice-cover-september-2019>*

Elith, J., Leathwick, J. R. & Hastie, T. (2008) A working guide to boosted regression trees. *Journal of Animal Ecology*, **77**(4), 802–813. doi:10.1111/j.1365-2656.2008.01390.x

Encyclopædia Britannica. (n.d.) Diagram explaining the formation of open-system and closed-system pingos. Encyclopaedia Britannica, Inc. 31.1.2020. *<https://www.britannica.com/science/pingo#/media/1/461054/210207>*

Ernst, C. M. & Buddle, C. M. (2013) Seasonal patterns in the structure of epigeic beetle (Coleoptera) assemblages in two subarctic habitats in Nunavut, Canada. *The Canadian Entomologist,* **145**(02), 171–183. doi:10.4039/tce.2012.111

Etzelmüller, B., Ødegård, R. S., Berthling, I. & Sollid, J. L. (2001a) Terrain parameters and remote sensing data in the analysis of permafrost distribution and periglacial processes: principles and examples from southern Norway. *Permafrost and Periglacial Processes*, **12**(1), 79–92. doi:10.1002/ppp.384

Etzelmüller, B., Hoelzle, M., Heggem, E.S.F., Isaksen, K., Mittaz, C., Mühll, D.V., Ødegård, R.S., Haeberli, W. & Sollid, J.L. (2001b) Mapping and modelling the occurrence and distribution of mountain permafrost. *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography*, **55**(4), 186–194. doi:10.1080/00291950152746513

Etzelmüller, B., Heggem, E. S. F., Sharkhuu, N., Frauenfelder, R., Kääb, A. & Goulden, C. (2006) Mountain permafrost distribution modelling using a multi-criteria approach in the Hövsgöl area, northern Mongolia. *Permafrost and Periglacial Processes,* **17**(2), 91–104. doi:10.1002/ppp.554

Farquharson, L. M., Mann, D. H., Grosse, G., Jones, B. M. & Romanovsky, V. E. (2016) Spatial distribution of thermokarst terrain in Arctic Alaska. *Geomorphology*, **273**, 116–133. doi:10.1016/j.geomorph.2016.08.007

Fick, S.E. & Hijmans, R.J. (2017) Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology. *<http://worldclim.org/version2>*

Flemal, R. C. (1976) Pingos and pingo scars: their characteristics, distribution, and utility in reconstructing former permafrost environments. *Quaternary Research,* **6**(01), 37–53. doi:10.1016/0033-5894(76)90039-9

French, H. (2018) *The periglacial environment.* Fourth Edition. 544 p. John Wiley & Sons, Chichester, UK.

Friedman, J., Hastie, T. & Tibshirani, R. (2000) Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors). *The Annals of Statistics*, **28**(2), 337–407. doi:10.1214/aos/1016218223

Fronzek, S., Luoto, M. & Carter, T.R. (2006) Potential effect of climate change on the distribution of palsa mires in subarctic Fennoscandia. *Climate Research,* **32**: 1–12.

Frye, C. (2007) Setting the Z Factor parameter correctly. Imagery & Remote Sensing. ArcGIS Blog, ESRI. 20.1.2020. *<https://www.esri.com/arcgis-blog/products/product/imagery/setting-the-z-factor-parameter-correctly/>*

Google Earth Pro. (2002) North Slope County, Alaska. 69°84´16.68”N, -157°17´56.11”W. Eye alt 776 m. Maxar Technologies 2020. 13.1.2020. *<http://www.google.com/earth/index.html>*

Google Earth Pro. (2015) Circumpolar region, Earth. 69°92´41.81”N, 03°38´92.81”W. Eye alt 13922.87 km. SIO, NOAA, U.S. Navy, NGA, GEBCO. IBCAO, Landsat/Copernicus. 13.1.2020. USGS. *<http://www.google.com/earth/index.html>*

Google Earth Pro. (n.d.) Yamalo-Nenetsk Peninsula, Russia. 69°62´16.40”N, 71°53´14.76”W. Eye alt 12.65 km. Landsat/Copernicus. 13.1.2020. *<http://www.google.com/earth/index.html>*

Google Earth Pro. (n.d.) Yukon, Canada. 66°03´73.55”N, -137°00´47.72”W. Eye alt 10.34 km. Landsat/Copernicus. 13.1.2020. *<http://www.google.com/earth/index.html>*

Grosse, G. & Jones, B. (2011) Spatial distribution of pingos in northern Asia*. Cryosphere,* **5**(1), 13–33. doi:10.5194/tc-5-13-2011

GTN-P. (2018) GTN-P global mean annual ground temperature data for permafrost near the depth of zero annual amplitude (2007-2016)*.* Pangaea. *<https://doi.org/10.1594/PANGAEA.884711>*

Guglielmin, M., Ellis Evans, C. J. & Cannone, N. (2008) Active layer thermal regime under different vegetation conditions in permafrost areas. A case study at Signy Island (Maritime Antarctica). *Geoderma,* **144**(1-2), 73–85. doi:10.1016/j.geoderma.2007.10.010

Guisan, A. & Zimmermann, N. E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**(2-3), 147–186. doi:10.1016/s0304-3800(00)00354-9

Guisan, A., Edwards, T. C. & Hastie, T. (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, **157**(2-3), 89–100. doi:10.1016/s0304-3800(02)00204-1

Guisan, A. & Hofer, U. (2003) Predicting reptile distributions at the mesoscale: relation to climate and topography. *Journal of Biogeography*, **30**(8), 1233–1243. doi:10.1046/j.1365-2699.2003.00914.x

Hastie, T. & Tibshirani, R. (1990) *Generalized Additive Models*. 335 p. Chapman and Hall, London.

Haugland, J. E. (2006) Short-term periglacial processes, vegetation succession, and soil development within sorted patterned ground: Jotunheimen, Norway. *Arctic, Antarctic, and Alpine Research*, **38**(1), 82–89. doi:10.1657/1523-0430(2006)038[0082:sppvsa]2.0.co;2

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M. & Blagotić, A. et al. (2017) Through ISRIC – WDC. SoilGrids250m: Global gridded

soil information based on machine learning. PLoS ONE 12(2): e0169748. doi:10.1371/journal.pone.0169748

Hinzman, L.D., Bettez, N.D., Bolton, W.R. et al. (2005) Evidence and implications of recent climate change in Northern Alaska and other Arctic regions. *Climatic Change* **72**, 251–298. https://doi.org/10.1007/s10584-005-5352-2

Hjort, J. & Luoto, M. (2006) Modelling patterned ground distribution in finnish lapland: an integration of topographical, ground and remote sensing information. *Geografiska Annaler: Series A, Physical Geography*, **88**(1), 19–29. doi:10.1111/j.0435-3676.2006.00280.x

Hjort, J., Luoto, M. & Seppälä, M. (2007) Landscape scale determinants of periglacial features in subarctic Finland: a grid-based modelling approach. *Permafrost and Periglacial Processes*, **18**(2), 115–127. doi:10.1002/ppp.584

Hjort, J. & Luoto, M. (2008) Can abundance of geomorphological features be predicted using presence–absence data? *Earth Surface Processes and Landforms,* **33**(5), 741–750. doi:10.1002/esp.1572

Hjort, J. & Marmion, M. (2008) Effects of sample size on the accuracy of geomorphological models. *Geomorphology,* **102**(3-4), 341–350. doi:10.1016/j.geomorph.2008.04.006

Hjort, J. & Marmion, M. (2009) Periglacial distribution modelling with a boosting method. *Permafrost and Periglacial Processes*, **20**(1), 15–25. doi:10.1002/ppp.629

Hjort, J., Etzelmüller, B. & Tolgensbakk, J. (2010) Effects of scale and data source in periglacial distribution modelling in a high arctic environment, western Svalbard. *Permafrost and Periglacial Processes*, **21**(4), 345–354. doi:10.1002/ppp.705

Hjort, J. & Luoto, M. (2013) Statistical Methods for geomorphic distribution modeling. In Shroder, J. (Editor in Chief), Baas, A.C.W. (Ed.), *Treatise on Geomorphology*, pp. 59–73. Academic Press, San Diego, CA, vol. 2, Quantitative Modeling of Geomorphology. doi:10.1016/B978-0-12-374739-6.00028-2

Hjort, J., Ujanen, J., Parviainen, M., Tolgensbakk, J. & Etzelmüller, B. (2014) Transferability of geomorphological distribution models: Evaluation using solifluction features in subarctic and Arctic regions. *Geomorphology*, **204**, 165–176. doi:10.1016/j.geomorph.2013.08.002

Hoberg, E. P., Galbreath, K. E., Cook, J. A., Kutz, S. J. & Polley, L. (2012) Northern host–parasite assemblages. *Advances in Parasitology*, 1–97. doi:10.1016/b978-0-12-398457-9.00001-9

Hodson, A. J., Nowak, A., Redeker, K. R., Holmlund, E. S., Christiansen, H. H., & Turchyn, A. V. (2019) Seasonal dynamics of methane and carbon dioxide evasion from an open system pingo: lagoon pingo, Svalbard. *Frontiers in Earth Science*, **7**. doi:10.3389/feart.2019.00030

Huggett, R.J. (2011) *Fundamentals of geomorphology.* Third Edition. 536 p. Routledge, London and New York.

