Georg Veh

On the cause of thermal erosion on ice-rich permafrost (Lena River Delta/ Siberia)

An assessment on geomorphological parameters for erosion susceptibility mapping using logistic regression

Master Thesis

supervised by

AR Dr. Florian Haas

to attain the academic degree
Master of Science (M.Sc.) in Geography
submitted to the
Mathematisch-Geographische Fakultät
Katholische Universität Eichstätt-Ingolstadt

Potsdam, August 2015

Georg Veh

Ursachen von Thermoerosion auf eisreichem Permafrost (Lena Delta/ Sibirien)

Ableitung geomorphologischer Parameter zur Kartierung der Erosionsanfälligkeit mittels logistischer Regression

Masterarbeit

Betreuer

AR Dr. Florian Haas

zur Erlangung des akademischen Grades
Master of Science (M.Sc.) Geographie
eingereicht an der
Mathematisch-Geographische Fakultät
Katholische Universität Eichstätt-Ingolstadt

Potsdam, August 2015

Verfasser:

Georg Veh

geboren am 6. März 1990 in Nördlingen (Deutschland)

Adresse: Geschwister-Scholl-Straße 40, 14471 Potsdam

Email: georg.veh@ku.de

Matrikelnummer: 380231

Gutachter:

Dr. Florian Haas

Lehrstuhl für Physische Geographie

Katholische Universität Eichstätt-Ingolstadt

Adresse: Ostenstraße 14, 85072 Eichstätt

Email: florian.haas@ku.de

ABSTRACT.

Vast parts of Arctic Siberia are underlain by ice-rich permafrost, which is exposed to different processes of degradation due to global warming. Thermal erosion as a key process for landscape degradation causes the recent reactivation and formation of new landforms like thermo-erosional valleys and gullies. However, a statistical assessment about the decisive factors and the locations most susceptible for this phenomenon is still missing. This study investigates the influence of different geomorphological parameters on the occurrence of recently observed thermal erosion using a GIS-based approach and statistical modeling by logistic regression. The study site is located on an island within the Arctic Lena River Delta and is mainly composed of ice- and organic-rich deposits. Field surveys and mapping of high-resolution remotely sensed data revealed that thermal erosion occurs predominantly i) on very steep slopes along the margins of the island, ii) in the upper reaches of deeply incised thermo-erosional valleys and iii) in thermo-erosional gullies. Several potentially influencing environmental parameters were derived by a combination of high-resolution satellite imagery and 2 m-DEM. The full set of parameters was reduced stepwise within the logistic regression model. This approach allows the selection of a parsimonious model, i.e. a best-fit model using as few variables as possible. The parameters *Contribution of warm open surface water, Relief ratio, Direct solar radiation and Snow accumulation* turned out be the decisive factors for thermal erosion. Uncertainties in the model due to sampling and model selection were evaluated statistically and spatially through the generation of 100 models. Receiver Operating Characteristics (ROCs) were used to validate the spatial predictive capability of each model run. The consensus map as the median of all susceptibility models represents the final susceptibility map. The agreement between mapped and predicted erosion is generally very high within the study site, confirmed by an Area under the ROC curve (AUC) of 0.957 for the consensus map. The variability of predicted erosion probabilities between the single models is about four percentage points per cell within the study site and thus, very low. Mismatches between observed and predicted erosion could be attributed to the generation of the explanatory environmental parameters and the modeling approach. Model results seem promising for the spatial prediction of susceptible sites for thermal erosion, but require external validation on other sites with comparable environmental conditions.

ZUSAMMENFASSUNG.

Große Teile der sibirischen Arktis gründen auf besonders eisreichem Permafrost, welcher als Folge des Klimawandels verschiedenen Degradationsprozessen ausgesetzt ist. Thermoerosion als ein Schlüsselprozess der Landschaftsdegradation führt zur Reaktivierung und Bildung neuer Geländeformen wie Thermoerosionstäler und –gullies. Eine statistische Auswertung über die entscheidenden Faktoren und die Standorte mit den höchsten Anfälligkeiten gegenüber diesem Prozess fehlt jedoch bislang. Diese Arbeit untersucht den Einfluss verschiedener geomorphologischer Parameter hinsichtlich des Auftretens von rezent beobachteter Thermoerosion anhand eines GIS-basierten Ansatzes und statistischer Modellierung mittels logistischer Regression. Das Untersuchungsgebiet liegt auf einer Insel im arktischen Lena Delta und besteht größtenteils aus eis- und organikreichen Ablagerungen. Geländearbeiten und Kartierungen in hochauflösenden Fernerkundungsdaten zeigten, dass Thermoerosion hauptsächlich i) auf sehr steilen Hängen entlang des Inselsaums, ii) im Oberlauf von tief eingeschnittenen Thermoerosionstälern und iii) in Thermoerosionsgullies auftritt. Mehrere potentiell einflussreiche Umweltparameter wurden mittels einer Kombination aus einem 2m-DHM und Satellitenbildern abgeleitet. Die komplette Reihe an Umweltparameter wurde innerhalb des logistischen Regressionsmodells stufenweise reduziert. Dieser Ansatz erlaubt die Auswahl der wichtigsten Faktoren, welche gleichzeitig das beste Modell erzeugen. Die Faktoren *Einfluss von warmen Oberflächenwasser, Reliefgradient, Direkte Sonneneinstrahlung und Schneeakkumulation* erwiesen sich als die entscheidenden Faktoren für Thermoerosion. Unsicherheiten im Modell aufgrund von Stichproben und Modellauswahl wurden statistisch und räumlich ausgewertet durch die Bildung von 100 Modellen. Die Receiver Operating Characteristic (ROC) Kurve wurde zur Validierung der räumlichen Vorhersagekraft sowohl für jedes einzelne, als auch für das Consensus Modell, welches den Median aus allen Modellläufen darstellt, verwendet. Die Übereinstimmung zwischen kartierter und vorhergesagter Erosion innerhalb des Untersuchungsgebiet ist im Allgemeinen sehr hoch, was eine Fläche unter der ROC Kurve von 0.957 für das Consensus Modell bestätigt. Die Variabilität der vorhergesagten Erosionswahrscheinlichkeit im Untersuchungsgebiet zwischen den einzelnen Modellläufen beträgt etwa 4 Prozentpunkte pro Pixel, was als sehr gering zu interpretieren ist. Unterschiede zwischen beobachteter und vorhergesagter Erosion konnten auf die Erzeugung der erklärenden Umweltparameter und den Modellierungsansatz zurückgeführt werden. Die Modellergebnisse wirken vielversprechend für die räumliche Vorhersage von Thermoerosion, bedürfen aber einer externen Validierung in Gegenden mit vergleichbaren Umweltbedingungen.

Table of contents

List of figures

List of tables

1   Introduction

The periglacial environment reacts very sensitively to thermal changes, particularly in recent times of global warming (Romanovsky et al., 2010a; 2010b) The average surface temperature in the Arctic has increased by c. 0.09°C per decade for the last century, which is about 50 % greater than the average of the whole Northern Hemisphere (AMAP, 2011). The effects on the polar geomorphology are shown by a broad variety of actively degrading landforms (Kokelj and Jorgenson, 2013). Thawing permafrost can lead to mass movements resulting in features like Active Layer Detachments (Lewkowicz and Harris, 2005; Lewkowicz and Kokelj, 2002; Rudy et al., 2013) and Retrogressive Thaw Slumps (Lacelle et al., 2010; Lantuit and Pollard, 2005; Lantz and Kokelj, 2008); to the subsidence of the surface and/ or the formation of lakes as examples of thermokarst (Bouchard et al., 2014; Grosse et al., 2011; Morgenstern et al., 2011); and to new drainage networks through changes in runoff regimes known as thermal erosion. Especially in regions with high ground ice content, these processes dominate the shape of Arctic coastal lowlands (Dallimore et al., 1996). Hence, they have an immense influence on the local and regional water balance (Karlsson et al., 2012), which can in turn affect the vegetation pattern and the ground thermal conditions (Schuur et al., 2007). Besides the alteration of the Arctic ecosystem, they also contribute to the release of soil organic carbon and thus, can be a decisive factor for global warming (Schuur et al., 2008; Walter et al., 2007; Zimov et al., 2006). Up to 1500 Pg of Soil Organic Carbon or 50 % of world's below belowground organic carbon, respectively, are stored in the permafrost regions, of which 800 Pg are perennially frozen. This demonstrates the vulnerability of permafrost deposits to thaw. Furthermore, the potential release of this carbon storage reveals its importance as a positive feedback mechanism to the global carbon cycle (Hugelius et al., 2014; Koven et al., 2011; Tarnocai et al., 2009).

While the formation and spatial distribution of thermokarst lakes have been studied at many sites in the circumpolar permafrost region (Lenz et al., 2013; Marsh et al., 2009; Pohl et al., 2009; Yoshikawa and Hinzman, 2003), comparatively few studies exist about thermo-erosional landforms. Thermal erosion is a combination of both mechanical forces, i.e. the hydraulic effect of flowing water across a surface, and thermal forces, i.e. heat transfer into the ground, causing the thaw of underlying permafrost (Costard et al., 2007; Dupeyrat et al., 2011). Several studies observed the effect of thermal erosion on Arctic coastlines (Günther et al., 2013; Lantuit et al., 2012; Wobus et al., 2011) and on river banks of large river systems draining into

_____

the Arctic Ocean (Costard et al., 2007; Gautier et al., 2003). Towards the interior of ice-rich permafrost lowlands, thermal erosion is often a mixture of two or more forms of landscape degradation. For instance, the growth of a thermokarst lake can lead to the drainage over a pre-existing thermokarst depression causing rapid valley incision (Labrecque et al., 2009; Lauriol et al., 2009). Some case studies (Fortier et al., 2007; Godin et al., 2014; Toniolo et al., 2009) focused in detail on the description and quantification of single processes in thermo-erosional gullies. Morgenstern (2012) qualitatively present the diversity of thermo-erosional landforms and considered general possible driving factors. They showed that the shape and dimensions of thermo-erosional landforms vary a lot, and that they always depend on multiple preconditions.

However, a statistical assessment about the decisive environmental parameters for thermal erosion on ice-rich permafrost is still missing. Spatial modeling, i.e. developing a model to explain the spatial distribution of a phenomenon, faces this problem (Heckmann et al., 2014). In this regard, logistic regression is a frequently chosen approach because it allows establishing a statistical relationship between potential controlling variables and the occurrence of the phenomenon. The second benefit of this method is that these susceptibility maps indicate also the future potential locations of erosion. The reason why logistic regression has not been applied for modeling erosion susceptibility in the Arctic might come from the remote and inaccessible locations, which make ground truth a difficult task. Nevertheless, the application of logistic regression for modeling the susceptibility for gully erosion (e.g. Akgün and Türk, 2011; Conoscenti et al., 2014; Lucà et al., 2011; Martínez-Casasnovas et al., 2004) or landslide hazard (e.g. Ayalew and Yamagishi, 2005; Dai and Lee, 2002; Lee, 2005; Ohlmacher and Davis, 2003) shows its valuable potential in geomorphological research and risk assessment. On the other hand, these results cannot be transferred to the Arctic where the environmental setting (i.e. climate, topography, lithology, soil, land use, etc.) is completely different.

This study focuses on two sites located on Sobo-Sise Island in the Arctic Lena River Delta. The environmental conditions are representative for a major landscape unit in the Siberian arctic lowlands, called Ice Complex. The study aims to close the knowledge gap regarding the causes for the degradation of ice-rich permafrost by thermal erosion in this landscape. The following research questions have been formulated:

1) What are the decisive environmental parameters, which cause thermal erosion on ice-rich permafrost?

2) Which locations in the study sites are most susceptible for thermal erosion?

A GIS-based approach is applied to derive datasets of thermo-erosional features and potential influencing environmental parameters using a DEM and satellite imagery. Stepwise logistic regression is used to detect the dominant influencing parameters for thermal erosion. The final susceptibility map serves to identify the most susceptible locations for thermal erosion within the study sites and to discuss the validity of the modeling results.

2 Scientific background

2.1 Permafrost

Permafrost is defined as perennially frozen ground that remains at or below 0°C for at least two consecutive years (Van Everdingen, 2005). About 24 % of the land area of the Northern hemisphere is underlain by permafrost (Brown et al., 1997; Zhang et al., 1999).

The permafrost distribution is represented as a series of quasi-concentric zones in which permafrost is 1) *continuous*, 2) *discontinuous* or 3) *sporadic* or *isolated*. Accordingly, this means that permafrost underlies either

1)   all terrestrial surfaces except local anomalies like taliks below major water bodies; or

2)   those in which the climate is conducive to permafrost but details of its geography are influenced by factors like slope gradient and aspect, vegetation pattern, thermal properties of the substrate, etc.; or

3)   those in which permafrost occurs only under localized circumstances favorable to its formation or preservation, e.g. in peat deposits (Nelson et al., 2002) (Figure 1).
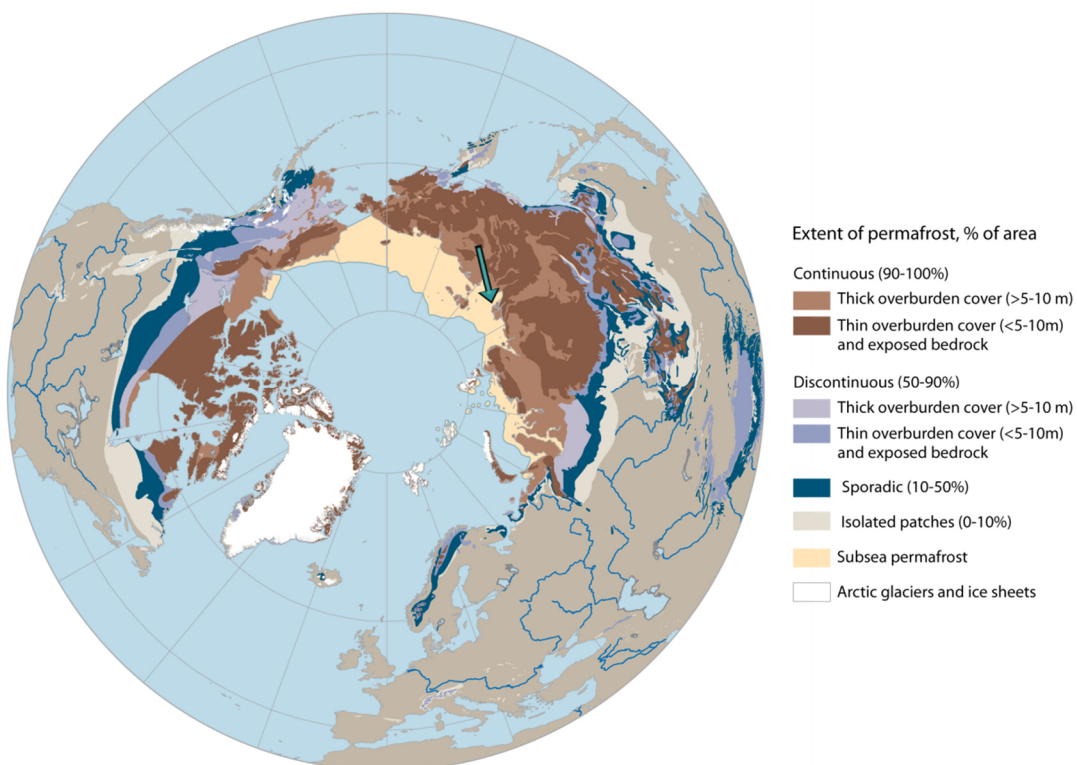
Figure 1: Map of distribution and properties of permafrost and ground ice in the Northern Hemisphere (20°N to 90°N). Permafrost extent is estimated in percent area (90-100%, 50-90%, 10-50%, <10%, and no permafrost) as suggested by Brown et al. (1997, 1998). The arrow indicates the study region Lena Delta (Datasource: AMAP, 2011).

_____

The depth of permafrost ranges from several meters or few decameters to a maximum of about 1600 m in parts of Siberia (Embleton and King, 1975). Permafrost depth represents a negative thermal balance between the surface, referred primarily to air temperatures, and the subsurface conditions, mainly controlled by the geothermal heat gradient (French, 2007).