Iwahana, G., Machimura, T., Kobayashi, Y., Fedorov, A. N., Konstantinov, P. Y. & Fukuda, M. (2005) Influence of forest clear-cutting on the thermal and hydrological regime of the active layer near Yakutsk, eastern Siberia. *Journal of Geophysical Research: Biogeosciences,* **110**(G2), n/a–n/a. doi:10.1029/2005jg000039

Japan Space Systems. (2012) ASTER Global Digital Elevation Model (ASTER GDEM). The Ministry of Economy, Trade and Industry of Japan (METI) and the National Aeronautics and Space Administration (NASA). *https://ssl.jspacesystems.or.jp/ersdac/GDEM/E/*

Jones, A., Stolbovoy, V., Tarnocai, C., Broll, G., Spaargaren, O. & Montanarella, L. (eds.) (2009) *Soil atlas of the northern circumpolar region*. European Commission, Office for Official Publications of the European Communities, Luxembourg. 142 pp.

Jones, A. (2013) Ibyuk - Canada's tallest pingo - viewed from Cessna 172 - outside Tuktoyaktuk - Northwest Territories – Canada. Viewed under the license CC BY-SA 2.0. 31.1.2020. *https://www.flickr.com/photos/adam_jones/9460732764/*

Jones, B. M., Grosse, G., Arp, C. D., Jones, M. C., Walter Anthony, K. M. & Romanovsky, V. E. (2011). Modern thermokarst lake dynamics in the continuous permafrost

zone, northern Seward Peninsula, Alaska. *Journal of Geophysical Research,* **116**. doi:10.1029/2011jg001666

Jorgenson, M. T. (2013) Thermokarst terrains. In Shroder, J. (Editor inChief), Giardino, R., Harbor, J. (Eds.), *Treatise on Geomorphology*, pp. 313–324. Academic Press, San Diego, CA, vol. 8, Glacial and Periglacial Geomorphology. 10.1016/B978-0-12-374739-6.00215-3

Jorgenson, M. T. & Grosse, G. (2016) Remote sensing of landscape change in permafrost regions. *Permafrost and Periglacial Processes,* **27**(4), 324–338. doi:10.1002/ppp.1914

Karjalainen, O., Aalto, J., Luoto, M., Westermann, S., Romanovsky, V. E., Nelson, F. E. & Etzelmüller, B. (2019) Circumpolar permafrost maps and geohazard indices for near-future infrastructure risk assessments. *Scientific Data*, **6.** doi:10.1038/sdata.2019.37

Kellerer-Pirklbauer, A. (2017) Solifluction rates and environmental controls at local and regional scales in central Austria. *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography,* **72**(1), 37–56. doi:10.1080/00291951.2017.1399164

Kempen, B., Poggio, L., De Sousa, L. et al. (2019) Through ISRIC – WDC. SoilsSoilGrids250m pre-release 2019-12 - Soil organic carbon content. *<https://data.isric.org/geonetwork/srv/fin/catalog.search#/metadata/713396f6-1687-11ea-a7c0-a0481ca9e724>*

Kim, K., Yang, J.-W., Yoon, H., Byun, E., Fedorov, A., Ryu, Y. & Ahn, J. (2019) Greenhouse gas formation in ice wedges at Cyuie, central Yakutia. *Permafrost and Periglacial Processes.* doi:10.1002/ppp.1994

Kinnard, C. & Lewkowicz, A. G. (2006) Frontal advance of turf-banked solifluction lobes, Kluane Range, Yukon Territory, Canada. *Geomorphology,* **73**(3-4), 261–276. doi:10.1016/j.geomorph.2005.06.010

Knowles, J. F., Blanken, P. D., Lawrence, C. R. & Williams, M. W. (2019) Evidence for non-steady-state carbon emissions from snow-scoured alpine tundra. *Nature Communications,* **10**(1). doi:10.1038/s41467-019-09149-2

Kuzmina, S. A., Sher, A. V., Edwards, M. E., Haile, J., Yan, E. V., Kotov, A. V. & Willerslev, E. (2011) The late Pleistocene environment of the Eastern West Beringia based on the principal section at the Main River, Chukotka. *Quaternary Science Reviews*, **30**(17-18), 2091–2106. doi:10.1016/j.quascirev.2010.03.019

Lininger, K. B., Wohl, E. & Rose, J. R. (2018) Geomorphic controls on floodplain soil organic carbon in the Yukon flats, interior Alaska, from reach to river basin scales. *Water Resources Research*, **54**(3), 1934–1951. doi:10.1002/2017wr022042

Luoto, M. & Seppälä, M. (2002) Modelling the distribution of palsas in Finnish Lapland with logistic regression and GIS. *Permafrost and Periglacial Processes,* **13**(1), 17–28. doi:10.1002/ppp.404

Luoto, M. & Hjort, J. (2004) Generalized linear modelling in periglacial studies: terrain parameters and patterned ground. *Permafrost and Periglacial Processes,* **15**(4), 327–338. doi:10.1002/ppp.482

Luoto, M. & Hjort, J. (2005) Evaluation of current statistical approaches for predictive geomorphological mapping. *Geomorphology,* **67**(3-4), 299–315. doiI:10.1016/j.geomorph.2004.10.006

Luoto, M., Marmion, M. & Hjort, J. (2010) Assessing spatial uncertainty in predictive geomorphological mapping: A multi-modelling approach. *Computers & Geosciences,* **36**(3), 355–361. doi:10.1016/j.cageo.2009.07.008

Lupachev, A. V. & Gubin, S. V. (2012) Suprapermafrost organic-accumulative horizons in the tundra cryozems of northern Yakutia. *Eurasian Soil Science*, **45**(1), 45–55. doi:10.1134/s1064229312010115

Mackay, J. R. (1998) Pingo Growth and collapse, Tuktoyaktuk Peninsula Area, Western Arctic Coast, Canada: a long-term field study. *Géographie Physique et Quaternaire,* **52**(3), 271. doi:10.7202/004847ar

Marmion, M., Hjort, J., Thuiller, W. & Luoto, M. (2008) A comparison of predictive methods in modelling the distribution of periglacial landforms in Finnish Lapland. *Earth Surface Processes and Landforms,* **33**(14), 2241–2254. doi:10.1002/esp.1695

Mathieu, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2008) Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions.* **15**(1), 59-69. doi:https://doi.org/10.1111/j.1472-4642.2008.00491.x

Matsuoka, N. (2001) Solifluction rates, processes and landforms: a global review. *Earth-Science Reviews,* **55**(1-2), 107–134. doi:10.1016/s0012-8252(01)00057-5

Matsuoka, N. (2011) Climate and material controls on periglacial soil processes: Toward improving periglacial climate indicators. *Quaternary Research*, **75**(02), 356–365. doi:10.1016/j.yqres.2010.12.014

Matthews, J. A., Shakesby, R. A., Berrisford, M. S. & McEwen, L. J. (1998) Periglacial patterned ground on the Styggedalsbreen glacier foreland, Jotunheimen, southern Norway: micro-topographic, paraglacial and geoecological controls. *Permafrost and Periglacial Processes*, **9**(2), 147–166. doi:10.1002/(sici)1099-1530(199804/06)9:2<147::aid-ppp278>3.0.co;2-9

McCullagh, P. & Nelder, J.A. (1989) *Generalized Linear Models*. Second Edition. 532 p. Chapman and Hall, New York.

McLintock, P., Coates, K.S. & Bumsted, J.M. (2019) Manitoba. Encyclopædia Britannica, Inc. 6.2.2020. <*https://www.britannica.com/place/Manitoba*>

Metcalfe, R. A. & Buttle, J. M. (2001) Soil partitioning and surface store controls on spring runoff from a boreal forest peatland basin in north-central Manitoba, Canada. *Hydrological Processes*, **15**(12), 2305–2324. doi:10.1002/hyp.262

Millar, S. (2013) Mass movement processes in the periglacial environment. In Shroder, J. (Editor in Chief), Giardino, R., Harbor, J. (Eds.), *Treatise on Geomorphology*, pp. 374–391. Academic Press, San Diego, CA, vol. 8, Glacial and Periglacial Geomorphology. doi:10.1016/b978-0-12-374739-6.00217-7

NaVostok. (2020) Chukotka autonomous region. Information portal on the ways of resettlement to the Russian Far East for the Old Believers. 6.2.2020. <*http://navostok.info/en/reg2.php*>

Olefeldt, D., Goswami, S., Grosse, G. et al. (2016) Circumpolar distribution and carbon storage of thermokarst landscapes. *Nature Communications*, **7**, 13043. https://doi.org/10.1038/ncomms13043

Oliva, M., Ortiz, A. G., Franch, F. S. & Catarineu, M. S. (2014) Present-Day Solifluction Processes in the Semi-Arid Range of Sierra Nevada (Spain). *Arctic, Antarctic, and Alpine Research,* **46**(2), 365–370. doi:10.1657/1938-4246-46.2.365

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

Oreskes, N., Shrader-Frechette, K. & Belitz, K. (1994) Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science*, **263**(5147), 641–646. doi:10.1126/science.263.5147.641

Ott, R.L. & Longnecker, M. (2010) An Introduction to Statistical Methods and Data

Analysis. Sixth Edition. 1284 p. Cengage, Belmont.

Otto, J-C. & Smith, M.J. (2013) *Geomorphological mapping.* Clarke, L. (Editor) Geomorphological Techniques (Online Edition). Section 2.6 Geomorphological Mapping, British Society for Geomorphology.

Pareta, K. & Pareta, U. (2015) Geomorphological interpretation through satellite imagery & DEM data. *American Journal of Geophysics, Geochemistry and Geosystems*, **1**(2), pp. 19-36.

Paul, C.K., Ussler, W., Dallimore, S.R. et al. (2007) Origin of pingo-like features on the Beaufort Sea shelf and their possible relationship to decomposing methane gas hydrates. *Geophysical Research Letters*, **34**(1). doi:10.1029/2006gl027977

Pearce, J. & Ferrier, S. (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, **133**(3), 225–245. doi:10.1016/s0304-3800(00)00322-7

Perreault, N., Lévesque, E., Fortier, D., Gratton, D. & Lamarque, L. J. (2017) Remote sensing evaluation of High Arctic wetland depletion following permafrost disturbance by thermo-erosion gullying processes. *Arctic Science*, **3**(2), 237–253. doi:10.1139/as-2016-0047

Peterson, R. A. & Krantz, W. B. (2003) A mechanism for differential frost heave and its implications for patterned-ground formation. *Journal of Glaciology*, **49**(164), 69–80. doi:10.3189/172756503781830854

Phillips, J. D. (2003) Sources of nonlinearity and complexity in geomorphic systems. *Progress in Physical Geography*, **27**(1), 1–23. doi:10.1191/0309133303pp340ra

Phillips, J. D. (2006) Evolutionary geomorphology: thresholds and nonlinearity in landform response to environmental change. *Hydrology and Earth System Sciences,* **10**, 731–742, https://doi.org/10.5194/hess-10-731-2006, 2006.

Post, E., Alley, R. B., Christensen, T.R., Macias-Fauria, M., Forbes, B.C., Gooseff, M.N., Iler, A., Kerby, J.T., Laidre, K.L., Mann, M.E., Olofsson, J., Stroeve, J.C., Ulmer, F., Virginia, R.A. & Wang, M. (2019) The polar regions in a 2°C warmer world. *Science Advances,* **5**(12). doi:10.1126/sciadv.aaw9883

Rea, K.J. (2019) Nunavut. Encyclopædia Britannica, Inc. 6.2.2020. <*https://www.britannica.com/place/Nunavut*>

Rea, K.J. (2019) Yukon. Encyclopædia Britannica, Inc. 6.2.2020. <*https://www.britannica.com/place/Yukon-territory*>

Rekacewicz, P. (2005) Definitions of the Arctic. UNEP/GRID-Arendal. Vital Arctic Graphics. 30.1.2020. <*http://www.grida.no/resources/7010*>

Rekacewicz, P. (2005) Permafrost distribution in the Arctic. GRIP-Arendal. Vital Arctic Graphics. 31.1.2020. <*http://www.grida.no/resources/7000*>

Ridgeway, G. (1999) The state of boosting. *Computing Sciences and Statistics,* **31**, 172–181.

Rowley, T., Giardino, J. R., Granados-Aguilar, R. & Vitek, J. D. (2015) Periglacial processes and landforms in the critical zone. *Developments in Earth Surface Processes*, **19**, 397–447. doi:10.1016/b978-0-444-63369-9.00013-6

Rupp, T. S., Chapin, F. S. & Starfield, A. M. (2000) Response of subarctic vegetation to transient climatic change on the Seward Peninsula in north-west Alaska. *Global Change Biology*, **6**(5), 541–555. doi:10.1046/j.1365-2486.2000.00337.x

Shiklomanov, A. I., Lammers, R. B., Lettenmaier, D. P., Polischuk, Y. M., Savichev, O. G., Smith, L. C. & Chernokulsky, A. V. (2012) Hydrological changes: historical analysis, contemporary status, and future projections. In Groisman P., Gutman G. (eds). *Regional Environmental Changes in Siberia and Their Global Consequences,* p. 111-154. Springer Environmental Science and Engineering. Springer, Dordrecht. doi:10.1007/978-94-007-4569-8_4

Schmidt, R.A.M. & Miller, M.M. (2016) Alaskan mountains. Encyclopædia Britannica, Inc. 6.2.2020. *https://www.britannica.com/place/Alaskan-Mountains*

Swets, J. (1988) Measuring the accuracy of diagnostic systems. *Science*, **240**(4857), 1285–1293. doi:10.1126/science.3287615

Thuiller, W., Midgley, G. F, Rouget, M. & Cowling, R.M. (2006) Predicting patterns of plant species richness in megadiverse South Africa. *Ecography,* **29**(5), 733–744. doi:10.1111/j.0906-7590.2006.04674.x

Uxa, T., Mida, P. & Křížek, M. (2017) Effect of Climate on Morphology and Development of Sorted Circles and Polygons. *Permafrost and Periglacial Processes,* **28**(4), 663–674. doi:10.1002/ppp.1949

Veremeeva, A. & Gubin, S. (2009) Modern tundra landscapes of the Kolyma Lowland and their evolution in the Holocene. *Permafrost and Periglacial Processes,* **20**(4), 399–406. doi:10.1002/ppp.674

Walker, H.J. (1986). Periglacial environments. In: Fookes, P.G., Vaughan, P.R. (Eds.). A *Handbook of Engineering Geomorphology*. 82–96 p. Chapman and Hall, New York, NY.

Walker, D. A., et al. (2008) Arctic patterned-ground ecosystems: A synthesis of field studies and models along a North American Arctic Transect. *Journal of Geophysical Research,* **113**(G3). doi:10.1029/2007jg000504

Warburton, J. (2013) Patterned ground and polygons. In: Shroder, J. (Editor in Chief), Giardino, R., Harbor, J. (Eds.). *Treatise on Geomorphology*, pp. 298–312. Academic Press, San Diego, CA, vol. 8, Glacial and Periglacial Geomorphology.

Westermann, S., Østby, T. I., Gisnås, K., Schuler, T. V. & Etzelmüller, B. (2015*)* A ground temperature map of the North Atlantic permafrost region based on remote sensing and reanalysis data. *The Cryosphere*, **9**(3), 1303–1319. doi:10.5194/tc-9-1303-2015

Wetterich, S., Schirrmeister, L., Nazarova, L., Palagushkina, O., Bobrov, A., Pogosyan, L., Savelieva, L., Syrykh, L., Matthes, H., Fritz, M., Günther, F., Opel, T. & Meyer, H. (2018) Holocene thermokarst and pingo development in the Kolyma Lowland (NE Siberia). *Permafrost and Periglacial Processes*, **29**(3), 182–198. doi:10.1002/ppp.1979

Williams, P.J. & Smith, M.W. (1991) *The frozen earth: fundamentals of geocryology.* Cambridge University Press, New York.

Wolf, J. & Fricker, A. (2013) A revised version of the topographic wetness index written for arcpython. *https://github.com/africker/Topographic-Wetness-Index/blob/master/README.md*

Wood, S. N. & Augustin, N. H. (2002) GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecological Modelling*, **157**(2-3), 157–177. doi:10.1016/s0304-3800(02)00193

Yang, M., Nelson, F. E., Shiklomanov, N. I., Guo, D. & Wan, G. (2010) Permafrost degradation and its environmental effects on the Tibetan Plateau: A review of recent research. *Earth-Science Reviews*, **103**(1-2), 31–44. doi:10.1016/j.earscirev.2010.07.002