In vertical dimensions, the ground can be subdivided into an *active layer*, which is affected by seasonal freeze and thaw cycles, and into the permanently frozen underground with an annual temperature amplitude nearly zero. The *permafrost table* delineates the boundary between these two layers, which is often a varying intermediate zone than a distinct boundary. The depth of the active layer can vary
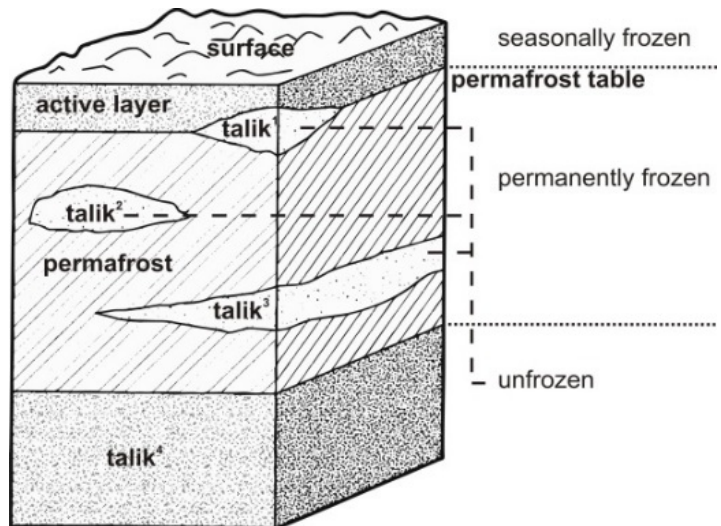
Figure 2: Schematic sketch of a vertical permafrost profile showing the relationship between permafrost, permafrost table, active layer and supra (1), intra (2, 3) and sub permafrost taliks (4) (after French, 2007).

significantly from year to year as well as between locations due to a broad variety of influencing factors, e.g. climate, topographic orientation, vegetation, soil properties or snow cover. The depth of seasonal thaw increases from North to South with few centimeters in the high Arctic to several meters in Subarctic environments (Schirrmeister et al., 2012b). Non-frozen areas within a permafrost zone are called *taliks*. They occur below water bodies (*supra-permafrost talik*), as enclosed lenses due to high salt content (*intra-permafrost talik*) or as the never frozen zone below the maximum permafrost depth (*sub-permafrost talik*) (French, 2007) (Figure 2).

_____

2.2   Permafrost aggradation

The type of ground ice can be classified by the source of the water prior to freezing and the processes, which transfer water to the freezing plane (Mackay, 1972).

The freezing of pore water in the soil leads to the formation of *pore ice* as a process of in-situ-freezing, or to the separation of ice in ice lenses, described as *ice segregation*. The respective type of ground ice depends on the amount and availability of moisture and the properties of the soil, mainly controlled by the grain size and thermal conductivity. Ice lenses may range in thickness from hairline to more than 10 m (Van Everdingen, 2005).

For this work, the focus is on a third type of ground ice, called *ice wedges*. They are considered as one of the most characteristic features in the Arctic periglacial environment.

Ice wedges are massive bodies of wedge-shaped ice and best develop in unconsolidated sediments of the poorly drained tundra in the zone of continuous permafrost. When air temperatures drop well below 0°C in early winter, the thermal gradient from the surface to the ground leads to the contraction of ice and, subsequently, to cracking, i.e. a vertical opening of the ice wedge from the top to its interior. The width of a crack is only about 1 cm, but can reach depths of about 5 m (Mackay, 1974). With the beginning of the melting season in spring/ early summer, the crack fills with water from melting snow. Refreezing and repeated cracking in the following years causes the ice wedge to grow progressively in horizontal and vertical direction. Ice wedges grow in upward direction. Their size and shape is a function of both horizontal and vertical growth rates based on the thermal changes in the upper 5 to 10 m of the permafrost (Black, 1976). Because ice is less dense than the surrounding frozen sediment, further growth of the wedge results in the deformation of adjacent sediment in horizontal and vertical directions (Lachenbruch, 1962). These strong stresses in the underground can alter the

Figure 3: Schematic diagram of the growth of epigenetic and syngenetic wedges. The wedges develop from youngest (1), to intermediate (2) and oldest (3) stage. In cross sections, ice wedges show a two-sided layering due to the seasonal growth along the crack (a, b, c each represent a freezing period) (Mackay, 1990).

_____

shape of an ice wedge from the exemplary, triangular wedge towards an irregular, columnar bulk (Lachenbruch, 1966). In general, the size of ice wedges can vary from less than 10 cm to more than 3 m in diameter and from one to more than 10 m in depth. Maximum dimensions are observed in northeastern Siberia with vertical extents of 30 to 50 m and horizontal extents of more than 10 m (Czudek and Demek, 1970) (Figure 7).

Ice wedges can be subdivided into syngenetic and epigenetic ice wedges due to differences in growth relative to the land surface (Mackay, 1990, 1972) (Figure 3).

Epigenetic ice wedges grow in pre-existing permafrost and are usually much younger than the surrounding material. The surface of the adjoining ground may be raised slightly because of both the accumulation of organic matter and the uplift of the adjacent ground caused by the volume addition of wedge ice. Usually,

Figure 4: Schematic cross section (left) and plane view (right) of an ice wedge polygonal network (Meyer, 2003).

they become about 1 to 1.5 m wide and not more than 4 m deep (Harry and Gozdzik, 1988).

Syngenetic ice wedges can achieve much greater dimensions. They grow contemporaneously with the accumulation of material on the ground surface. They typically occur on floodplains as fluvio-aeolian deposits, beneath peat in polygonal tundra or at the bottom of slopes as gelifluction deposits. In general, syngenetic ice wedges reach their maximum volumes where the ice-accretion rate is high and sediment accumulation is low (French, 2007).

A clearly detectable network of regularly arranged polygons develops on the surface with a

Figure 5: Oblique view on a partially water filled polygonal network on Samoylov Island (Lena Delta). General diameter of polygons is c. 15 m (Source: K. Piel, AWI).

raised rim on either side of the crack (Figure 4). The slight depression above the crack is called *polygon trough*. The spatial pattern of pentagonal, hexagonal, or orthogonal polygons is a complex, site-specific system, depending on several influences like snow and vegetation cover or the occurrence of a water body in the vicinity. In regions with high ice content, the centers of the polygons are often filled with water because the growing ice wedges reach the table of seasonally thaw (Figure 5).

2.3   The Siberian Ice Complex

About 290.000 km² of Siberia are underlain by a very ice-rich type of permafrost called *Ice Complex* (Grosse et al., 2013) (Figure 6). The combination of long-term stable, cold continental climate since the late Pleistocene and the absence of glaciation in this region allowed the deposition of sediments with a volumetric ice contents up to 90 % (Figure 7).

The distribution of the Ice Complex is linked to the interaction of several geomorphological, geological and climatic processes described as the process of *nival lithogenesis* (Kunitsky, 2007). The formation of the Ice Complex started about 80 ka BP and ended with the transition to the Holocene at 12 ka BP (Schirrmeister et al., 2012a).
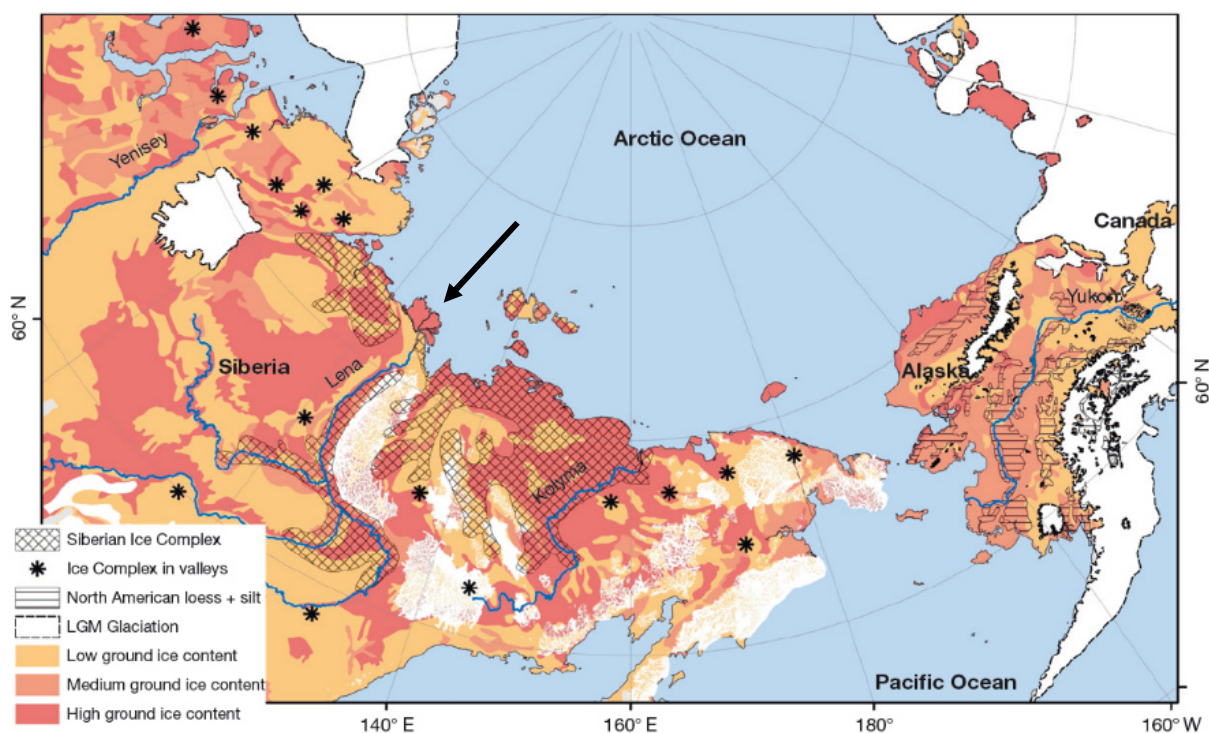
Figure 6: Map of the distribution of ice-rich permafrost deposits in Arctic and Subarctic lowlands. The arrow indicates the study region Lena Delta (Schirrmeister et al., 2012a).

_____

At the beginning, mixtures of windblown snow, plant and mineral detritus accumulated in numerous perennial snowfields in topographically protected zones of hills and low mountain ranges (Kunitsky et al., 2002). Besides the aeolian transport, more silt-sized mineral detritus was produced by intense frost weathering around the snowfields. Over the time, large

Figure 7: Syngenetic ice wedges and intrapolygon frozen sediment columns of an Ice Complex exposure on Bolshoy Lyakhovsky Island. Person for scale. Picture from Schirrmeister et al. (2012).

amounts of plant and clastic detritus additionally accumulated in these snowfields. This material was transported downslope as runoff from meltwater. Several processes of sediment transport (alluvial, proluvial, colluvial, slope wash, solifluction, permafrost creep, and eolian) deposited and reworked this mixture of fine-grained sediment in the foreland of the mountain ridges over thousands of years. Large plains and alluvial fans developed in Northern Siberia. The cold climate conditions promoted the growth of thick syngenetic ice wedges on these supersaturated flat accumulation plains, building up the 40 to 50 m thick Ice Complex deposits. The formation of the Ice Complex is described as *polygenetic* according to the various processes of transport, accumulation and re-sedimentation (Schirrmeister et al., 2011a).

2.4 Degradation of ice-rich permafrost

With the transition from Late Glacial to Holocene, warmer summer temperatures initiated the thaw of permafrost deposits. During the Early Holocene Optimum thawing of Ice Complex reached its maximum (Grosse et al., 2007). About 70 % of the whole Arctic Ice Complex terrain was affected by thermokarst-related processes (Walter-Anthony et al., 2014), while smaller subsets around the Laptev Sea show slight variations thereof, e.g. 78 % in the Lena-Anabar coastal lowlands (Grosse et al., 2006) or 65 % in the Kolyma lowlands (Veremeeva and Gubin, 2009). Nevertheless, several studies in other permafrost landscapes also observed a significant, partially abrupt increase of thermokarst activity in recent decades (Agafonov et al., 2004; Jones et al., 2011; Jorgenson et al., 2006; Osterkamp et al., 2009).
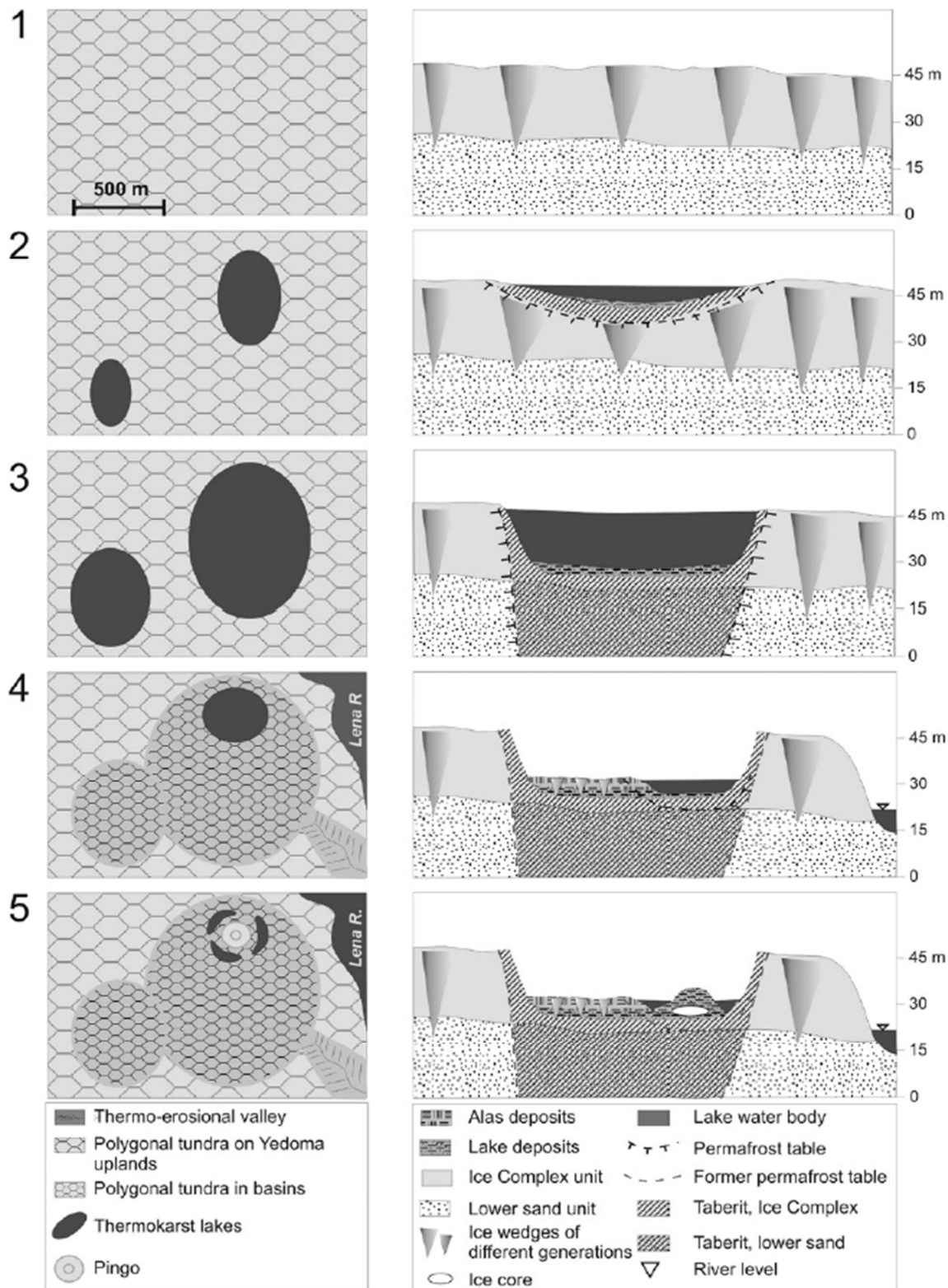
Figure 8: Scheme of thermokarst development in Ice Complex landscapes, adapted for the Lena River Delta, in plain view (left) and cross section (right). 1: Undisturbed polygonal tundra. 2: Initiation of Thermokarst lakes on Ice Complex uplands with lateral and vertical thermokarst development. 3: Maturity of thermokarst lakes with lateral expansion only, lake sedimentation and talik development. 4: Partial drainage of thermokarst basin. Refreezing of former lake bottom with ice aggradation and peat accumulation. 5: Partial drained coalesced thermokarst basin with pingo (Source: Morgenstern et al, 2011).

_____

The degradation of ice-rich permafrost proceeds in a succession of different stages, first described by Soloviev (1962) and Czudek and Demek (1970) for Central Yakutia and refined by Jorgenson and Osterkamp (2005) for boreal permafrost ecosystems. Morgenstern et al. (2011) specified the development of thermokarst on Ice Complex landscapes in Siberian coastal lowlands by a conceptual model (Figure 8).

Starting from the original low center polygonal tundra (Figure 8-1), water accumulates along the ice wedge troughs or in polygonal ponds which grow and coalesce. The polygon rims subside due to melt of subsurface ice wedges. The sediment packages between the wedges keep their height and form a flat top. Thus, former low-center polygons change to high-center polygons. This implies first a shift of the local hydrological characteristics and, subsequently, a change in the regional runoff regime of the tundra surface (Figure 9).
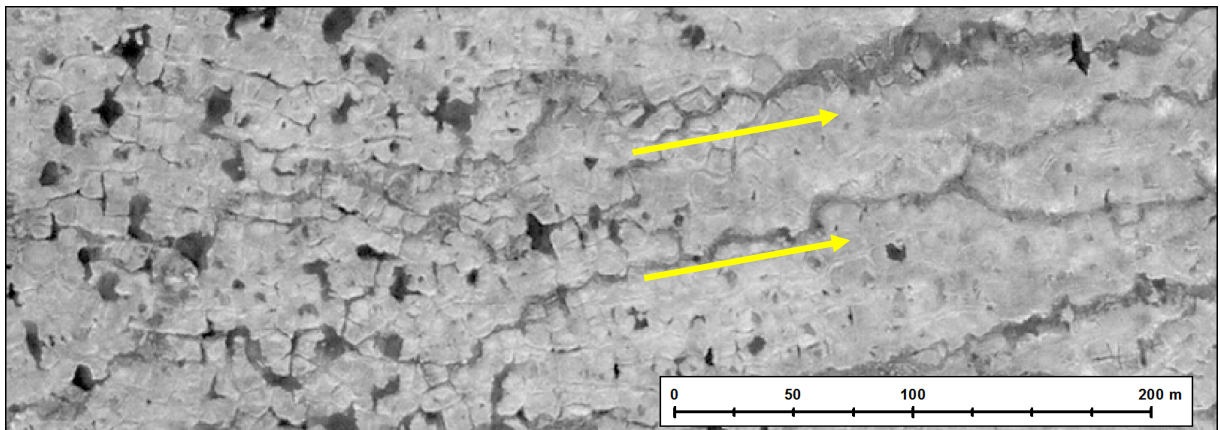
Figure 9: Transition from low-center polygons with water-filled ponds towards high-center polygons with water-filled troughs. These troughs serve as initial pathways for discharge towards lower relief levels. The arrows indicate direction of currently developing discharge channels. Example from Sobo-Sise Island, Lena Delta (Source: Quickbird02, IR-Channel).

Continuous thawing and erosion leave isolated thermokarst mounds (or *baydjarakhs* in Yakutian language) consisting of fine-grained sediment or peat remnants of Ice Complex deposits. The occurrence of baydjarakhs is linked to a relief gradient where water from thawing ice wedges is directly discharged towards lower relief units, e.g. a lake, a depression or a major river (Figure 10).

A thermokarst lake will develop where a sufficient amount of water is concentrated (Figure 8-2). Due to the high heat storage capacity of the growing lake and its heat transfer into the ground, a talik can develop, if the lake does not freeze completely in winter. The taliks

_____

Figure 10: Alas with distinct transition from the floor towards slopes with stabilized baydjarakhs. Baydjarakhs are typically 3 to 10 m in diameter. Scene from Sobo-Sise Island, Lena Delta.

extend sometimes several decametres to more than 100 m in depth (West and Plug, 2008). The ice loss below the lake causes a subsidence of the lake bottom and a compaction of both Ice Complex and lake sediments, called *taberit* (Figure 8-2 & 3). The lake laterally may increase its size by lateral expansion (i.e. thermal abrasion on the shores). It can coalesce with other thermokarst lakes to form larger inland water bodies with several kilometers in diameter (Figure 11).
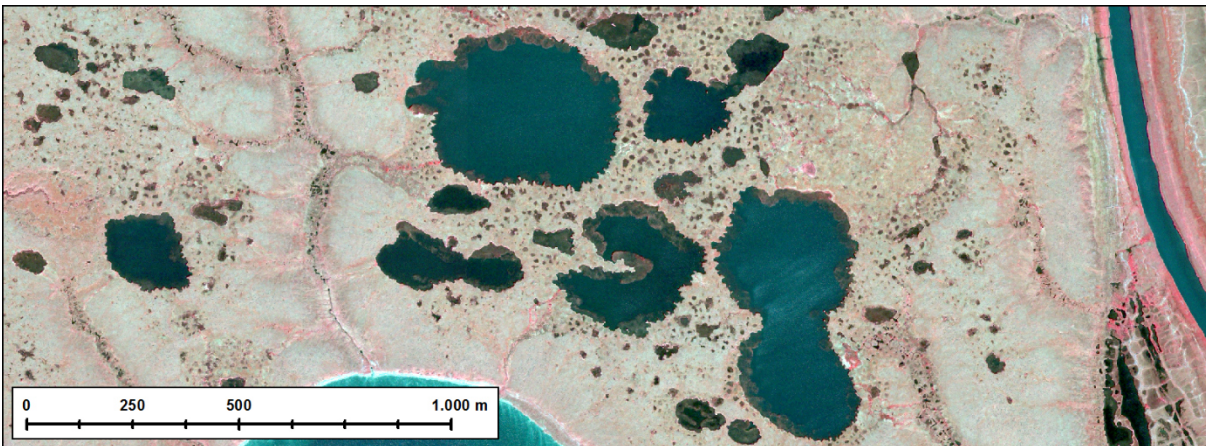
Figure 11: Series of coalescent small-size lakes on the Ice Complex surface on Sobo-Sise, Lena Delta. All these lakes may once form a larger waterbody (Source: Quickbird02, Channels 4-3-2).

However, these lakes can drain by one single rapid, catastrophic event or slowly over a longer period for several years, shaping a distinct valley after the outlet. This process represents one type of thermal erosion (Figure 8-4).

A thermokarst depression (or *alas* in Yakutian language) remains, which consists of a flat floor with a distinct transition towards comparatively steep slopes. The floor can be covered with eroded material from the alas slopes (Figure 12). New permafrost aggradation can start in the partially or fully drained lake bottom. Epigenetic ice wedges with thicknesses of 5 to 7 m

_____

can grow into the flat ground by seasonal freeze-and-thaw-cycles and a new polygonal net may develop (Morgenstern et al., 2011). These second-generation surface features can degrade again, form new thermokarst lakes etc., thus representing a polycyclicity of relief genesis.
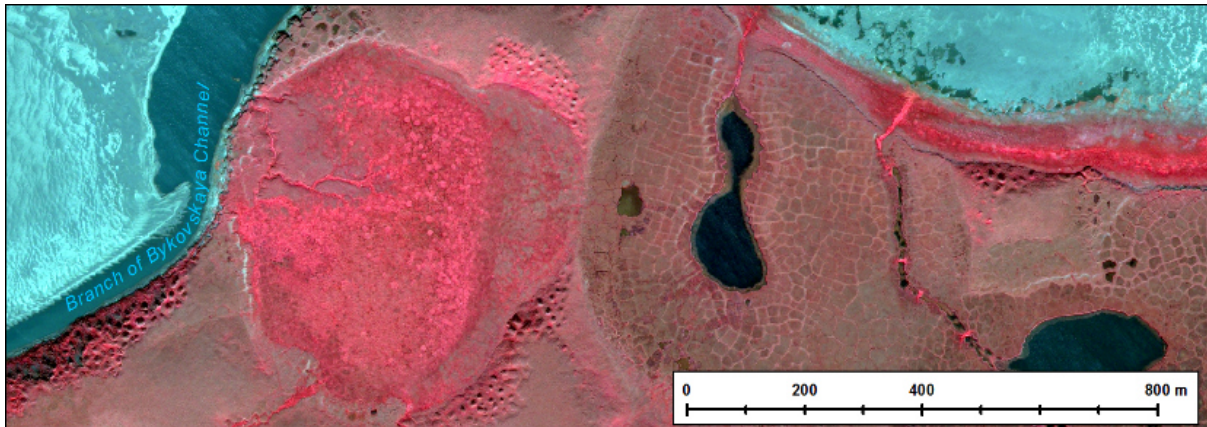
Figure 12: Example for two drained lakes on Sobo-Sise Island, Lena Delta. Left: Recently, completely drained lake surrounded by fresh bayjdarakhs. Right: Well-developed, second generation polygonal network in an old, partially drained lake (Source: Quickbird02, Channels 4-3-2).

2.5   Thermal erosion

Thermal erosion is defined as "the erosion of ice-bearing permafrost by the combined thermal and mechanical action of moving water" (Van Everdingen, 2005). This process differs from the development of thermokarst, which occurs due to thermal melting followed by subsidence of the ground. Thus, thermal erosion also contains the removal and transport of thawed sediment.

Typical locations for this process are Arctic coasts (Günther et al., 2013; Lantuit and Pollard, 2008), riverbanks of large rivers in permafrost landscapes (Costard et al., 2003; Gautier et al., 2003) and the surfaces of ice-rich sediments (Fortier et al., 2007; Godin et al., 2014). In this study, the focus is on ice-rich permafrost, where thermal erosion causes the rapid formation of thermo-erosional gullies and valleys (e.g. Morgenstern, 2012). Here, common stages in the erosion of underlying ice wedges include slumping, piping and the creation of small tunnels (French, 2007).

Morgenstern (2012) observed eight different types of thermo-erosional landforms on ice-rich permafrost, ranging from small water tracks to actively eroding gullies and stabilized valleys.

Table 1: Categorization of thermo-erosional valleys and corresponding hydrological features (from Morgenstern, 2012).

Category	Occurrence	Characteristics	Hydrologic regime
Short, straight gullies	On alas and thermokarst lake slopes	Radially arranged around lakes and alasses; v- to u-shaped; steep gradient; up to few meters deep and wide; dense, fresh vegetation	Intermittent streams
Drainage pathways in alasses	On alas floor	Connect residual and secondary thermokarst lakes in partly drained alasses with the stream network outside the alasses; slightly intented into the alas floor; low gradient; up to a few meters wide; dense, vital vegetation	Intermittent and small permanent streams
V-shaped ravines	Along steep coasts and cliffs; often due to lake drainage	V-shaped; steep to moderate gradient, upt o tens of meters deep and wide; vegetation cover on floor and lower slopes often disturbed	Intermittent streams
V-shaped valleys	In upper parts of the watersheds on Yedoma surfaces	Mostly tributary valleys; V-shaped; moderate to low gradient, up to tens of meters deep and hundreds of meters wide; intact vegetation cover	streams
U-shaped valleys	On Yedoma Surface	U-shaped; low gradient, up to tens of meters deep and several to tens of meters wide; flat valley floor with vital vegetation	Intermittent and small permanent streams
U-shaped valleys of permanent streams and rivers	Lower parts of long streams close to their mouth	U-shaped; low gradient, up to tens of meters deep and hundreds of meters wide; broad floors with distinct floodplains; often bare sediment exposed; oxbow and small thermokarst lakes	Permanent, meandering streams

Broad valley floodplains	Lower parts of long streams close to their mouth	Low gradient, up to tens of meters deep and hundreds of meters to kilometers wide; broad floors with distinct floodplains; often bare sediment exposed; oxbow and small thermokarst lakes	Permanent, meandering streams
Water tracks	On gently sloping Yedoma surfaces; on large, slightly inclined alas floors	Arranged in parallel; low gradient; not or only slightly indented into the surface; dense,vital vegetation	Poorly developed runoff systems

_____

3   Regional setting and study site

3.1   Study region Lena Delta

3.1.1   General regional setting

The Lena Delta covers about 32000 km^2 and is the world's largest Arctic delta (Walker, 1998). The Lena River originates in the Baikal mountains, flowing about 4400 km to the North and discharges on average 520 km³/ yr into the Laptev Sea within a corridor between 72° and 74° N and 123° and 130° E (Schwamborn et al., 2002) (Figure 13). The enormous sediment load (17.6*10^6 t/yr) from the 2.5*10^6 km² catchment has built up a semi-circular accumulation plain (Gordeev and Sidorov, 1993; Rachold and Grigoriev, 1999). It is bounded to the East, North and West by the Laptev Sea, a shelf sea of the Arctic Ocean, and to the South by the Chekanovsky and Kharaulakh ridges with maximum elevations up to 500 m a.s.l (Schirrmeister et al., 2011b). The whole region is underlain by continuous permafrost with a maximum
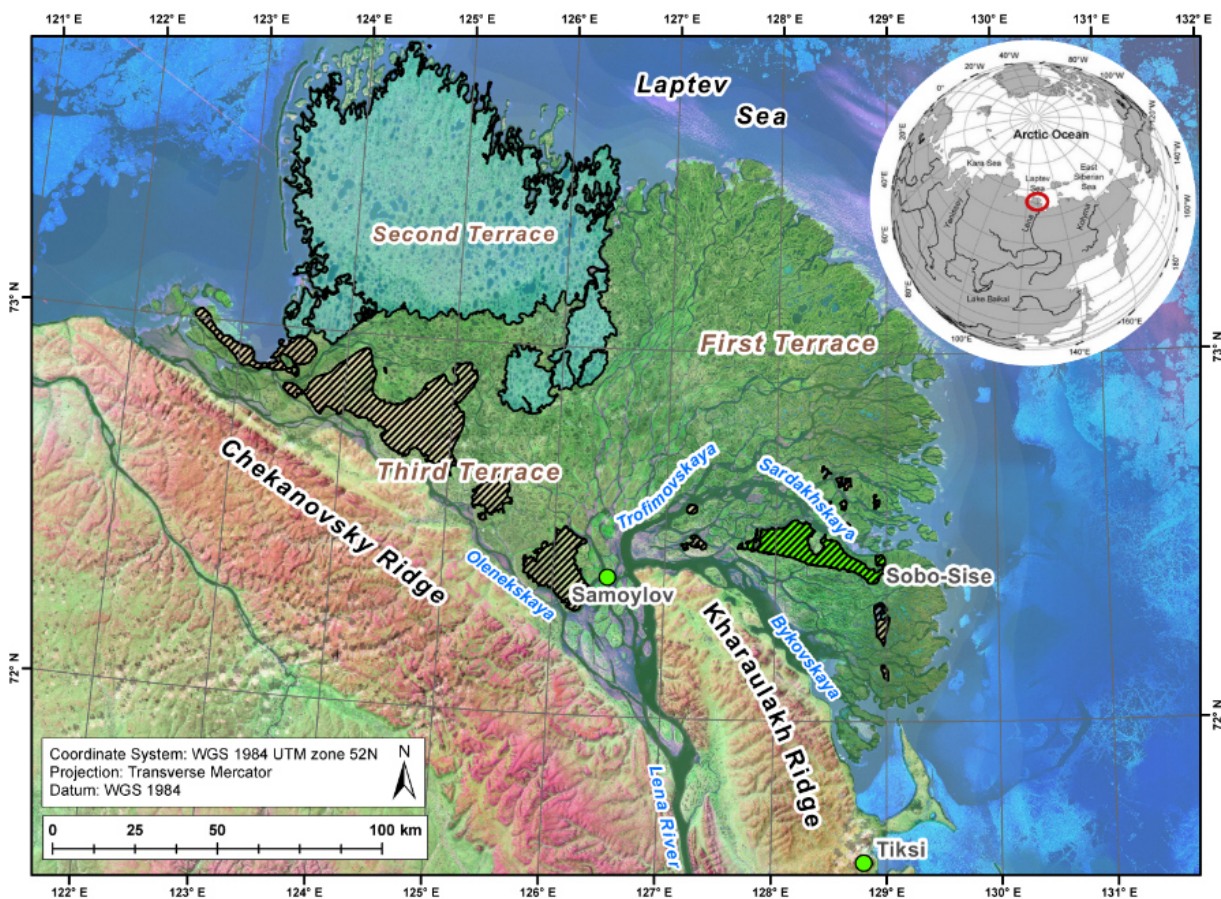
Figure 13: Map of the Lena Delta including the main river branches, the geomorphologic terraces (3rd terrace fasciated in black) and locations mentioned in the text (Study site Sobo-Sise highlighted in green). Data source: Landsat ETM+, SRTM.

16

_____

thickness of about 500 to 700 m (Romanovskii, 2004). The delta consists of several hundreds of river branches, which form more than 1500 islands (Are and Reimnitz, 2000). The combination of tectonic stresses, eustatic sea level rise and sediment transport formed its present shape and sedimentary structure since the late Pleistocene (Schwamborn et al., 2002).