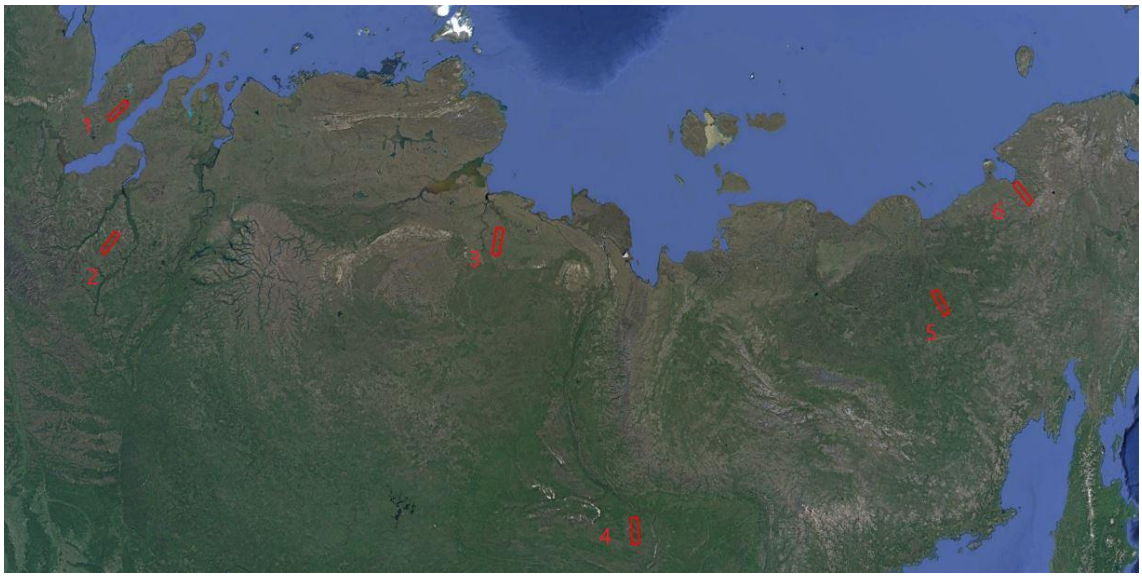
Yee, T. W. & Mitchell, N. D. (1991) Generalized additive models in plant ecology. *Journal of Vegetation Science*, **2**(5), 587–602. doi:10.2307/3236170

Appendices

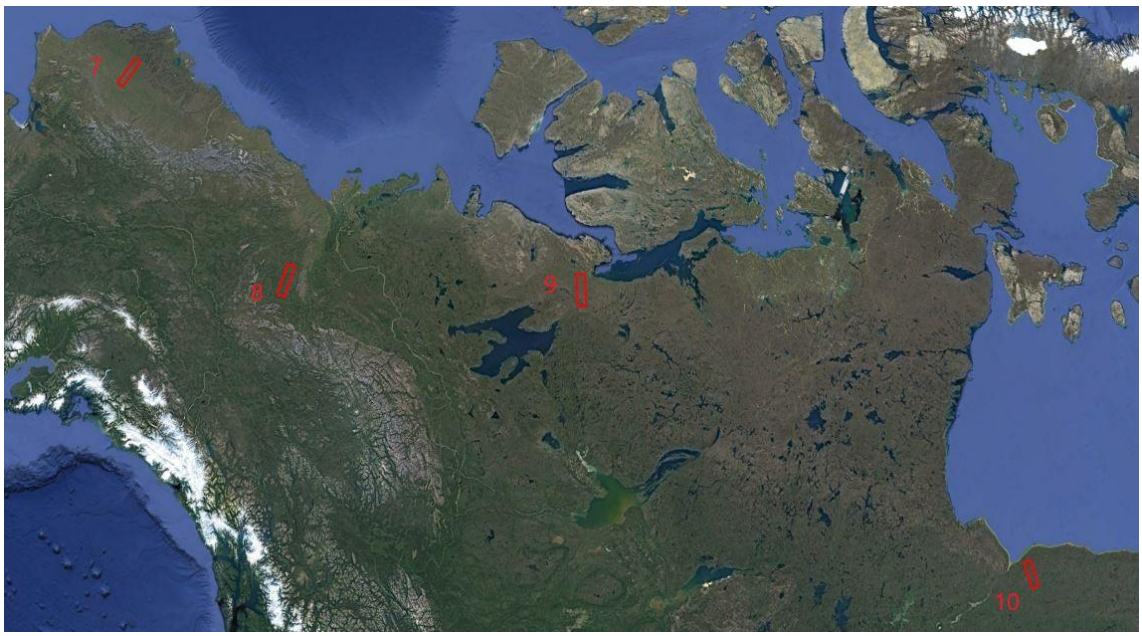
Appendix 1. Study area coordinates based on Google Earth Pro data and WGS 1984 Major Auxiliary Sphere datum.

Study Areal Coordinates		
Study area	Latitude	Longitude
1	69.719068°	71.091522°
	69.720064°	71.771145°
	68.827749°	71.765572°
	68.833150°	71.138827°
2	66.453629°	79.748851°
	66.460072°	80.346490°
	65.567187°	80.360323°
	65.566630°	79.818098°
3	72.539921°	115.110327°
	72.539339°	115.884835°
	71.649936°	115.869187°
	71.647864°	115.144286°
4	63.279186°	128.075937°
	63.280386°	128.600044°
	62.382664°	128.594374°
	62.382815°	128.107565°
5	67.605416°	156.598595°
	67.607849°	157.193009°
	66.711496°	157.211597°
	66.710488°	156.631113°
6	68.765434°	169.261848°
	68.767193°	169.890839°
	67.863479°	169.851925°
	67.865827°	169.255127°
7	69.973259°	-157.345141°
	69.977002°	-157.345141°
	69.081688°	-156.688868°
	69.078500°	-157.293094°
8	66.882662°	-137.454076°
	66.882674°	-136.871182°
	65.977021°	-136.891504°
	65.975062°	-137.446073°
9	67.925586°	-116.442665°
	67.931354°	-115.814137°
	67.039240°	-115.769363°
	67.038693°	-116.355347°
10	56.962513°	-91.714921°
	56.964496°	-91.283008°
	56.066350°	-91.266679°
	56.066485°	-91.670719°

Appendix 2. More precise image of study area locations in North Asia. Eye alt 3907.14 km. Google Earth Pro, Data SIO, NOAA, U.S. Navy, NGA, GEBCO. Landsat/Copernicus, IBCAO.



Appendix 3. More precise image of study area locations in North America. Eye alt 3566.01 km. Google Earth Pro, Data SIO, NOAA, U.S. Navy, NGA, GEBCO. Landsat/Copernicus, IBCAO, U.S. Geological Survey.

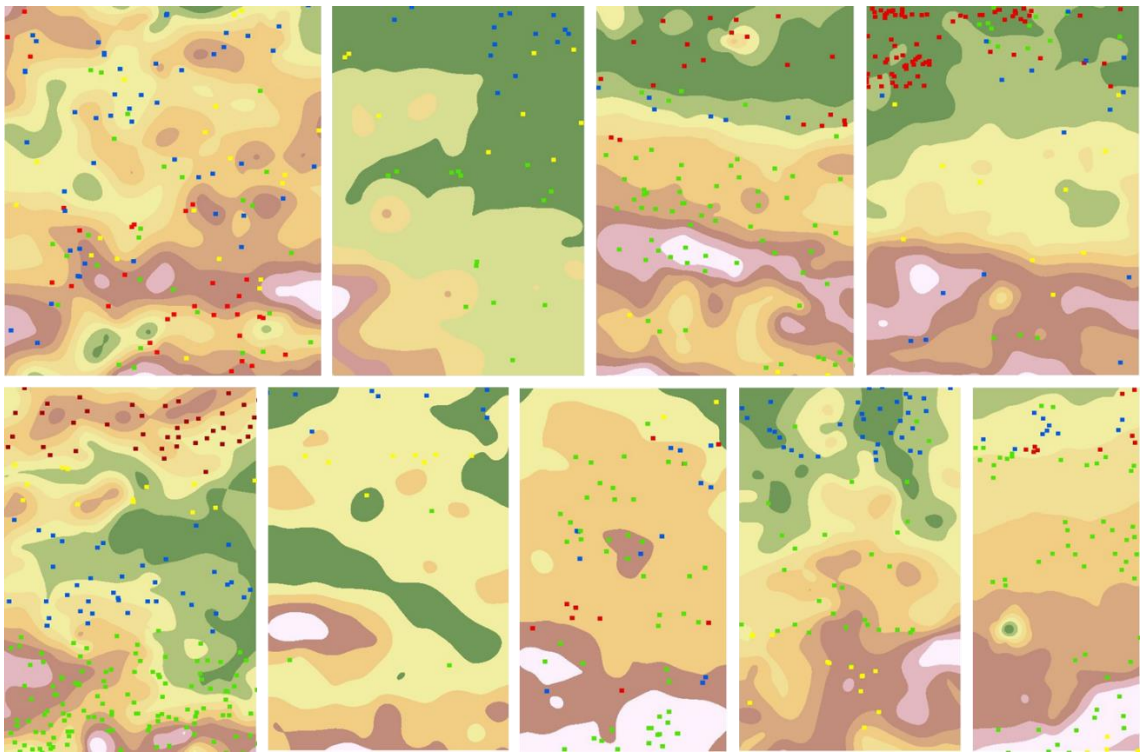


Appendix 4. Python Script used to calculate Topographic Wetness Index (TWI)