3.1.2 Geology and sediment succession in the Lena Delta

The position in the Arctic Rift Zone has led to vertical block tectonics with narrow long horsts, grabens and synclines in the Lena Delta and the surrounding marine region. The modern seismic activity is still very high, especially in the shallow, broad shelf of the Ust-Lena-Rift system in the north-eastern offshore region (Drachev, 2000; Franke et al., 2000) (Figure 14). The tectonic stresses caused channel migration in the delta, which is apparent in a remarkable linearity of the main branches like the Olenyokskaya or the Bykovskaya channel (Schwamborn et al., 2002) (Figure 13). Such shifts in the main runoff direction occurred several times since the late Pleistocene. In this period of extreme dry-continental climate conditions, the position of the Lena Delta extended far northwards due to regression of the Laptev Sea with a sea level about 80-100 m lower than today (Hubberten et al., 2004; Romanovskii et al., 2000). The Holocene represents a phase of transgression with a rapid eustatic sea level rise. The current sea level was reached by the middle of the Holocene, c. 5 ka BP (Bauch et al., 2001). Consequently, the interaction of climatic changes and tectonic stresses caused periods of activity and inactivity in different parts of the delta. Thus, three geomorphological terraces developed in the Lena Delta (Figure 13 and Figure 14).
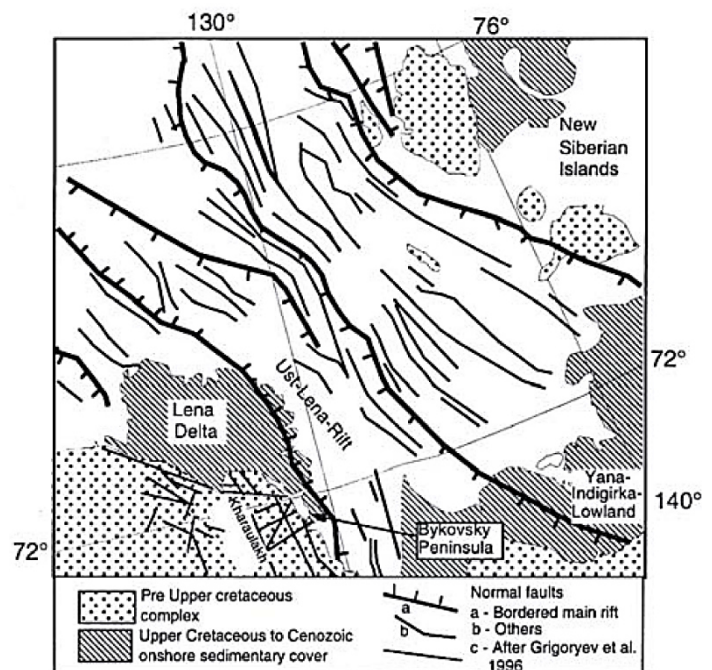
Figure 14: Schematic sketch of main faults in the Lena Delta Region (Schirrmeister et al. 1999, based on citations therein).

The first terrace represents the modern active zone of sedimentation, located mainly in the eastern part of the delta. The Holocene-aged sediments (max 6 to 5 ka BP) with elevations of 1 to 12 m a.s.l. show inverse layering which demonstrate present high energetic sedimentation conditions in the delta plain (Schwamborn et al., 2002). Sediment accumulation on the floodplains occurs in early summer when the
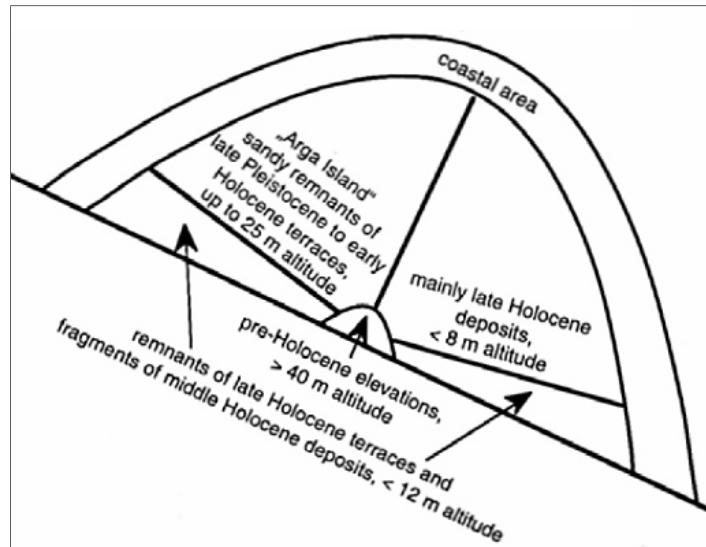
Figure 15: Schematic sketch of the main geomorphic terraces in the Lena Delta (Schwamborn and Griegoriev, 1999).

Lena River reaches its yearly peak discharge after snow melt (Figure 16). Discharge amounts in June can be 55 times greater than the low basal winter discharge (Yang et al., 2002). Ice barriers can raise the river level for more than 10 m and cause vast inundations.

The second terrace with heights ranging from 20-30 m a.s.l. formed during the transition from late Pleistocene to early Holocene (17 to 12 ka BP). Results of sedimentological analysis state a braided river system, which deposited enormous amounts of fine-sandy sediments. This period of increased river discharge is assumed to have lasted for only a short time (maximum discharge around 13 ka BP), representing an abrupt climatic warming (Schwamborn et al., 2002). Tectonic uplift in the western part of the delta is supposed to cause the current inactivity of the second terrace (Schirrmeister et al., 2011a). Located in the northwestern part of the delta, it covers about 23 % of the delta area.

Figure 16: Mean discharge of the Lena River at Stolb in close vicinity to Samoylov Island. Note the distinct peak in June, which leads to severe flooding (R-ArcticNET, 2015).

For this study, the third terrace is of peculiar interest. The elevation of the third main terrace ranges from 20 to 66 m a.s.l.. They are located at the

18

southern margins of the Lena River Delta. It is the smallest terrace by area with about 1700 km² and covers nearly 6 % of the whole delta surface (Morgenstern et al 2011). The third terrace represents the oldest part of the delta. It consists of two distinct layers. In the lower deposits, the facies consists of fine-grained fluvial sands with alluvial peaty layers (ca. 88 to 43 ka BP), representing a Paleo-Lena River deposition. After an abrupt facies change, Ice Complex deposits of polygenetic origin follow on top with extremely ice-rich sediments containing thick ice wedges (Schwamborn et al., 2002). The Ice Complex in the Lena Delta developed between 43 and 14 ka BP due to cold climate conditions in this region. In this period, the Chekanovsky and Kharaulakh Ridge represent the main sediment source for the growth of syngenetic ice wedges in a poorly drained accumulation plain. Tectonic movement forced a tilt of the islands of the third terrace with uplift in the western part of the Lena Delta and subsidence in the east, showing a relative relief gradient of about 20 m from west to east. Schwamborn et al. (2002) regard tectonic activity in the development stage as the decisive factor for the strong facies boundary between fluvial sands and overlying Ice Complex (Figure 17).

With the transition to the Holocene, the degradation of the Ice Complex started due to the ameliorated climate conditions, which is evident in the occurrence of thermokarst and thermo-erosional landforms throughout the islands of the third terrace. These negative relief features became filled with eroded material and new polygonal ice-wedge systems developed in these

Figure 17: Ice Complex bluff on Kurungnakh Island near Samoylov Island with Olenekskaya Channel and eroded bluff deposits in the foreground. The dashed black line indicates the boundary between fluvial sands and Ice Complex deposits. Note the bulky peat inclusions between the wedges. Scale: Vertical extent of the Ice Complex wall c. 45 m.

19

sediments. Peat accumulated on the surface between 4 and 1 ka BP, which is again affected by thermokarst processes in recent time (Schirrmeister et al., 2011b; Schirrmeister et al., 2002a, 2002b; Wetterich et al., 2008).

3.1.3   Climate characteristics of the Lena Delta

The Lena Delta is dominated by an Arctic continental climate regime with maritime influence due to its vicinity to the Laptev Sea. The longest consecutive climate record representative for the Lena Delta is located at the southern margin of the delta near Tiksi (71.63 N, 128.87 E) (Figure 18). While the mean annual air temperature is negative (−13.5 °C), there exists a high annual amplitude from −32 °C in January to 6.5 °C in July. This is a consequence of the position in the Arctic with low or zero insulation in winter during polar night and inverse conditions in summer during polar day.

The mean annual precipitation at Tiksi is 323 mm. Due to its location near the Laptev Sea with surrounding mountains this value would probably not represent the general precipitation characteristics in the Lena Delta. In this regard, the meteorological station on Samoylov Island (see Figure 13) delivers data that are more reliable. Here, observations from 1999 to 2011 show a mean annual amount of rainfall of about 125 mm. Most rainfall occurs between middle of May to the end of September. 70% of these rainfall events are characterized as light rainfall events with 1 mm precipitation and only 1% of the collected events are described as heavy precipitation events > 16mm (Boike et al., 2013).

The transition from rain to snow starts between September and October. The maximum snow depth varies on average between 30 and 40 cm, but it is highly variable due to micro-topographic effects. Strong Arctic winds redistribute
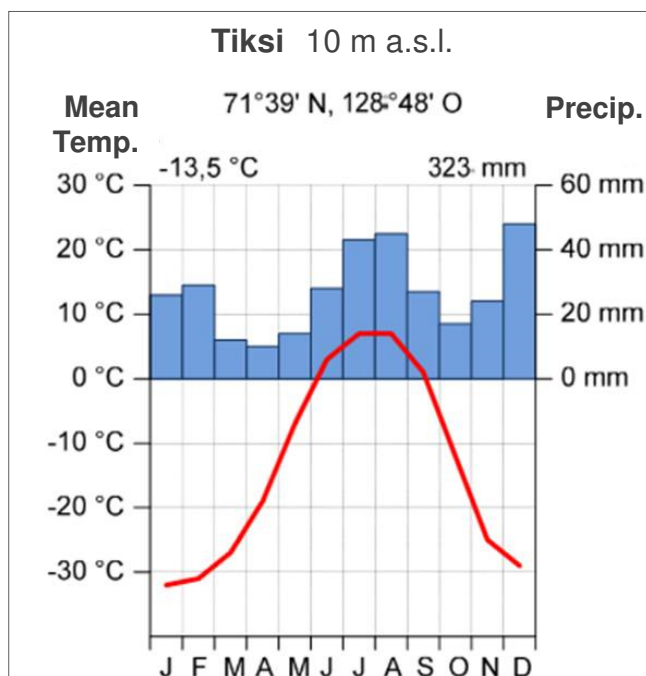
Figure 18: Mean annual temperature and precipitation at Tiksi (Roshydromet, 2015).

the snow and can lead to accumulation in polygon centers on the one hand and to snow-free sites on polygons rims on the other hand. Thus, snow patches form, which can even outlast the complete summer season under favorable conditions (Kunitsky et al., 2002). Snowmelt usually starts in the second half of May, causing an over-saturation of the flat polygonal tundra in the following months due to low evaporation (Boike et al. 2013).

In contrast to the general strong Arctic warming trend (AMAP, 2011), no clear increase in annual air temperatures was observed on Samoylov Island over the past few years, even though some winters were not as cold as mean winter temperatures (Boike et al., 2013) (Figure 19).
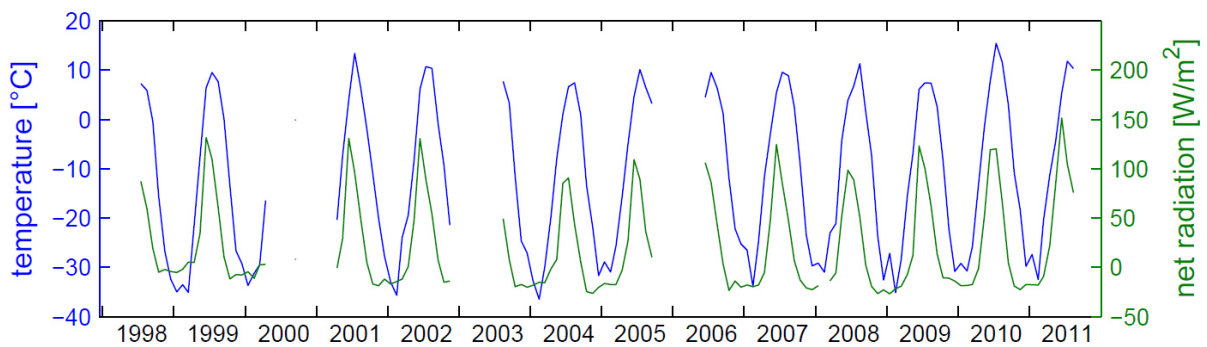
Figure 19: Mean monthly air temperature and net radiation record for Samoylov Island, 1998- 2011 (Boike et al. 2013).

3.1.4 Vegetation and soils

The vegetation period in the Lena Delta lasts up to three months. The Arctic vegetation shows a low diversity and is adapted to extreme environmental conditions like low solar radiation and temperatures, high wind speeds and wet soils through both low annual rates of growth and low maximum growth heights. The Lena Delta is covered by typical tundra vegetation, consisting of grasses, sedges, mosses, lichens, herbs and dwarf shrubs. The composition of species varies locally between *wet* and *dry* tundra based on differences in surface wetness (Boike et al., 2013; Muster et al., 2012) (Figure 20).

Long-term stable cold-climate conditions impeded soil genesis and led to a considerable accumulation of organic matter in the Quaternary, which is additionally supported by fluvial and/or aeolian sediment input (Zubrzycki et al., 2014, 2013). This organic-rich surface layer

generally ranges between 1 to 3 m in thickness. In general, the depth of the active layer upon Ice Complex deposits ranges between 15 and 70 cm. Deep frost penetration and frost-action processes reshape the active layer. Permafrost soils are classified as cryosols (FAO, 2014) or gelisols (Soil Survey Staff, 2014). Soil conditions in cryosols can differ greatly

Figure 20: Differences in the composition of vegetation between polygon rim and polygon center. Scale: diameter of polygon c. 5 m (Picture by N. Bornemann, 2014).

between single sites due to a varying depth of the water table in the soil, leading to oxidative or reducing conditions. The most common soil types in the Lena Delta are *Glacic Aquiturbels*, *Typic Aquiturbels* and *Typic Historthels* (Kutzbach et al., 2004).

3.2   Key region Sobo-Sise

3.2.1   Geology and geomorphology

Sobo-Sise is an island in the southeastern part of the Lena Delta. Its west to east trending, elongated shape mainly developed by the interaction of two main branches of the Lena River. Towards the north, the Sardakhskaya Channel borders the island, whereas the Bykovskaya Channel forms the southern margins (Figure 21).
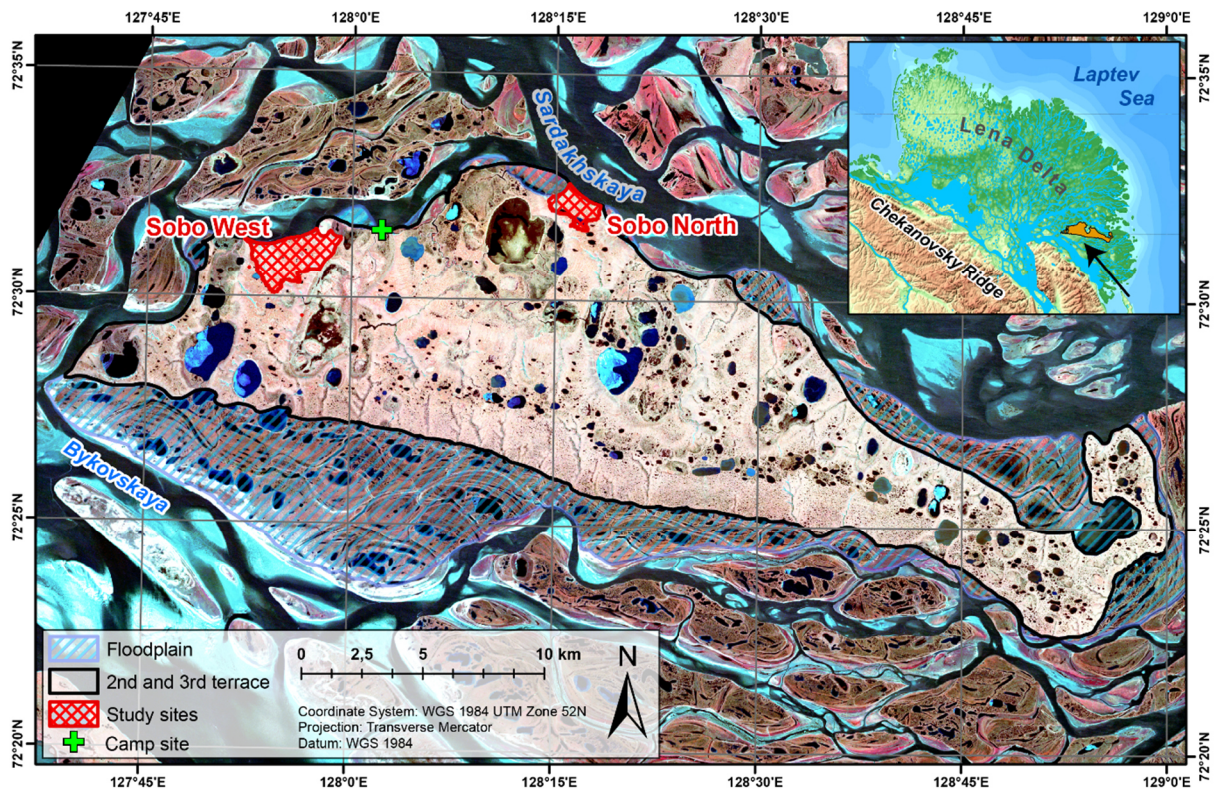
Figure 21: Sobo-Sise Island in the eastern Lena Delta with Ice Complex uplands and surrounding floodplains. Thermokarst features are clearly visible on the Ice Complex surface. A histogram stretch was applied in order to highlight the differences of reflectance between floodplains and Ice Complex uplands (Source: RapidEye, Channels 5-3-2).

Although there is currently no detailed information about the island's geological setting, Sobo-Sise can be roughly subdivided into two sections. Large parts of the southern as well as smaller parts of the western reach belong to the first terrace. This terrace consists of floodplains with a different annual sedimentation activity. Lower parts (max. 4 m a.r.l.) of the first terrace comprise sandbanks that change their shape and spatial distribution after the annual Lena flood in late spring. Higher levels of the first terrace comprise shallow lakes, ponds and oxbows. Late Pleistocene Ice Complex deposits underlie the main part of Sobo-Sise. This study focuses exclusively on Ice Complex terrain, which comprises a size of 319 km².

23

According to their origin as an accumulation plain in the foreland of the Kharaulakh Ridge, the Ice Complex deposits form uplands with homogeneous elevations between 25 - 35 m a.r.l. (max. elevation: 39.2 m a.r.l.). These deposits show a very high absolute ice content of 80 – 90 %. They consist of thick ice wedges, with a maximum depth of 20 m and width up to 5 m, and bulky peat inclusions. The subsurface structure of the Ice Complex is well demonstrated on a nearly vertical exposure at the northern tip of Sobo-Sise, where the Sardakhskaya Channel hits the Ice Complex and leads to massive thermal river erosion (Figure 22). In contrast to other Ice Complex sites, no underlying fluvial sands are present at this spot, which is a result from tectonic subsidence of the Ice Complex islands in the eastern Lena Delta as described in Sec. 3.1.2 (Schwamborn et al., 2002).

Various types of permafrost degradation exist on Sobo-Sise. The island shows the highest number of lakes among all Ice Complex islands in the Lena River Delta. Morgenstern et al. (2011) reported 841 major lakes (> 900 m²), covering 32.2 % of the total island surface. They occur on the Yedoma uplands as well as in the broad alasses and can exceed a size of 3 km². The alasses, which cover 27.2 % of the island total area, dissect the landscape into several distinct Yedoma uplands.

Figure 22: Ice Complex outcrop at the northern tip of Sobo-Sise. During higher water levels, the Sardakhskaya Channel forms a thermo-erosional niche, which favours massive block failure at the bluff. Note the high ice content and the bowl-shaped peat layers on top (image by A. Morgenstern 2014).

_____

3.2.2 Study sites on Sobo-Sise

On Sobo-Sise, two sites on Ice Complex terrain were chosen during the expedition in August 2014 for the study on controlling factors of thermal erosion. A broad corridor (ca. 9 km) of coalesced alasses separates the two Yedoma uplands into a western and northern area of interest (in the following Sobo West and Sobo North) (Figure 21).

Sobo West comprises a size of 5.23 km² with an absolute altitude difference of 30.5 m within the study site. The Lena River erodes the northern margin thereby creating a steep bluff with prominent baydjarakhs. The northeastern and eastern boundary of the study site is marked by a moderate slope from the Yedoma uplands towards a flat alas floor (max.elev. 3-5 m a.r.l.). The watersheds of the investigated streams delineate the southern and western borders of the study site.

The second study site Sobo North is situated at the northeastern tip of Sobo-Sise. Sobo North is bounded from its westernmost until its northernmost point by a floodplain with elevations between 1 and 3 m a.r.l.. The Lena River forms a 1.6 km long vertical bluff from the northern tip until the easternmost point of the study site. The watersheds of the streams draining from the Ice Complex into the Lena River delineate the eastern and southern margins of the study site. With approximately 2 km², the size of Sobo North is less than half than that of Sobo West, whereas the maximum elevation difference within the study site is the same (31 m).

_____

4   Material and methods

The methodological framework to address the aim of this work contains

I. Field work to characterize the study sites and map erosional features (Sec. 4.1)

II. Generation of a high-resolution DEM and orthorectified satellite images on the basis of GeoEye-2 stereo satellite imagery (Sec. 4.2);

III. GIS-based mapping of thermo-erosional features and determination of potentially influencing environmental parameters (Sec. 4.3);

IV. Statistical modeling of controlling factors for thermal erosion using logistic regression and generation of susceptibility maps (Sec. 4.4).

Field work
- Definition of the study sites
- Mapping of erosional features
- GCP collection

Raw data processing
- DEM generation & editing
- Orthorectification of GeoEye-1 image pairs

GIS dataset
- Mapping of thermoerosional features
- Deviation of influencing parameters

Statistical modeling
- Detection of decisive factors for erosion via logistic regression
- Generation of susceptibilty maps
- Validation

Figure 23: Overview of the methodological approach applied in this study.

4.1   Field work

Two Ice Complex sites were defined as key sites for mapping thermo-erosional features during the expedition in August 2014. This study focuses on thermal erosion under recent relief-shaping conditions. Thus, only relief entities that indicated contemporary erosion in terms of exposed bare soil were considered as erosional features. These manually mapped erosional features served as a validation database for automated GIS-mapping (Section 4.3.1). Furthermore, fieldwork included a detailed description of geomorphological characteristics and relief conditions. Observed environmental processes at recently eroding sites provided a valuable input for determining the potential influencing parameters in the logistic regression model (Section 4.3.2). Exemplary discharge and water temperature measurements were conducted on accessible locations at the outlet of streams with different sizes in order to investigate the relationship between these parameters and the occurrence of erosion.

The remote sensing datasets (Section 4.2) require georeferencing with known in-situ coordinates (Ground Control Points, GCPs). 22 GCPs were collected from stable immobile surface features like the intersections of polygonal troughs using a tachymeter (Zeiss ELTA C30), handheld GPS (Garmin GPSmap 62stc) and real time kinematic differential GPS (Leica Viva GNSS GS10) instruments. The absolute accuracy of each survey method ranges from c. 2 cm (differential GPS) to ≤ 1 m (tachymetry) and c. 2 m (averaging function of handheld GPS).

4.2   Remote sensing data processing

4.2.1   Raw DEM processing with GeoEye-1 data

The accurate identification and mapping of thermo-erosional features requires both very high resolution planimetric (2D) and topographic (3D) data. The commercial satellite GeoEye-1 delivers one of the world's highest-resolution satellite imagery products available. It has a pansharped resolution of 0.5 m and provides four channels (RGB, NIR). The concept of stereophotogrammetry allows the extraction of height information from overlapping satellite images (stereopairs), enabling the generation of high resolution DEMs. A raw DEM for Sobo-Sise Island was built by Günther (2015; unpublished data) from three GeoEye-1 stereopairs recorded on July 27 (West), August 15 (Central) and August 24 (East) 2014 (Figure 24). The raw DEM was georeferenced with the 22 GCPs collected during expedition with a RMSE of 1.58 m (Günther 2015, pers. comm.).

_____

Stettner (2015) described the process of DEM generation with PCI Geomatica in detail. It consists of

1)   converting the raw images into epipolar pairs;

2)   extracting DEMs from the overlap between the epipolar pairs;

3)   geocoding the epipolar DEMs on the basis of the geometric model and merge them into one DEM; and

4)   editing poorly correlated areas in the DEM.

Steps 1-3 were performed by Günther (2015; unpublished data). This study focused on DEM editing and refinement (Step 4), which is described in the following section.
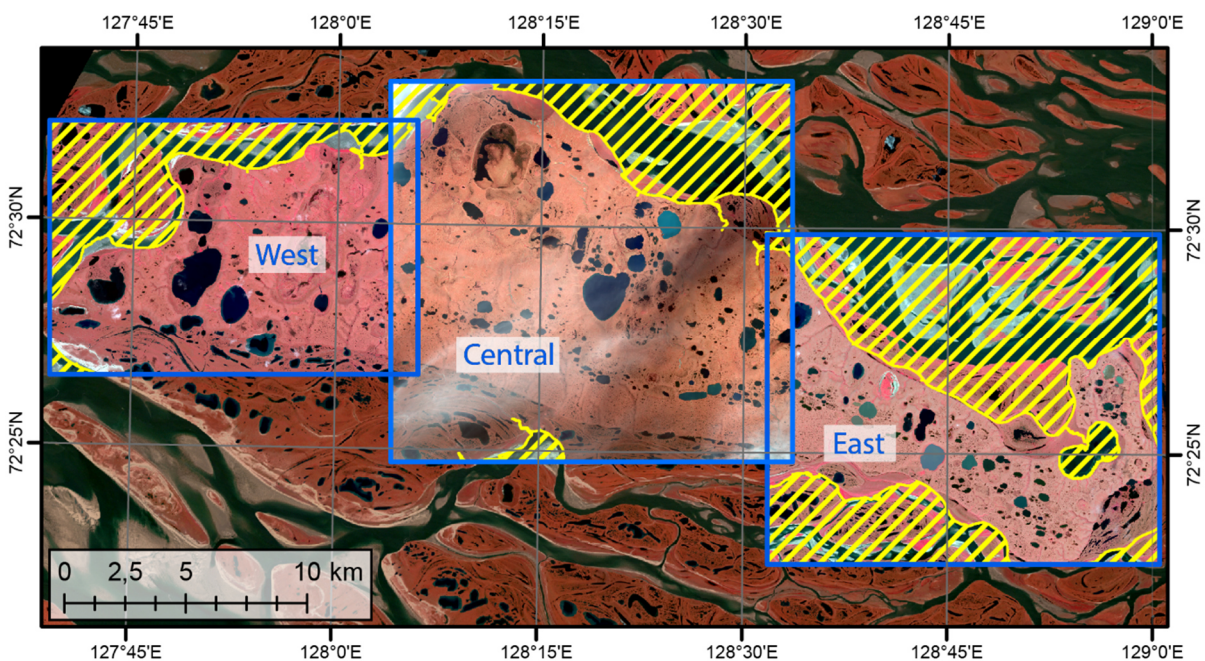
Figure 24: Extent of acquired GeoEye-1 stereopairs (Channels 4-3-2). River branches and enclosed floodplains that were excluded from DEM generation (see Sec. 4.2.2) are illustrated in yellow. Background: RapidEye, 30 June 2014, Channels 5-3-2.

4.2.2 DEM editing

The raw DEM contained pixels with failed or incorrect values (artifacts) which had to be replaced. As a first step, two filters were applied on the entire DEM to remove noise (i.e. single or few spurious pixels). The first filter calculates the average and variance in a 9 x 9 moving window excluding failed and background pixels. If the center pixel had a higher deviation than two standard deviation from the mean, it was set to NoData. The second filter counted the

number of NoData-values immediately surrounding each pixel. If five or more NoData-pixels border the center pixel, the center pixel is also set to NoData. These failed pixels are then interpolated using an inverse distance algorithm from the surrounding pixels (PCI Geomatics, 2014).

Small parts (< 2 %) of the island were covered by clouds and cloud shadows. The DEM showed a strongly undulated surface in these areas. Elevation values diverged highly due to failed image alignment during DEM generation. The stereopairs were visually checked for the occurrence of clouds and affected areas were masked in the DEM. These masks were filled by an alternating application of the "Remove Bumps" and "Remove Pits" filter. Both filters use a 7 x 7 kernel, which replaces all pixels with slope gradients greater than 5 % (PCI Geomatics, 2014).

Water bodies were the most common source for artifacts. The image matching process produced an inconsistent value range for lake pixels differing strongly both in positive and negative direction from the mean expectable lake elevation. This was due to the special topographic and illumination conditions on lakes and lake boundaries. The spectral properties of lakes were observed to be generally homogeneous within one scene, but their reflectance values changed significantly between two satellite scenes (i.e. within one stereopair) due to the different angle of view of the satellite sensor. Additionally, strong winds caused waves, which biased image alignment results. Aquatic vegetation in the nearshore environment raised the local lake level above the expected mean lake level. Bluffs with a vertical elevation shift of several meters between lake and land surface within few pixels (e.g. where the Sardakhskaya Channel hits the Ice Complex, see Figure 21) are another source of errors. Maximum deviations compared with the "true" elevations occurred where clouds superimposed lakes. These errors were corrected manually.

Open water surfaces showed a very strong absorption of the incident infrared radiation. In the greyscale infrared band of satellite data, water bodies appeared very dark in contrast to the highly reflecting surrounding vegetation and soil cover (Pietroniro et al., 2005). For each stereopair, one scene was selected for a threshold-based reclassification of the infrared channel. The values of the Digital Number (DN) in the infrared band of each scene differed slightly due to haze in the atmosphere. Hence, the water threshold values had to be defined by visually and for each scene separately (Table 2).

_____

Covered part of GeoEye-1 scene	DN value for greyscale threshold in the IR band
West	0 – 210
Central	0 – 260
East	0 – 220

First, each satellite image selected for masking had to be orthorectified using the cloud corrected DEM as topographic input. This step ensured the correct planimetric position of the water mask on the DEM surface. After the threshold application, the raster datasets were converted into a polygon vector shapefile. All water bodies with sizes smaller than 100 m² were removed, because misaligned elevation values only occured on water bodies with markedly larger aerial extents. The polygons were converted to lines to deviate the shore line of each lake. Subsequently, the mean elevation of each shore line and its standard deviation were calculated. The difference of these values was assigned to each polygon. Finally, the elevation of each polygon was burnt as a constant value into the DEM. The intention for this process was to create a hydrologically correct DEM. If simply the mean values of each polygon would be considered, "lake plateaus" (i.e. lake levels situated above the surrounding land surface) could have been created resulting from wrong interpolation in the nearshore zone. Instead, the method chosen here suggested the following. If elevation values of the shoreline pixels only fluctuated within a small range (within a decimeter range) around the mean, the lake level were decreased only by this comparatively small value. If shoreline pixels showed a strong deviation from the mean (in the order of several meters), they were burnt by far deeper into the DEM. However, in terms of watershed analysis, hydrological correctness was still guaranteed, as each lake should represent a depression in the surrounding landscape (for an example see Figure 25).

The final water mask for Sobo-Sise Island contained more than 12.000 water bodies. The automatic procedure presented here reflected a cost-benefit effective approach as ground based water level measurements seemed to be an inadequate effort.

A second mask comprised the Lena River branches and was developed based on three orthorectified GeoEye scenes where the elevation was set to zero. This adjustment ensured the correct planimetric representation of river pixels, which are supposed to exhibit the sea level (i.e. a total elevation of 0 m.a.s.l.). The Lena River Delta channels were extracted using the same grey-value threshold values as presented in Table 2 and were burnt into the DEM using a constant elevation of 0 m.a.sl.
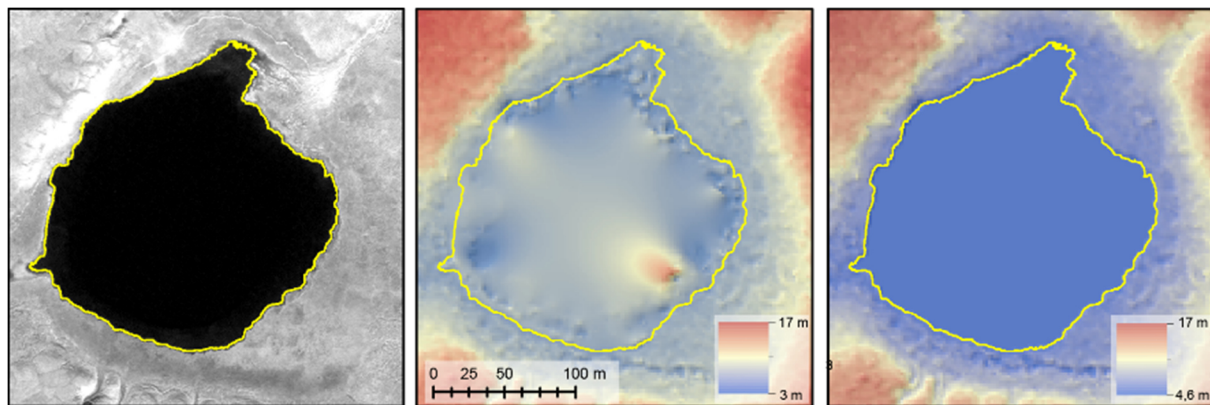
Figure 25: Masking of a medium sized lake in the western part of Sobo-Sise (scale in the middle picture applies to all). The yellow polygon represents the lake mask derived from the threshold-based reclassification of the NIR-band (left). The mean elevation of the shoreline (6.27 m a.s.l.) and its standard deviation (0.59 m) were extracted from the raw DEM (middle) and their difference (5.68 m a.s.l.) was burnt as the new elevation into the DEM (right). Note the strongly undulated lake surface (middle) which is now replaced by a constant elevation value (right).