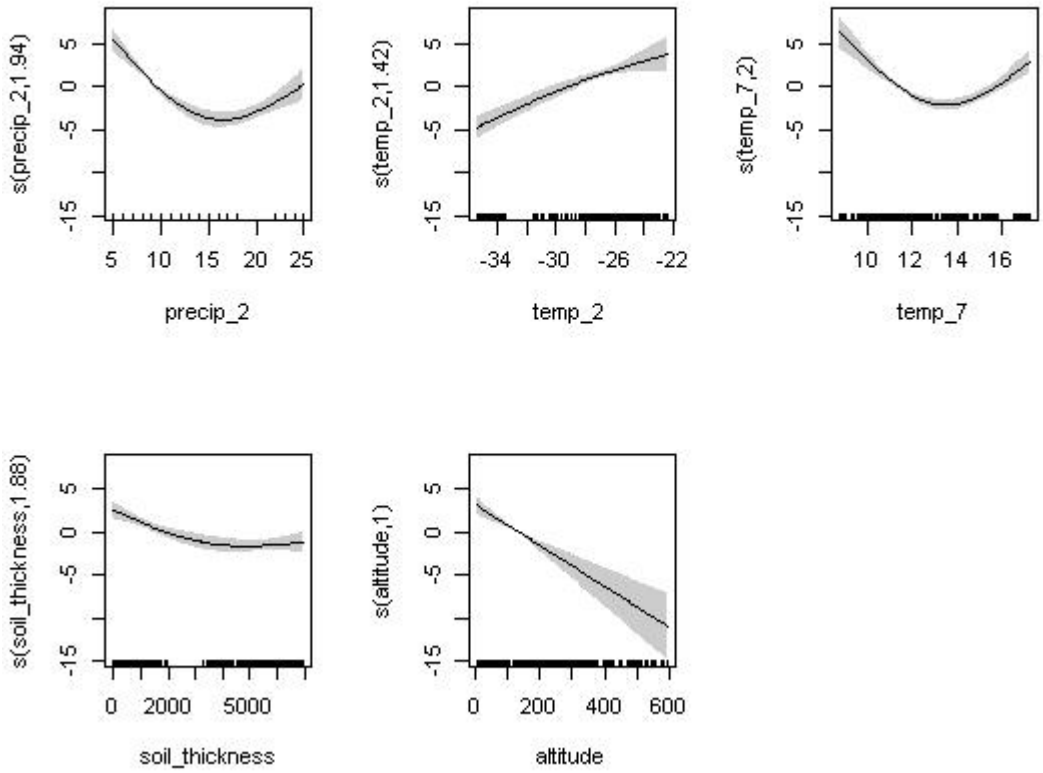
```
import arcpy, math

if __name__ == '__main__':
        arcpy.CheckOutExtension("Spatial")
        # Define workspace and set input and output files
        arcpy.env.workspace = arcpy.GetParameterAsText(0)
        inDEM = arcpy.GetParameterAsText(1)
        outTWI = arcpy.GetParameterAsText(2)
        # Intermediates
        arcpy.AddMessage("Filling DEM.\n")
        DEM_filled = arcpy.sa.Fill(inDEM)
        arcpy.AddMessage("Creating flow direction.\n")
        outFlowDirection = arcpy.sa.FlowDirection(DEM_filled, "FORCE")
        arcpy.AddMessage("Creating flow accumulation.\n")
        #outFlowAccumulation = arcpy.sa.FlowAccumulation(outFlowDirection, "",
"FLOAT") + 1
        outFlowAccumulation = arcpy.sa.FlowAccumulation(outFlowDirection, "",
"INTEGER") + 1
        arcpy.AddMessage("Creating slope.\n")
        slope = arcpy.sa.Slope(DEM_filled)
        arcpy.AddMessage("Converting slope in degrees to slope in radians")
        # 2Pi radians = 360 degrees
        # Pi radians = 180 degrees
        # conversion: Pi radians/180 degress
        slope_radians = slope * math.pi/180.0
        # Output
        arcpy.AddMessage("Creating TWI\n")
        TWI = arcpy.sa.Ln(outFlowAccumulation / (arcpy.sa.Tan(slope_radians)+.01))
        TWI.save(outTWI)
        arcpy.AddMessage("Saved TWI. Done.")
```

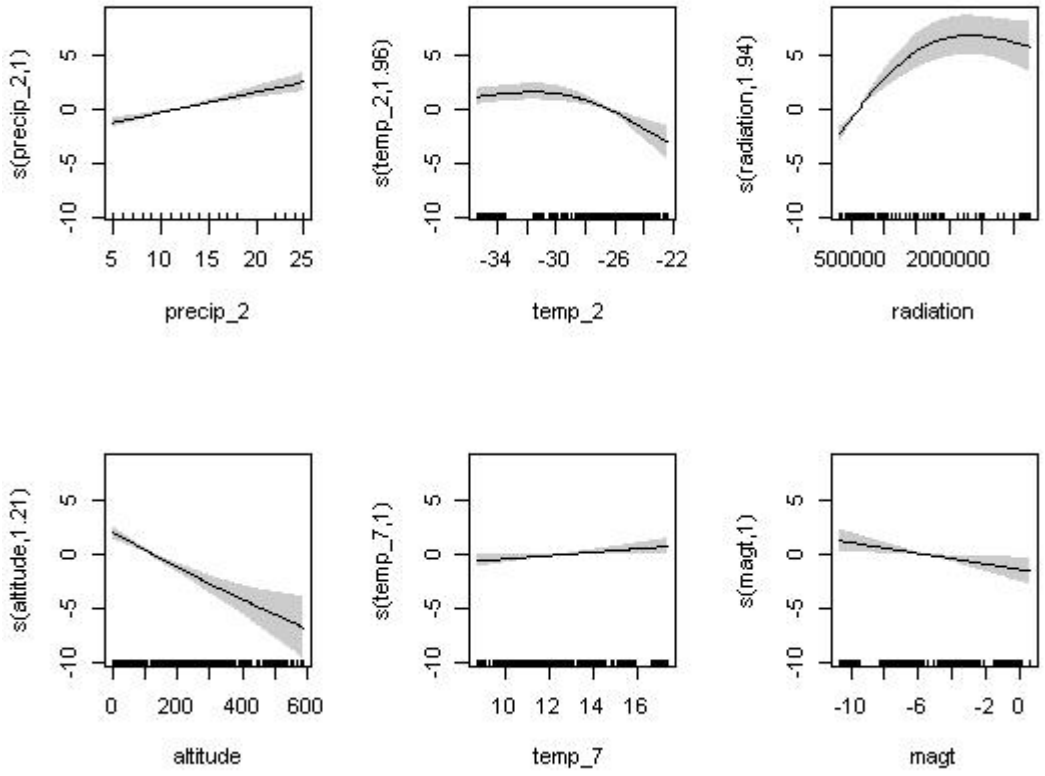
Appendix 5. Examples of landform and process distributions observed in study areas and projected onto DEM. From top left: area 3, area 6, area 9, area 7. From bottom left: area 1, area 5, area 4, area 8 and area 10. Square colors indicate landform type: red = patterned ground, blue= pingo, green = thermokarst, yellow = solifluction. Scale of all study areas is approximately 50 x 250 km.



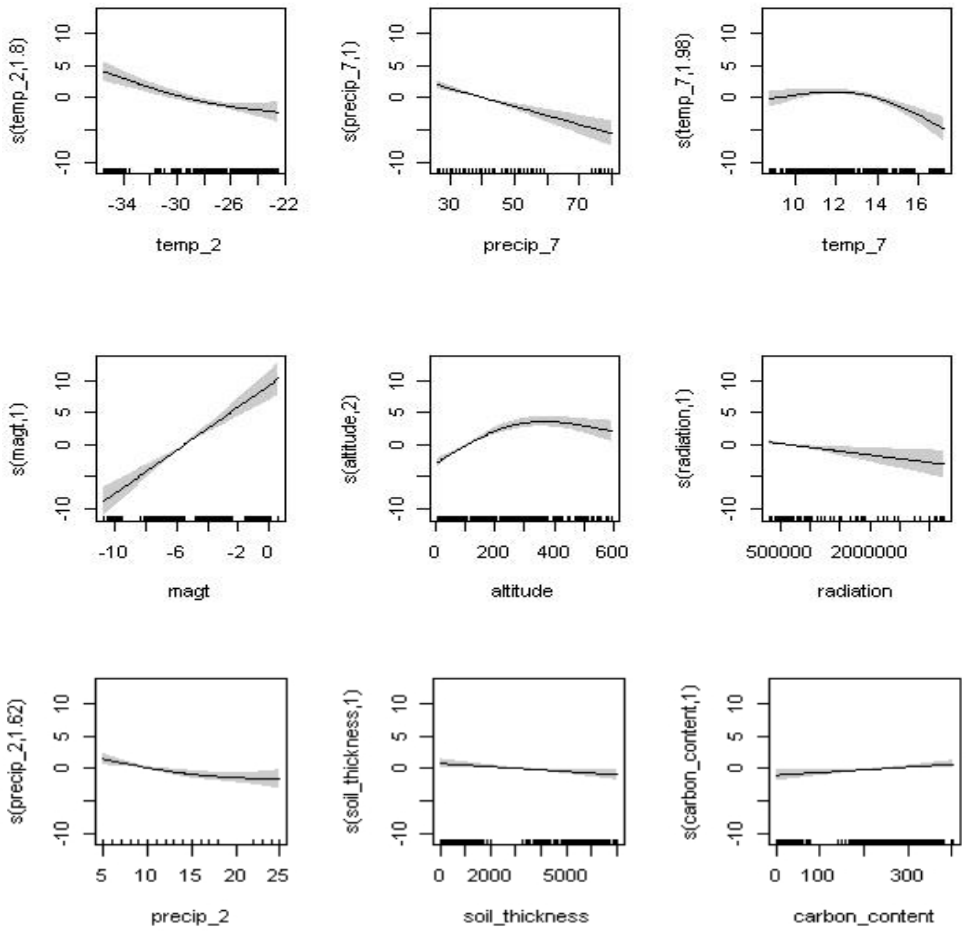
Appendix 6. GAM curves for patterned ground. Y-axis shows the smoothing function values and the environmental variables in the model.