The southern and northern margins of the central scene were covered by dense cirrostratus clouds, which prevented greyscale thresholding. Major lakes were manually masked and filled by the procedure described above.

The resolution of the final DEM is 2.0 m. Its validity was checked visually due to missing comparative data in this region and showed a satisfying level of detail. All steps were performed using the packages Ortho Engine and Focus of the software PCI Geomatica (V. 2014).
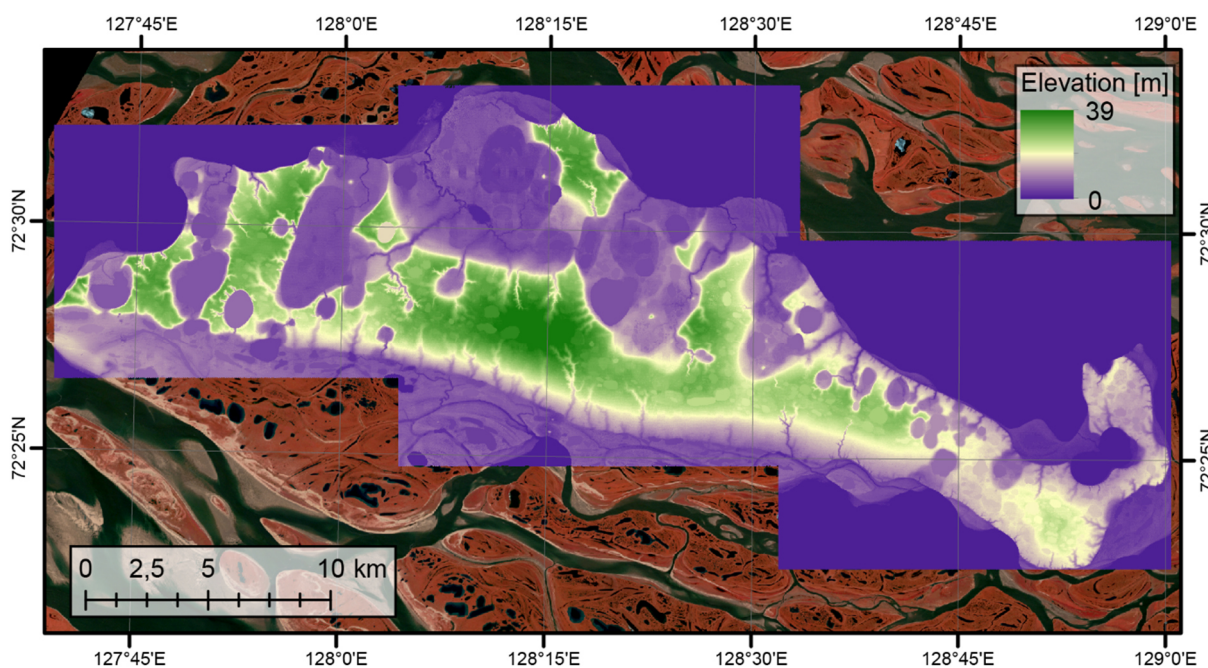
Figure 26: Final DEM of Sobo-Sise. Background: RapidEye, Channel combination 5-3-2.

4.3   GIS datasets

4.3.1   Mapping of thermo-erosional features

Thermo-erosional features showed a small scaled, heterogeneous pattern in the study sites. Manual mapping of each single erosional patch, especially along the numerous baydjarakh outcrops, proved an inefficient and subjective, user-dependent task. Instead, the Modified Soil-Adjusted Vegetation Index (MSAVI) was chosen as an alternative approach for automatic mapping of erosion pixels. The MSAVI (Qi et al., 1994) was developed to increase the vegetation signal while reducing the influence of the soil background to a minimum. It is a further development of the Soil Adjusted Vegetation Index (Huete, 1988), as it replaced the former user-defined soil brightness factor *L* with a self-adjusted *L*-factor. MSAVI depends on the relationship between the red and near-infrared channel and was calculated from the GeoEye-1 scene following the equation (1).

$$MSAVI = \frac{(2 * NIR + 1 - \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - RED)})}{2}$$

The index value range lies between -1 and 1, where pixels with vegetation (higher reflectance in the NIR channel) strongly tend towards 1.

The soil signal in the tundra vegetation cover is comparatively high due to active cryoturbation (Sec. 3.1.4). Two well-identifiable classes remained in the adjusted MSAVI maps that could be clearly distinguished by a threshold: vegetation and, consequently, no vegetation. The no-land-use-class could only consist of water and bare soil, since there exists no land use on the Yedoma uplands.. All pixel values in the MSAVI maps below an empirical threshold (Table 3) were assigned to this class and clipped with the water and cloud mask, generated in Sec 4.2.2. Different thresholds were necessary due to emerging vegetation between the image acquisitions, which changed the reflectance values within the satellite bands.

The remaining soil pixels showed a good agreement with the conditions in the field (Figure 27). In Sobo West, the frequency of erosion cells was $n_{erosion} = 24101$ and $n_{non-erosion} = 1282498$ for non-erosion cells, respectively. For Sobo North, the corresponding values were $n_{erosion} = 4877$ and $n_{non-erosion} = 484452$. Thus, the relative frequency of erosion was 1.845 % for Sobo West, 0.997 % for Sobo North and 1.614 % for both study sites.
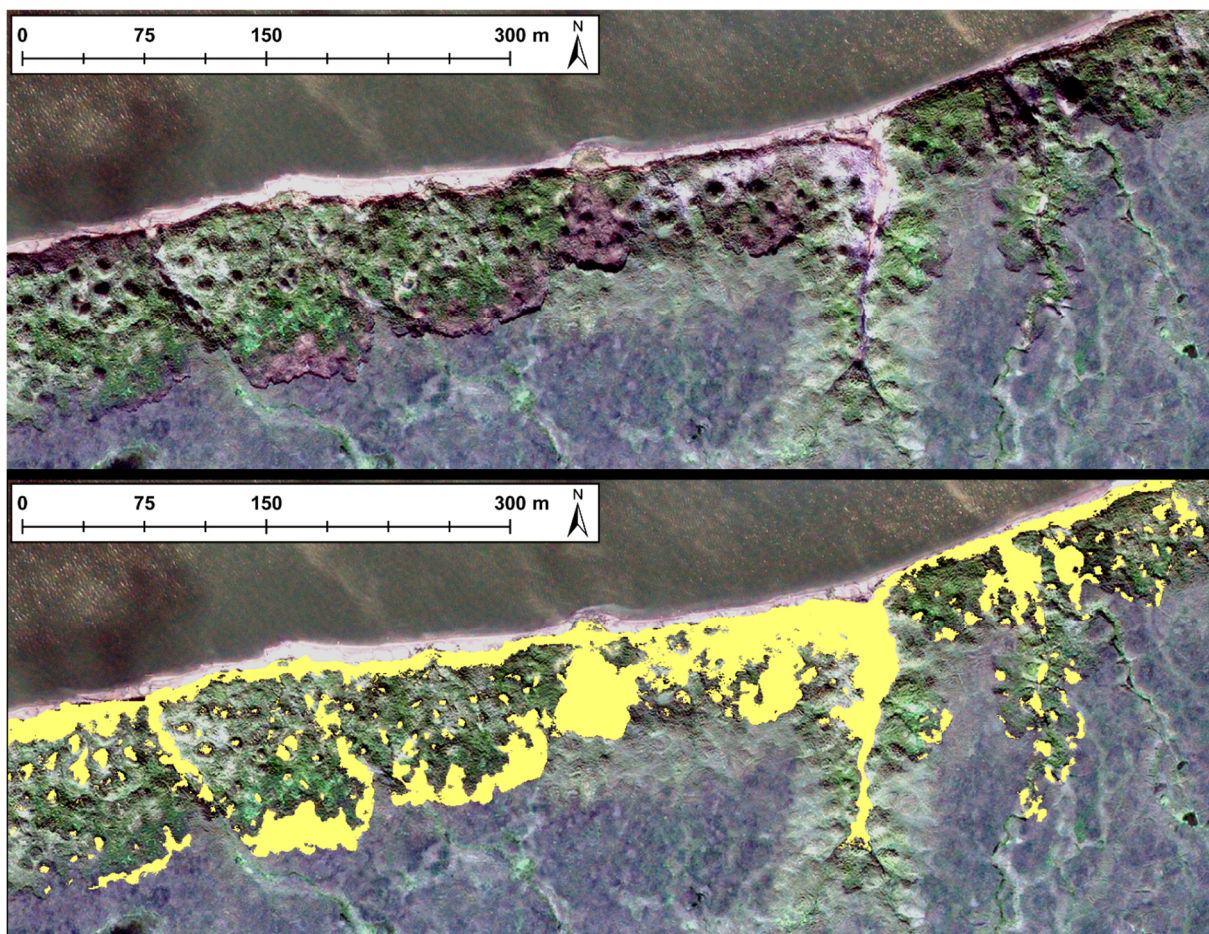
Figure 27: Example for semi-automatic mapping of thermo-erosional features using MSAVI thresholding along the Ice Complex bluff in the western test site. Notice the good distinction between the stabilized slopes with green vegetation and erosion on baydjarakhs and on the headscarps of the bluff.

Table 3: Thresholds in GIS data sets.

Covered part of GeoEye scene	MSAVI	NDVI	NIR channel
	Detection of erosional features	Detection of water tracks/ moist surfaces	Delineation of open water bodies
West	0.78	0.75	0 – 400
Central	0.62	0.68	0 – 440

4.3.2   Selection of environmental parameters

The purposeful selection of environmental parameters is a prerequisite for the generation of a logistic regression model (Hosmer Jr et al., 2013). This step required a detailed conceptual knowledge about the environmental system and involved processes therein and, furthermore, a precise spatial calculation and quantification of the latter. The chosen parameters had to comprise *all* different natures of the phenomenon (topographic, hydrological, climatic, etc.), which were expected to affect its occurrence (Van Den Eeckhaut et al., 2006). The parameters chosen in this study were the following:

1)   Contribution of surface open water

2)   Contribution of water tracks

3)   Slope

4)   Profile curvature

5)   Relief ratio

6)   Potential incoming solar radiation

7)   Snow accumulation

In situ measurements of these variables were impossible, as the distribution of these data was spatially continuous across the study sites. They were derived from the 2m-DEM and GeoEye-1 data either directly or via proxies. Erosion susceptibility mapping on this high level of detail enabled a very fine spatial discrimination of the influence for each parameter on pixel scale. However, this came with the price of increasing computation time.

Contribution of surface open water

Costard et al. (2007) and Wobus et al. (2011) observed increasing thermo-erosion rates along perennially frozen riverbanks and coastlines as a consequence of rising water temperatures due to climate change. Laboratory experiments on ice-rich soils and mathematical modeling (Costard et al., 2003; Dupeyrat et al., 2011; Randriamazaoro et al., 2007) supported the empirical observations. Randriamazaoro et al. (2007) tested ablation rates by raising ground ice temperature, discharge and water temperature, and found the latter to be the predominant parameter.

Warm running surface water along Yedoma landscapes was considered as a driving agent for thermal erosion in this study, though it became apparent that single water temperature measurements at the stream outlets during the expedition did not represent the whole season. A

_____

high degree of surface disturbance (gullying, block failure, in channel erosion) was explicitly observed in thermo-erosional valleys where water from the upper drainage basins accumulated. The evaluation of satellite and DEM data demonstrated an increased surface cover by open water bodies like thermokarst lakes and polygonal ponds in these catchments. Temperature measurements in randomly selected waterbodies proved that they partly heat up to more than 15°C on warm summer days. Provided the coupling to the stream network, the assumption was that the higher amount of contributing open warm surface water, the likelier is the occurrence of erosion along the drainage pathway.

The spatial dataset for this variable required preprocessing of both the DEM and satellite imagery. In order to provide a continuous flow from the waterbody to the stream outlet, spurious or artificial (cf. the burnt lakes) sinks had to be filled in the DEM using the algorithm of Wang and Liu (2006). While water bodies < 100 m² were not considered in the water mask during DEM editing (Sec. 4.2.2), they were now re-included to form a grid of discharge initiation cells. The surface area of each pixel (4 m²) was routed and cumulatively accumulated downstream using the Multiple Flow Direction algorithm (MFD; Freeman, 1991) as it was recommended by Erskine et al. (2006) for comparable terrain. The convergence factor which limits flow partitioning was set to 4 according to Holmgren (1994). The resulting grid differed from a "common" flow accumulation grid, because a stream network could only evolve from predefined input cells. Thus, only those cells that were coupled to the modeled stream network cells were assigned a positive value, all other decoupled cells contained the value 0 (i.e. no flow accumulation).

Contribution of water tracks

A further insight from fieldwork was the presence of a complexly branched, capillary system of small water tracks on the Yedoma uplands. These water tracks flew along the polygon structure, but they did not cut into the surface. Thus, they could not be interpreted as a stream network in its original sense, consisting of visible running water in a well-developed valley. These water tracks concentrated at some point in the relief, which caused thermal erosion and the formation of an incised valley. Vital, dense vegetation covered the shallow streambeds of the water tracks. The same type of vegetation stabilized the slopes of distinct valleys where the active layer was supersaturated due to both thawing subsurface permafrost and contributing water tracks.

Both origins of dense vegetation cover were grouped into one class, which explained the effect of "non-visible" surface water on thermal erosion. The Normalized Difference Vegetation Index (NDVI) proved an effective proxy to detect affected surfaces. This index uses the rationale that vital green plants highly absorb visible red light and, while they strongly reflect near infrared (NIR) radiation. It is calculated by the following band ratio:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

In this study, vital vegetation in water tracks and on slopes could be clearly distinguished from the surrounding tundra vegetation, because NDVI values increased with higher vitality. Further data processing followed the approach as described for *Contribution of surface open water*. All vegetation pixels below an empirical threshold (Table 3) were discarded and the surface area of the remaining pixels was accumulated downstream.

Slope

Slope determines the rate of change of elevation in the direction of the steepest descent. Slope is a key component in landscape evolution, because it affects the velocity of both surface and subsurface flow, and thus, the erosion potential (Bou Kheir et al., 2008; Conforti et al., 2011; Valentin et al., 2005). The slope map was directly derived from the DEM using the algorithm of Zevenbergen and Thorne (1987).

Profile curvature

Profile curvature describes the rate of change of slope gradient parallel to the direction of maximum slope. It is a proxy for local flow acceleration and thus, for the potential of erosion (Wilson and Gallant, 2000). In terms of thermal erosion, profile curvature accounts for the mechanical flow power of water, which was not yet considered in this study as a potential influencing factor. The profile curvature map was derived with the same algorithm as for slope, but the input DEM had to be smoothed with a 3x3 median filter. This step was necessary because calculations with the original DEM resulted in a very noisy surface. Positive curvature values characterize local convexity, whereas negative values represent slope concavity.

Relief ratio

The stability of Yedoma slopes is strongly affected by the Lena River. The fluvial erosion at the toe of the slopes causes an increase in stress on the back of the slope, which counteracts

a steady state topography. This process is especially active during spring flood when thermo-erosional valleys are inundated several hundreds of meters in some cases, depending on the local topography and the maximum river level during the flood. The flood removes the sediment in the valleys, resulting in a change of the former stream gradient. The proximity to streams is regarded as a decisive factor in landslide and gully erosion susceptibility research (e.g. (Nefeslioglu et al., 2008; Yalcin, 2008; Yilmaz, 2009). However, this study also considered the difference in elevation between the base level and any specific location within the catchment, because erosion was assumed not to be only a function of distance, but also of potential gravitational energy at this location. The gradient between each grid cell and the base level is calculated using the relief ratio (Rr):

$$Rr = \frac{\delta H}{\delta L}$$

where δH represents the elevation difference between the specific location and the outlet. δL is the overland flow distance to the outlet. Note that δL is *not* the direct/ Euclidean distance, but the distance of the flow path along the DEM surface from the respective grid cell to the outlet.