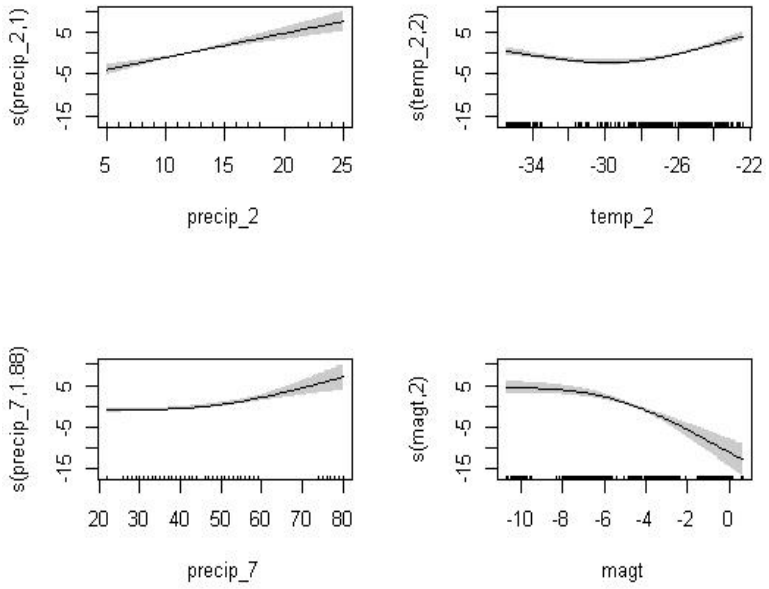
Appendix 7. GAM curves for pingos. Y-axis shows the smoothing function versus the environmental variables in the model.



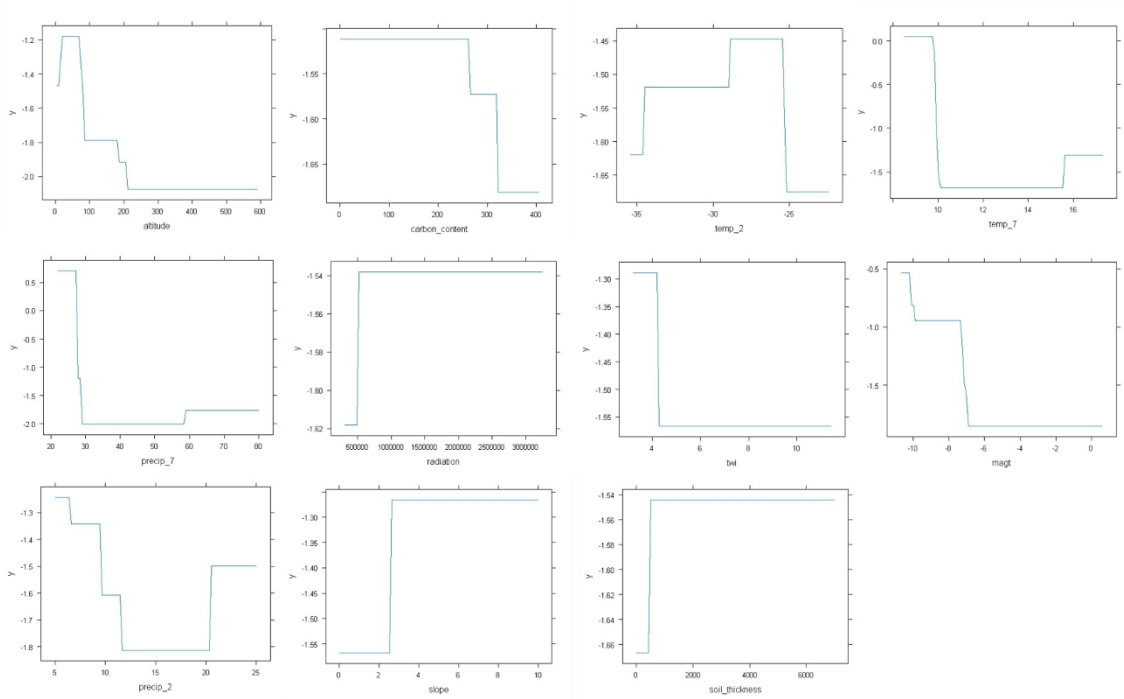
Appendix 8. GAM curves for thermokarst activity. The Y-axis shows the smoothing function versus the environmental variables in the model.



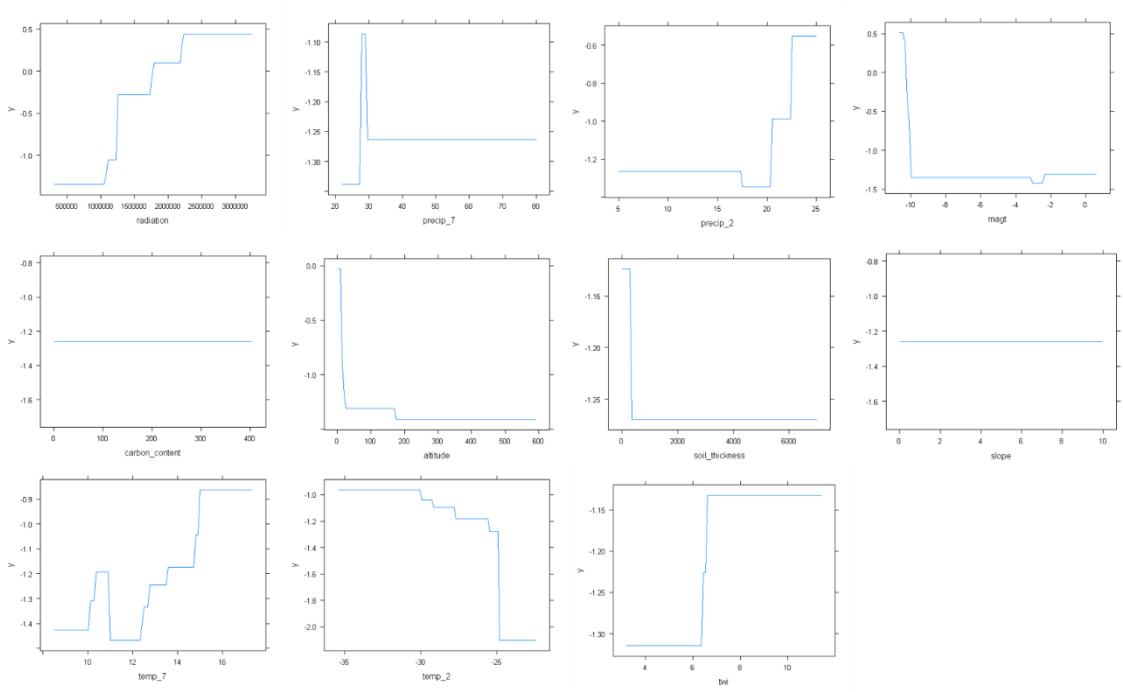
Appendix 9. GAM curves for solifluction activity. The Y-axis shows the smoothing function versus the environmental variables in the model.



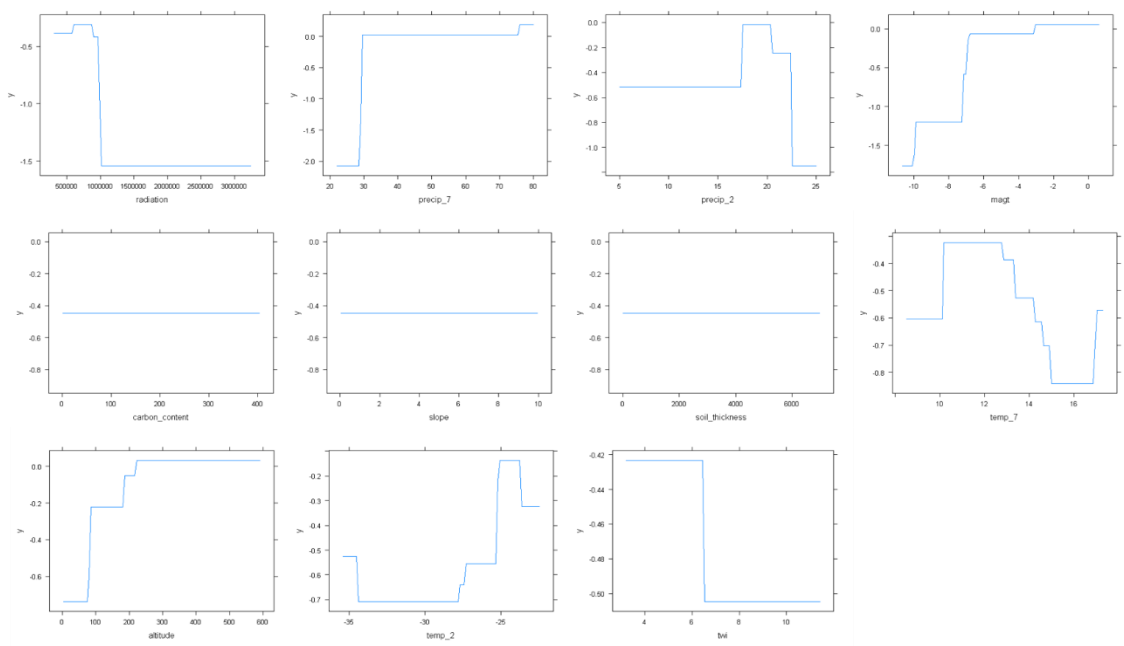
Appendix 10. Patterned ground GBM response curves. X-axis shows the predictor variable and y-axis the modelled values of the response variable plotted against an even distribution of the predictor.



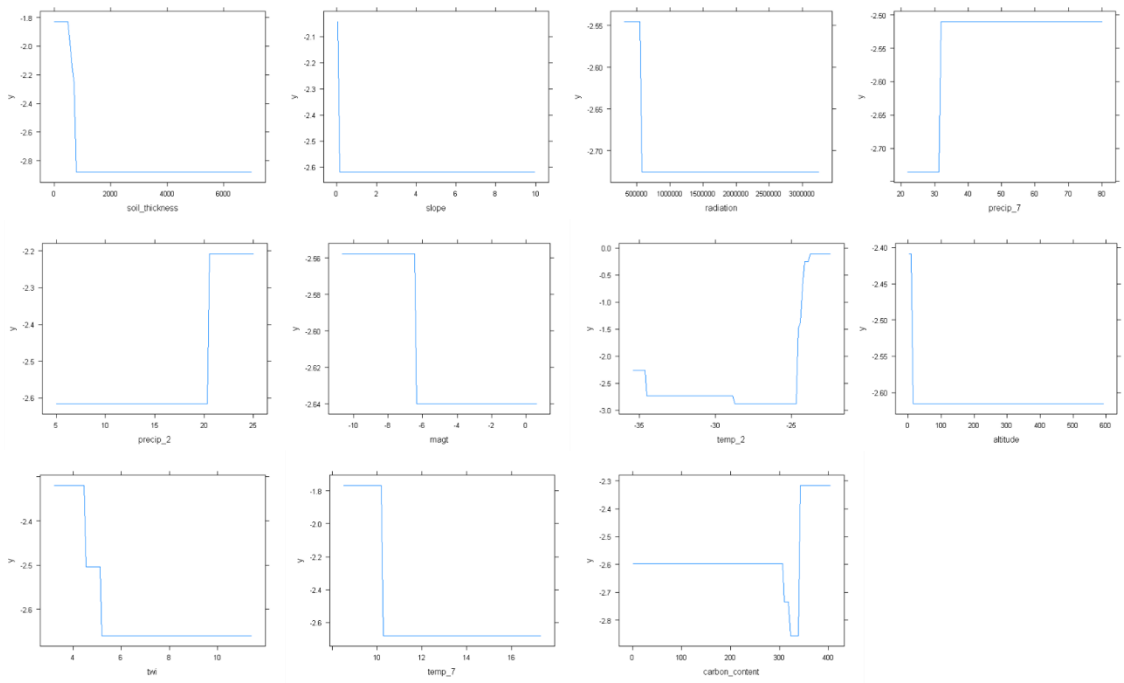
Appendix 11. Pingo GBM response curves. The X-axis shows the predictor variable and y-axis the modelled values of the response variable plotted against an even distribution of the predictor.



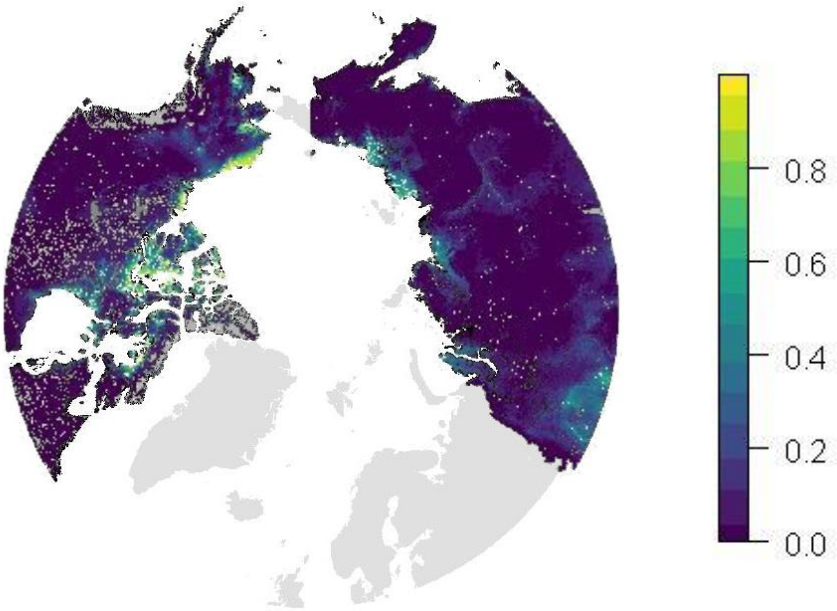
Appendix 12. Thermokarst activity GBM response curves.



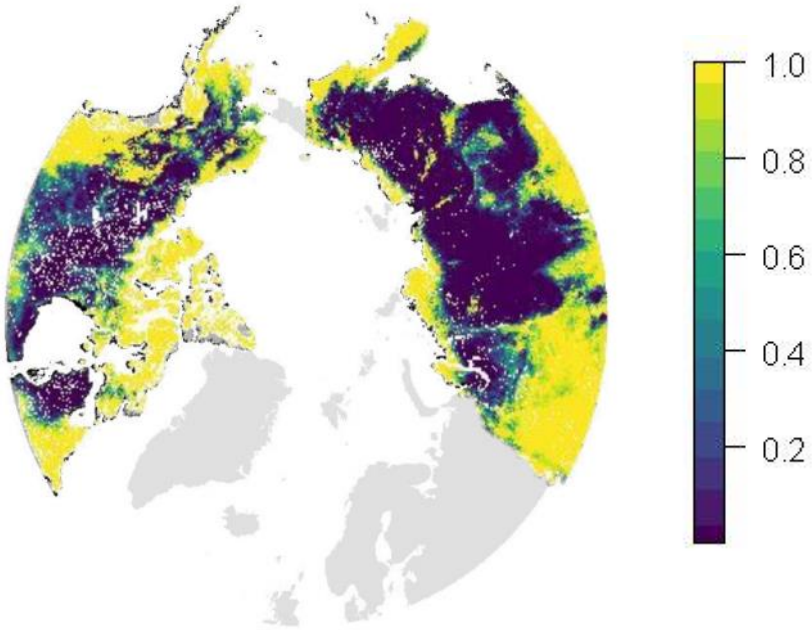
Appendix 13. GBM response curves for solifluction activity



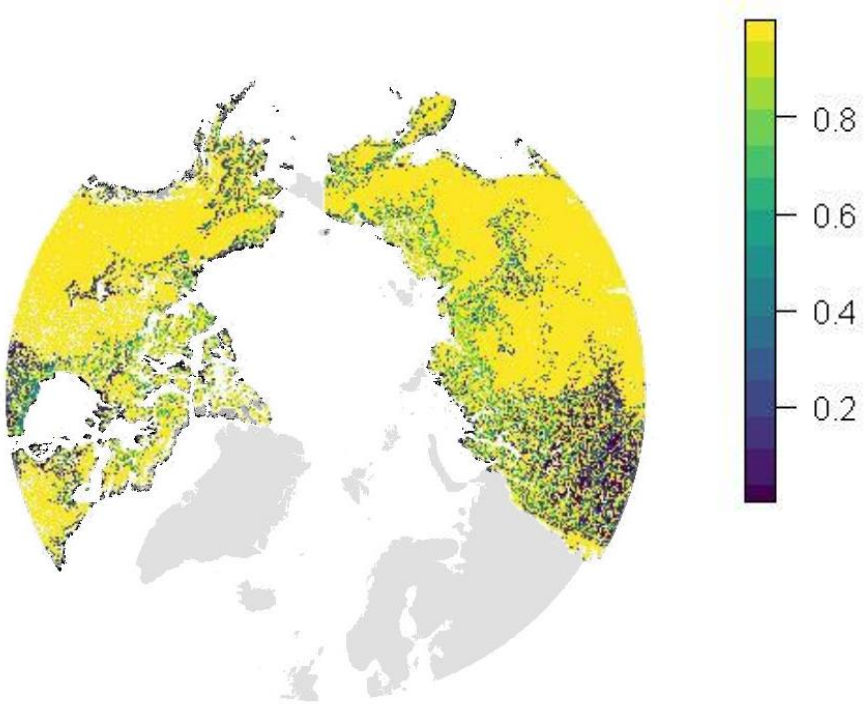
Appendix 14. Predicted occurrence of patterned ground in the circumpolar region using GLM plot. Scale indicates probability of occurrence where 1 = highest probability and 0 = lowest probability.