Potential incoming solar radiation

Slope asymmetries were observed in thermokarst basins (Ulrich et al., 2010) and in thermo-erosional valleys (French, 1971). Different degrees of slope stabilities were associated with the influence of incoming solar radiation. French (1971) stated that northeast facing slopes in thermo-erosional valleys are most susceptible for solifluction due to lower reception of solar
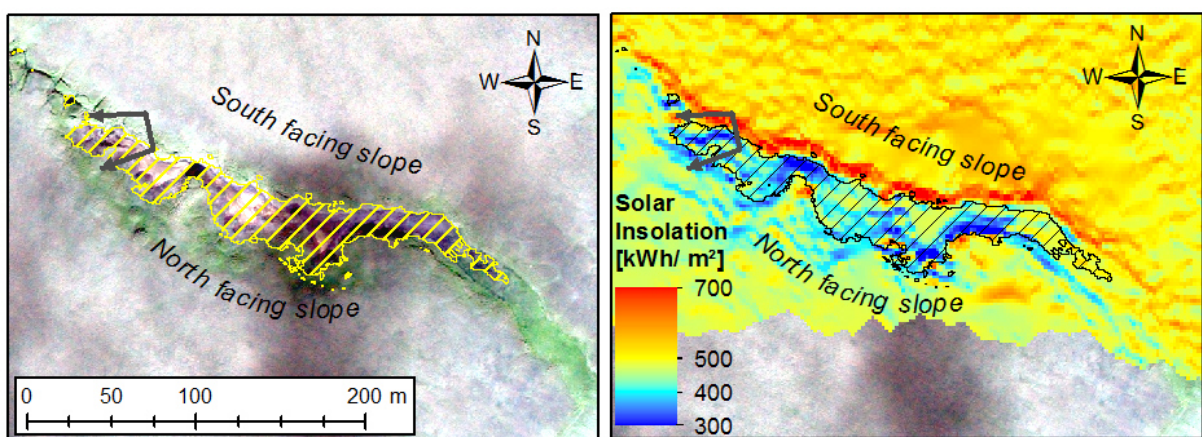
Figure 28: Asymmetric thermo-erosional valley (Sobo West, E in Figure 32). Left: Mapped erosional features above True Color Composite (GeoEye-1). Right: Modelled solar insolation. Note the distinct differences between solar insolation on south and north facing slopes and the dominance of erosion on the north facing slopes. The grey frame represents the angle of view from Figure 30.

radiation. The decreased evaporation on these slopes resulted in higher soil moisture, which promoted erosion. The same effect was observed within the study (Figure 28 and Figure 29).

In this study, the direct potential solar radiation was calculated for each cell of the 2m-DEM, using the algorithm of Böhner and Antonić (2009). Insolation data were modelled for each hour between 01 June and 01 October. The evaluation of long-term climate data from Tiksi Meteorological Station proved that the maximum daily temperature generally exceeded 0°C in this period, which was regarded as a potential trigger for thermal erosion.

Figure 29: Difference in thermal erosion on north and south facing slopes along the valley from Figure 29. Note the snow patch from winter along the north facing slope in this picture, taken on 10 Aug 2014 (A. Morgenstern).

Snow accumulation

French (1971) furthermore postulated that the above-mentioned microclimatic differences on the two slopes could be related to the dominant westerly winds in his study site (Beaufort Plain, N.W.T., Canada). These strong winds caused snowdrift and deposition preferentially on the lee sides of valleys and promote evaporation from exposed slopes during summer season. Woo et al. (1983) noted high snow densities in incised valleys and gullys due to snow compaction.

The same process was observed in the Lena Delta, where strong, gusty winds in winter redistributed the snow and caused a low snow cover on the tundra surface (Boike et al., 2013,

2008; Kutzbach et al., 2007; Sachs et al., 2008). Thermo-erosional valleys and lee sides of the bluffs acted as snow traps due to the decrease of wind speeds.

This study considered the effect of snow accumulation as well. It was argued that compacted snow in incised valleys provides a continuous surface moisture beyond the period of snowmelt in early June. Additionally, snow insulates the underlying permafrost from surface freezing during winter. This promotes a consistent surface instability especially on north facing slopes, which are additionally shielded from solar radiation and winds. The presence of snow patches on the north facing slopes in two valleys during the summer expedition supported this idea. Both sites were strongly affected by thermal erosion (see Figure 29 as one example). Several physically-based models have been developed to calculate the spatial distribution and accumulation of snow in Arctic terrain (Essery et al., 1999; Liston and Sturm, 2002; Pomeroy et al., 1997). Determining their input parameters would require detailed in situ measurements, which is why the model presented here is very simplified. The assumption was that snow accumulates on lee sides and is blown away from windward sides. The Windward/ Lee Index implemented in SAGA GIS (Böhner and Antonić, 2009) calculates this wind effect for a given
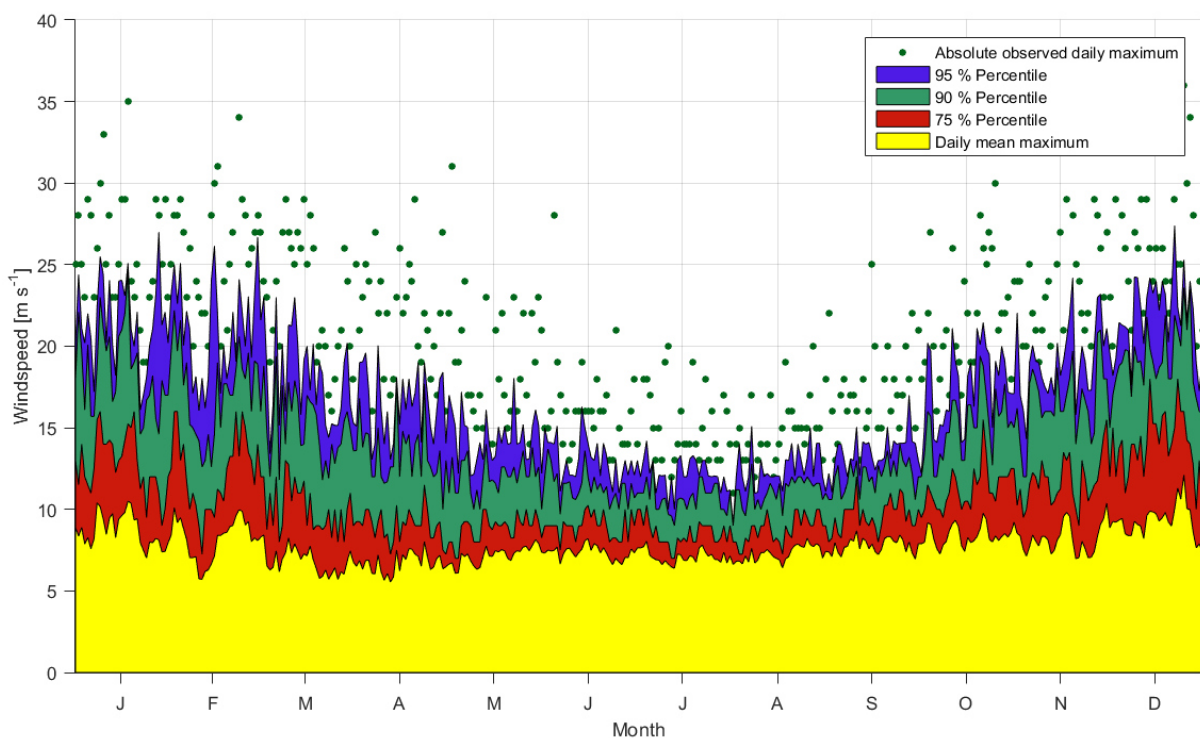
Figure 30: Daily maximum wind speeds at Tiksi Meteorological Station between 1966 and 2014. The mean daily maximum wind speed shows a rather constant variability during the year whereas the gusts (represented by the percentiles and the absolute observed daily maximum) significantly increase during winter.

topography using a DEM and a constant wind direction as input parameters. Values below 1 indicate wind shadowed areas, whereas values above 1 indicate areas exposed to wind. The Tiksi Meteorological Station provides a full set of wind speeds and directions from 1966 to 2014. It was evident that wind speeds were higher in the winter season, underlining the influence on snow redistribution by wind gusts (Figure 30). Winter was
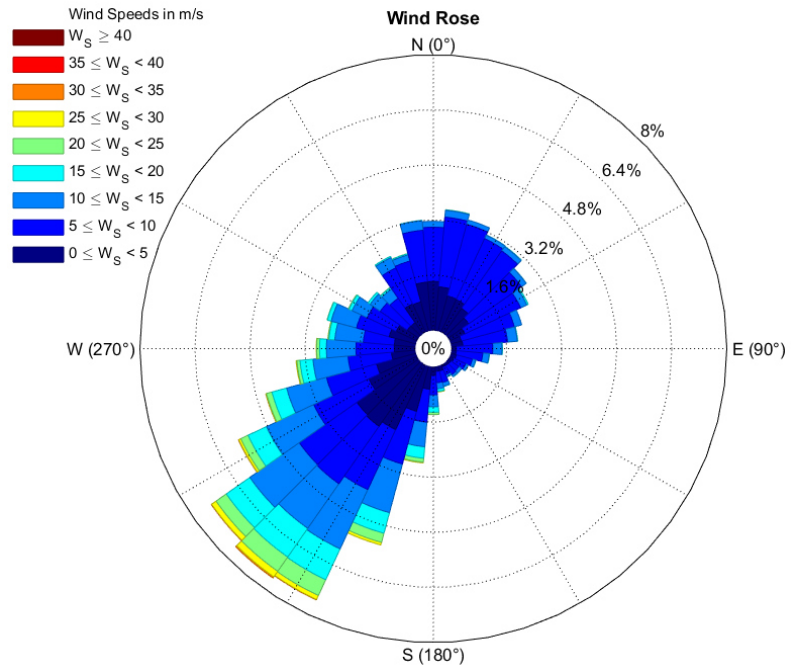
Figure 31: Maximum wind speeds and wind directions for Tiksi between 1961 and 2014. All pairs in this period, measured in 3h intervals, were taken into calculation. This explains the slight discrepancy between the yearly average (221°) and the winter average (264°).

defined as the period where air temperatures dropped below 0°C, which was on a long-term average between 01 October and 01 June. The average wind direction in winter of 264° (WSW; Figure 31) was used for running the model.

All environmental parameters were calculated in ArcGIS (V. 10.3) and SAGA GIS (V. 2.1.4).

_____

4.4 The susceptibility model

4.4.1 The principle of logistic regression

This study used binary logistic regression to measure the relationship between a categorical dependent variable and a set of independent variables. In logistic regression, the dependent (or response) variable is binary or Bernoulli distributed. A grid cell i ($i \in \{1, ..., n\}$ with n as the number of observations could either be affected by erosion ($Y_i = 1$) or not ($Y_i = 0$). The independent variables can have different scales of measurement (categorical or continuous); in this case, all environmental parameters were continuous. The aim of logistic regression is to construct a function, which predicts the probability for an event using a linear combination of the predictor variables (Menard, 2010).

Referred to the present study, a logistic regression model was not used to calculate the expected value of the response variable for every grid cell i ($Y_i = 0$ or $Y_i = 1$), but to model the *probability* for erosion for each cell to be affected by erosion: $p_i = P(Y_i = 1)$. This is achieved by implementing a link function, which specifies a function of the expected value of Y. The use of this link function represents a flexible generalization of linear models, called generalized linear models (GLMs). A GLM generalizes ordinary regression models, because it allows the response variable Y to have a distribution other than the normal (e.g. in this study binomial distribution) (Agresti, 2007).

In the following, the aim is first to demonstrate properties of ordinary regression models, since GLMs represent an extension of ordinary linear regression models. A simple univariate logistic regression model is defined as

$$Y = \beta_0 + \beta_1 x$$

or

$$E(Y) = \beta_0 + \beta_1 x$$

1) Y_i represents the response (or dependent) variable with a normal distribution.
2) X is the explanatory (or independent) variable, which predicts the dependent variable. It is linear in the parameters $\beta_0 + \beta_1 x$, with β_0 representing the intercept or the constant of the equation.
3) The link function represents the mean μ of the probability distribution of Y, by $\mu = E(Y)$. The link function defines a function that relates the mean to the linear

predictor. Here, it is the Identity Link, $g(\mu) = \beta_0 + \beta_1 x$. It is called *Identity* because the mean is modelled directly, thus it is identically with the linear combination of regression coefficients and variables (Agresti, 2007).

In binary logistic regression, the response is binary by 1 (here: erosion) and 0 (here: no erosion). The distribution of Y is defined as the probabilities $P(Y = 1) = \pi$ for success (erosion) and $P(Y = 0) = (1 - \pi)$ for failure (no erosion). Its mean is $E(Y) = \pi$. While the relationship between predictors and response variable is linear in ordinary regression, the response in the binary logistic regression model is non-linear, because Y is bounded between 0 and 1 (Menard, 2010). The value of π can vary as the value of x changes. π is expressed by $\pi(x)$ to demonstrate its dependence on from the explanatory variable.

In ordinary regression, $\mu = E(Y)$ is a linear function of x. For a model with a binary response, the corresponding expression is

$$\pi(x) = \beta_0 + \beta_1 x$$

This is called a linear probability model, because the probability of success changes linearly in x. The parameter β_1 represents the change in the probability per unit change in x. This model represents a GLM with a binary response variable and an identity link function (Agresti, 2007). Again, the problem is that this linear function predicts probabilities of $\pi(x) < 0$ and $\pi(x) > 0$. The step for solving this problem is to replace the linear response variable $\pi(x)$ by a link function, which represents the chances or *odds* for being a success: $\frac{\pi(x)}{1 - \pi(x)}$. Unlike $\pi(x)$, the range of the odds can take values towards positive infinity, but are have a minimum value of 0. Another transformation of this term generates a variable that also covers the range from 0 towards negative infinity. The natural logarithm of the odds is called the logit link function

$$g(x) = \log\left(\frac{\pi(x)}{1 - \pi(x)}\right)$$

The logit becomes negative and larger in absolute value as the odds decrease from 1 towards 0, and becomes increasingly large in the positive direction as the odds increase from 1 towards infinity. Using the natural logarithm of the odds that $\pi(x) = 1$ as the dependent variable circumvents the problem that the estimated probabilities might exceed the range between 0 and 1 (Menard, 2010).

The relationship between response and explanatory variable is

_____

$$\log\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \beta_0 + \beta_1 x$$

The logit can be converted back to $\pi(x)$ by exponentiation, which represents the probability of success (Menard, 2010)

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

This study represents a multivariate case, where the probability of erosion is explained by a set of several predictors $(x_1, x_2, \dots, x_k)$, as described in the previous section. The resulting probability for erosion $\pi(x)$ is linear in the combination of these variables with their estimated regression coefficients $(\beta_0, \beta_1, \beta_2 \dots \beta_k)$:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x \beta_0 + \beta_1 \, x_1 + \beta_2 \, x_2 + \dots + \beta_k \, x_k}}{1 + e^{\beta_0 + \beta_1 x \beta_0 + \beta_1 \, x_1 + \beta_2 \, x_2 + \dots + \beta_k \, x_k}}$$

The β coefficients are derived by a maximum likelihood estimation. This method maximizes the value of the likelihood function, which indicates how likely it is to obtain the observed value of the response variable, given the values of the explanatory variables and their coefficients $\beta_0, \beta_1, \beta_2 \dots \beta_k$. Details how to construct the likelihood function for binary logistic regression are given e.g. in (Agresti, 2007; Hosmer Jr et al., 2013).

4.4.2 Stratified sampling and multicollinearity analysis

Within the study sites, erosion and non-erosion cells were by far not equally distributed (ratio 1: 61), which is described as a rare-event dataset. (King and Zeng, 2001) stated that building a model with this ratio causes an underestimation of the probability for the erosion events. They gave recommendations how to correct the dataset for these rare events. One suggestion was to take stratified random samples with a defined ratio of erosion and non-erosion cells, which was also applied in this study.

Heckmann et al. (2014) highlighted the two most important prerequisites for logistic regression that are

1) the independence of the observations and;
2) uncorrelated explanatory variables.

cf. 1) The independence of observations must be neglected for both response and explanatory variables, because spatial data are highly autocorrelated. This is due to the fact, that

neighboring cells show very similar or equal values. Thus, if a model only is built only from adjacent cells containing the same information, the independence of each observation cannot be maintained. To encounter this problem, a stratified random sampling approach was chosen (cf. Van Den Eeckhaut et al., 2010) using the strata algorithm in R's sampling package (Barcaroli, 2014). A grid with a cell size of 10000 m² (100 m resolution) was imposed on the original dataset (2 m res./ 4 m² cell size). Both erosion and non-erosion cells existed in 90 grid tiles. From each of the 90 grid tiles, one erosion cell and four non-erosion cells were randomly selected together with the values of the respective environmental variables in this cell, resulting in a total sample size of 450 cells (90 events vs. 360 non-events). Heckmann et al. (2014) observed in a comparable dataset, that this ratio produces the lowest diversity in the model (i.e. the lowest number of included environmental parameters after stepwise selection; see Chapter 4.4.3). The total sample size avoided the sample being to large (which causes overfitting of the model) and was agreement with a sample size of 200– 600 observations recommended by Hjort and Marmion (2008). However, the sample size must still be large enough (Heckmann et al., 2014). Few observations result in high uncertainties during the estimation of the parameter coefficients. Additionally, an insufficient sample size may not cover the full variety of the environmental parameter. Green (1991) analyzed the variability of explanatory variables in regression models and postulated a minimum sample size of $n_{min} > 50 + 8m$, where m represents the number of explanatory variables. Thus, this criterion (here $n_{min}: 50 + 57 = 107$) is also fulfilled. However, note that the chosen stratified random sampling was a simplified approach, which rather minimized than prevented spatial autocorrelation. An alternative sampling strategy is given e.g. in Brenning (2005), who uses correlograms to estimate the sampling range between cells.

cf. 2) The second prerequisite accounts for the independence of the explanatory variables. The presence of correlated parameters in logistic regression hampers the estimation of the model coefficients and increases their variance. A small change in the observations of correlated parameters can cause large errors during the maximum likelihood estimation of the parameter coefficients. Moreover, if an explanatory variable can be directly replaced by another, it becomes unclear, which one to choose as the appropriate model (Agresti, 2007).