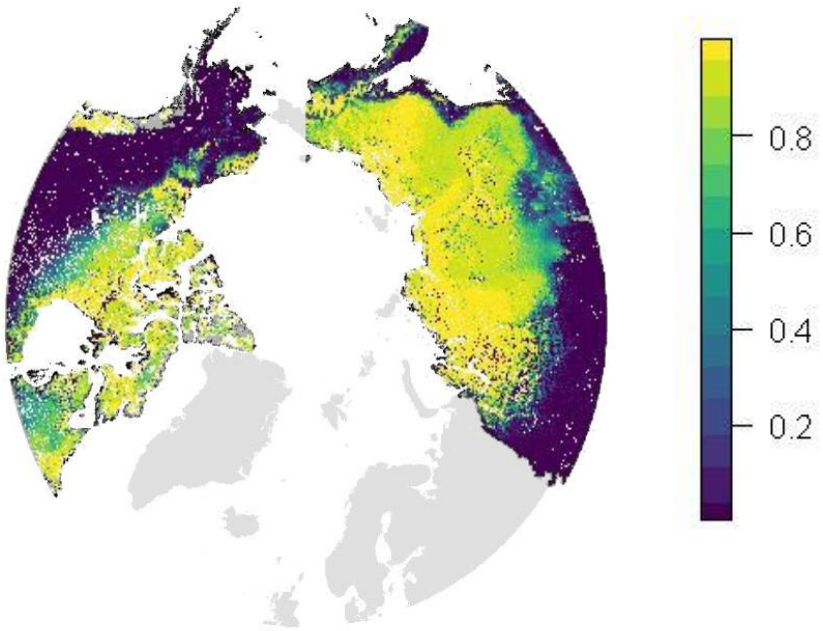
Appendix 15. Predicted occurrence of patterned ground in the circumpolar region using GAM plot. Scale indicates probability of occurrence where 1 = highest probability and 0 = lowest probability.



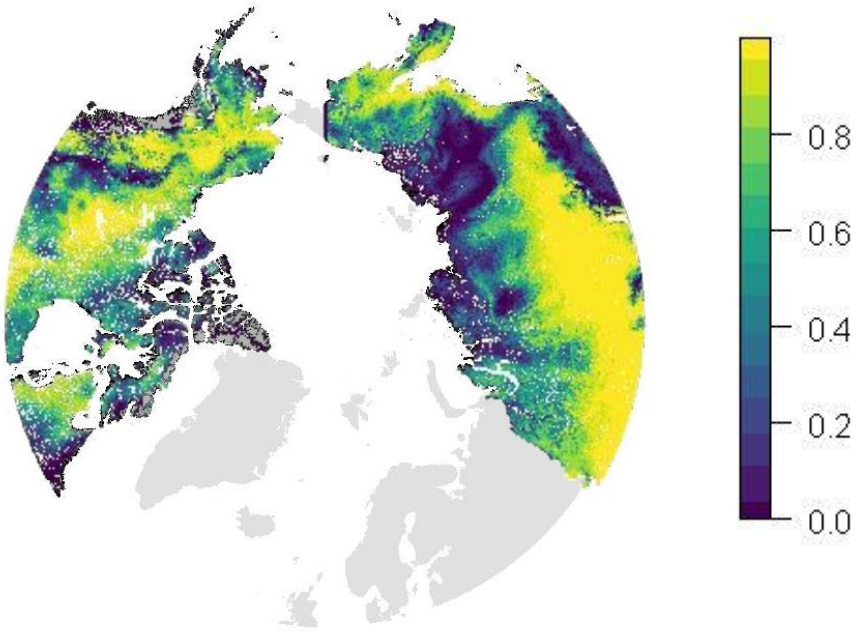
Appendix 16. Predicted occurrence of pingo in the circumpolar region using GLM plot. Scale indicates probability of occurrence where 1 = highest probability and 0 = lowest probability.



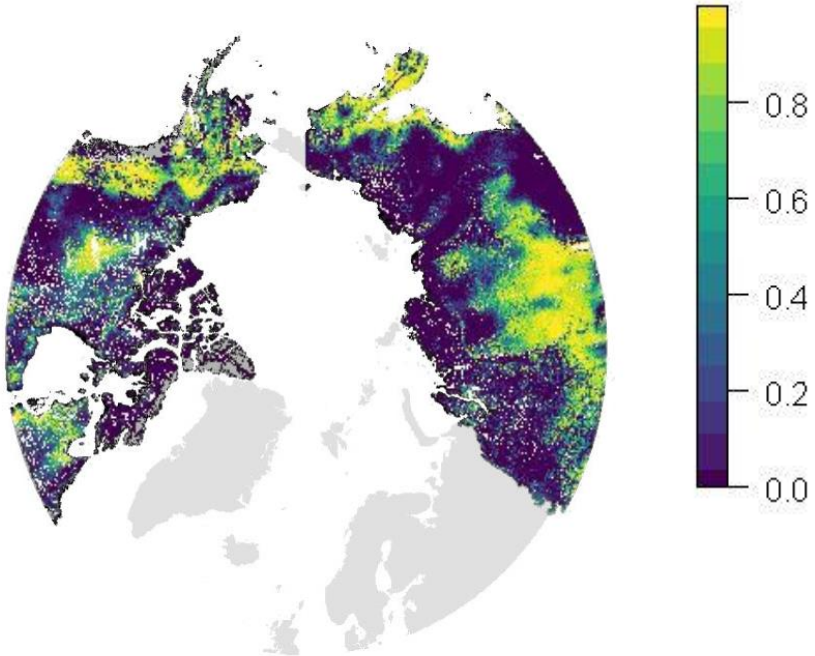
Appendix 17. Predicted occurrence of pingo in the circumpolar region using GAM plot. Scale indicates probability of occurrence where 1 = highest probability and 0 = lowest probability.



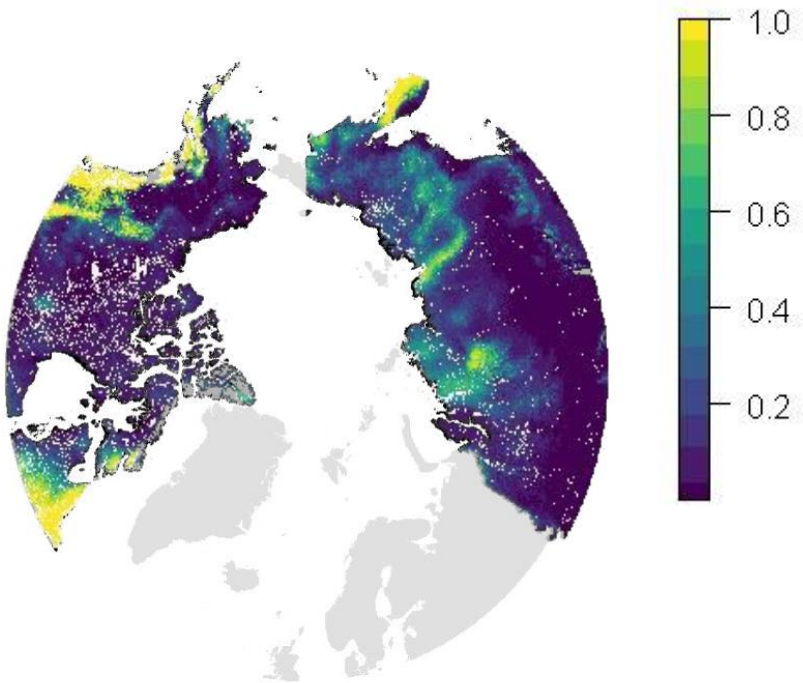
Appendix 18. Predicted occurrence of thermokarst activity in the circumpolar region using GLM plot. Scale indicates probability of occurrence where 1 = highest probability and 0 = lowest probability.



Appendix 19. Predicted occurrence of thermokarst activity in the circumpolar region using GAM plot. Scale indicates probability of occurrence where 1 = highest probability and 0 = lowest probability.



Appendix 20. Predicted occurrence of solifluction in the circumpolar region using GLM plot. Scale indicates probability of occurrence where 1 = highest probability and 0 = lowest probability.



Appendix 21. Predicted occurrence of solifluction in the circumpolar region using GAM plot. Scale indicates probability of occurrence where 1 = highest probability and 0 = lowest probability.