First, the correlation between each single environmental parameter was tested by constructing a cross correlation matrix. Each ith observation of any environmental parameter was compared with the ith observation of any other environmental parameter. The respective

coefficient of determination R^2, calculated from the residuals of their linear relationship, was recorded in the correlation matrix. It appeared that the parameters *Contribution of surface open water* and *Contribution of water tracks* were highly correlated ($R^2 = 0.79$), because they were both calculated with the same flow routing algorithm. The initiation cells for the downstream routing may show a different spatial pattern and quantity in both datasets. However, as these cells accumulated in the same main channels, the correlation within the main channels was nearly perfect. Studies that encountered the same problem of parameter correlation decided to combine the correlated datasets into a single parameter. However, the field observation in this study implied that thermal erosion was more likely to be provoked by streams from open water bodies, because water tracks also occurred in stable catchments with no evidence for thermal erosion. Thus, the parameter *Contribution of water tracks* was excluded from further analysis.

Furthermore, correlation can arise from the combination of several parameters, which is called multicollinearity. Perfect multicollinearity occurs when one explanatory variable can be linearly predicted with at least one other explanatory variable. This was checked by calculating the variance inflation factor (VIF). The VIF for each explanatory variable was calculated by setting up a linear regression model of one explanatory variable versus all other explanatory variables. The Coefficient of Determination R_i^2 represents the proportion of variance for each variable i, i.e. how well each tested variable i can be explained by the set of the other explanatory variables (O'Brien, 2007). The VIF for the ith variable was calculated by the formula

$$VIF_i = \frac{1}{1 - R_i^2}$$

The VIF for each environmental parameter is given in Table 4.

Table 4: Calculated Variance Inflation Factors (VIF) for the selected environmental parameters.

Contribution of surface open water	Slope	Profile curvature	Relief ratio	Potential incoming solar radiation	Snow accumulation
1.142	2.793	1.146	1.818	2.995	1.541

Generally, VIFs greater than 10 indicate strong multicollinearity and variables above this threshold should be omitted in logistic regression. This rule of thumb is questioned by O'Brien (2007) who suggests to include variables with even higher VIFs. However, as the highest observed VIF was 2.995 for the variable *Potential incoming solar radiation*, all remaining six

_____

variables were kept for further analysis. VIFs were calculated using the vif algorithm of R's car package (Fox and Weisberg, 2010).

4.4.3   Stepwise selection of environmental parameters

In order to find the decisive parameters, the six environmental parameters were reduced stepwise to a best minimum for generating the logistic regression model. The final model of each run is characterized by a parsimonious amount of predictor variables, which describe the response variable as well as possible. Stepwise reduction of explanatory variables helps to avoid model overfitting. An overfit occurs when many variables predict the occurrence of comparatively few events. The resulting susceptibility map may show a very good success rate for the present study for which the model was fitted, but the model parameters may perform poorly in new study sites (Petschko et al., 2014, referring to Hosmer Jr et al., 2013).

The Akike Information Criterion (AIC) measures the model's goodness of fit, while it penalizes the model for redundant parameters (Akaike, 1974). The AIC is defined as

$$AIC = 2k - 2\log(L)$$

where k is the number of model parameters and L the maximum likelihood of the regression model. The preferred model is the model with the *lowest* AIC value. The selection process starts with a full model including all randomly sampled explanatory variables. Then, the number of model variables is iteratively reduced. A decrease in the AIC value causes the exclusion of another variable (*forward exclusion*). If the AIC value increases, then the variable is resumed (*backward selection)* and another variable is excluded etc. This backward and forward selection procedure continues until the model finally reaches the best combination of maximum likelihood and minimum model parameters. The remaining parameters after AIC-based selection are used for fitting the final model and predicting the probability for erosion in the study sites.

Note that there is no statistical test to compare AIC-values of different models and that parameter selection in this study depends on an information criterion, not a significance based test. However, the AIC has found a widespread application in geoscience. AIC-based parameter selection was used in both forward direction (starting with one single predictor in the model, e.g. Brenning, 2009) and backward direction (eliminating predictors from a full model, e.g. Gorsevski et al., 2006). The stepwise combination of forward and backward parameter selection, as chosen in this study, is the most common method (e.g. Goetz et al., 2015; Lepore

et al., 2012; Petschko et al., 2014; Van Den Eeckhaut et al., 2010). Additionally, Brenning (2005) found stepwise variable selection using the AIC superior compared to other machine-learning classifiers because the latter favored model overfitting.

AIC-based parameter selection was performed using the algorithm stepAIC in the R package MASS (Venables and Ripley, 2015).

4.4.4 The consensus map and the IQR90 map

Constructing a model from randomly sampled data reduces spatial autocorrelation and fulfils the independence assumption of single observations. The resulting model from that sample might have a good fit due to stepwise AIC-based selection of explanatory variables. Nevertheless, the final model itself depends on the randomly selected cells of response and explanatory variables. Applying the same selection procedure on another random sample can lead to the selection of different environmental parameters and different model coefficients when fitting a new model.

Heckmann et al. (2014) and Marmion et al. (2009) recommended to repeat the modeling procedure many times and to construct a consensus map from all susceptibility maps. To avoid coincidentally choosing a model with a poor fit, the sampling, stepwise variable selection, and model fitting was repeated 100 times. Each fitted model was applied on the data from its selected environmental parameters to generate the corresponding susceptibility map for the study sites. From the stack of all 100 susceptibility maps, the consensus map was calculated. This map represents the median of all predicted probabilities for each cell in the study site. The uncertainty per cell is indicated by the interquantile range IQR90 map ($p_{0.95} - p_{0.05}$), which reflects 90% of the modelled susceptibility values. The IQR90 is an indicator of model dispersion and quantifies the uncertainty for each raster cell due to the variability of each model run. The analysis of the IQR90 is a helpful tool to demonstrate the spatial distribution of model uncertainty (Heckmann et al., 2014).

The consensus and IQR90 maps were generated with the R package raster (Hijmans and van Etten, 2012). The R script for model generation and susceptibility prediction was partially adopted from Heckmann et al. (2014) and modified with the friendly permission by the principle investigator.

4.4.5 Model evaluation: Decisive environmental parameters

Besides the generation of susceptibility maps, the aim of this study was to detect the decisive environmental parameters for thermal erosion. In this regard, stepwise selection of model parameters was a good indicator about the overall importance of a single parameter. A high significance in the models was attributed to those environmental, which formed part in (nearly) all model runs. Thus, the percentage occurrence of each parameter during the 100 model runs represented one measure of parameter importance.

Another possibility of assessing the importance of a parameter is to analyze its standardized coefficient. This study encountered a typical problem in logistic regression, that the explanatory variables do not share the same units and scales. For example, the effect on the dependent variable by a one-unit change of e.g. *Slope* could not be compared with a one-unit change of e.g. *Solar Radiation*, because *Slope* was measured in degree and *Solar Radiation* in kWh per square meter. If the coefficients are standardized, their values can be interpreted in terms of the relative influence of an independent variable on the dependent variable. In this study, the easiest way to standardize the coefficients was chosen, which is to standardize the values of the environmental parameters with $\mu = 0$ and $\sigma^2 = 1$. After conversion, a one-unit-change in a standardized explanatory variable corresponds to a one-standard-deviation change in the original explanatory variable. Consequently, each coefficient in the regression model represents the effect of a one-standard-deviation change in an explanatory variable (Agresti, 2007; Menard, 2004).

The number of the final parameters in each model run varied due to random sampling and stepwise parameter selection. It was not the scope to focus on the precise values of the single parameters. This was not possible, because model parameters do not co-exist consistently in all models after stepwise selection and each run draws on different observations due to random sampling. The aim of standardization was to compare the overall *relative* influence of the different environmental parameters on the dependent variable across the model runs.

4.4.6 Model validation

The produced susceptibility models required validation in order to investigate the agreement of the model results with the real occurrence of erosion. Chung and Fabbri (2003) stated that any prediction model without validation is useless and has hardly any scientific

_____

significance. Validation can be performed by splitting the dataset either randomly or systematically into test and training areas (e.g. Hjort and Marmion, 2008). Another possibility for checking the model transferability is to apply the model in another, spatially and/ or temporally independent dataset (Chung and Fabbri, 2003). The dataset in this study was not partitioned. Only few observations existed for erosion in thermo-erosional valleys, whereas the data basis for erosion along the bluff was estimated to be sufficient. The full dataset was considered for model generation because finding the decisive factors and predicting the susceptibility for the variety of erosion occurrence was one of the main goals of this study. No comparable spatially independent datasets were available with respect to spatial resolution, temporal comparability and ground-truthed erosional features.

Thus, the goodness of fit of each model run was evaluated within the existing dataset to explore the success of the predicted erosion susceptibility with regard to mapped erosional features (Chung and Fabbri, 2003). Hosmer Jr et al. (2013) recommended analyzing the area under the Receiver Operating Characteristic (ROC) curve to interpret model accuracy. ROCs plot the proportion of positive cases (i.e. erosion) correctly predicted *(*≙ true positive rate), versus the proportion of false positives in the total of negative observations *(*≙ false positive rate) (Beguería, 2006). The values below the ROC curve range between 0.5 and 1. A diagonal line in the ROC plot, which indicates complete randomness in predicting erosion and non-erosion events, represents the "null" situation and hence, a very poor predictive power of the model (Menard, 2010). The closer the area under the curve (AUC) converges towards 1, the better is the model's capability to discriminate between correctly predicted erosion and non-erosion cells. The benefit of ROCs is that the user does not have to define a cutpoint, i.e. a distinct threshold, which separates the modelled probabilities into two fixed categories indicating "erosion" and "no erosion". The ROCs and their respective AUC were calculated for all 100 model runs using the R package ROCR (Sing et al., 2005).

5 Results

5.1 Field observations

This section describes the characteristics of thermo-erosional landforms and associated environmental processes of the two study sites, which were observed during the expedition to Sobo-Sise in August 2014. During the field season, the weather changed from dry and hot to more rainy conditions, resulting in a large variability of discharge amount and water temperature. This fact obviates an interpretation concerning the influence of these parameters on erosion. In the following, selected discharge values are presented to provide an overall impression of discharge quantities within the study sites.

5.1.1 Sobo West

A bifurcated, well-developed stream dominated the western part of the study site (Figure 32 A). At the stream outlet, the highest discharge rate of Sobo West was determined ($16.6 \, \mathrm{l \, s^{-1}}$). Erosion occurred in the lower reaches due to accumulated discharge from the

Figure 32: Study site "Sobo West". The capital letters indicate different geomorphological units as mentioned in the text. Numbers in circles indicate the location of the images in Figure 33 (Source: GeoEye-1, Natural Color Composite; contour lines were derived from the generated DEM).

large catchment as well as in the upper reaches, where surface runoff from the Yedoma surface concentrated. The middle reaches of this valley were covered by vital vegetation and showed no evidence for erosion.

Several thermo-erosional gullies cut deep into the bluff along the course of the Lena River (Figure 32 B, Figure 33 ①). They showed a high degree of disturbance with block failure and smaller slumps in their head cut zones. In contrast to the mature valley A in Sobo West, erosion was present along the full length of the gullies. Subsurface flow (piping) led to the formation of ice caves, which collapsed and promoted further erosion. Discharge rates ranged between 0.72 and $2.14\,l\,s^{-1}$. The baydjarakhs along the bluff tended to topple due to their steep faces and undercutting by small streams ($< 0.02\ l\ s^{-1}$ discharge). These runlets converge along the bluff as meltwater from the Ice Complex bluff. The bluff eroded retrogressively towards the Yedoma surface and formed a distinct, concave scarp at the transition between Yedoma surface and the bluff. The turbidity of streams flowing over the Ice

Figure 33: Examples for thermo-erosional landforms in study site Sobo West: ① thermo-erosional gully at the Ice Complex bluff; ② slight incision of water tracks flowing towards the alas; ③ thermo-erosional valley. Locations are given in Figure 32. All photos by A. Morgenstern.

Complex bluff indicated a high degree of sediment suspension, and thus recent thermal erosion.

In the northeastern section, two stabilized valleys (in terms of the absence of contemporary erosion) drained onto a densely vegetated floor of a former lake (Figure 32 C). The cross-sections of the valleys showed a very smooth profile with discharge rates comparable to those of the gullies along the bluff.

Short drainage pathways with very low discharge rates (0.1–$0.5\,\mathrm{l\,s^{-1}}$) flew down the slopes on the eastern margin (Figure 32 D, Figure 33 ②). Their streambeds followed the structure of polygonal tundra and were covered with vital, dense vegetation. Slight incision occurred at the toe of the slopes, whereas the upper slopes showed a straight profile. Another thermo-erosional valley, developed by lake drainage, existed in the southeastern margin (Figure 32 E, Figure 33 ③). Although this drainage had occurred several decades ago, the degree of erosion especially in the upper and middle reaches of the stream was still very high.

Flattened vegetation in the incised head reaches of the valleys A and E implied a long snow cover at these locations. The degree of surface disturbance was also highest at these locations. Erosion occurred predominantly along slopes exposed in northern direction (NW-NE) of these valleys.

5.1.2 Sobo North

Three thermo-erosional valleys discharged towards the floodplain in the western section of the study site (Figure 34 A). They cut deep into the Yedoma surface with evidence for erosion especially in the upper reaches of the valleys. Here, bare soil was visible along the steep
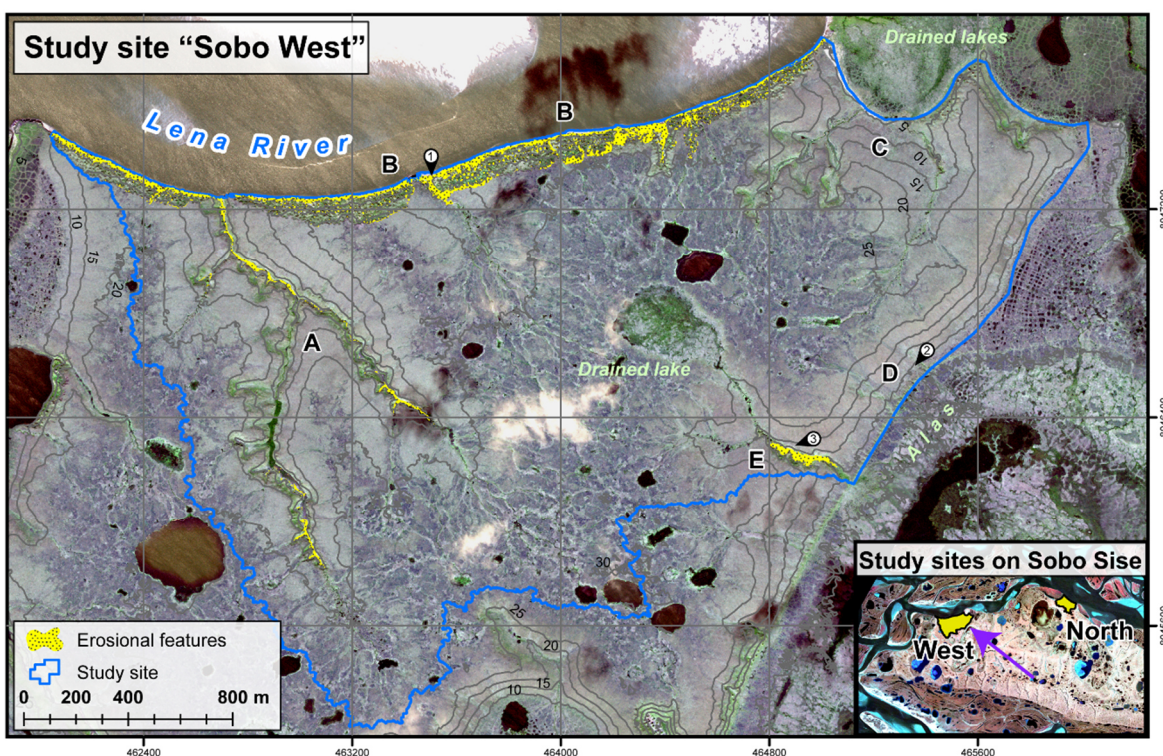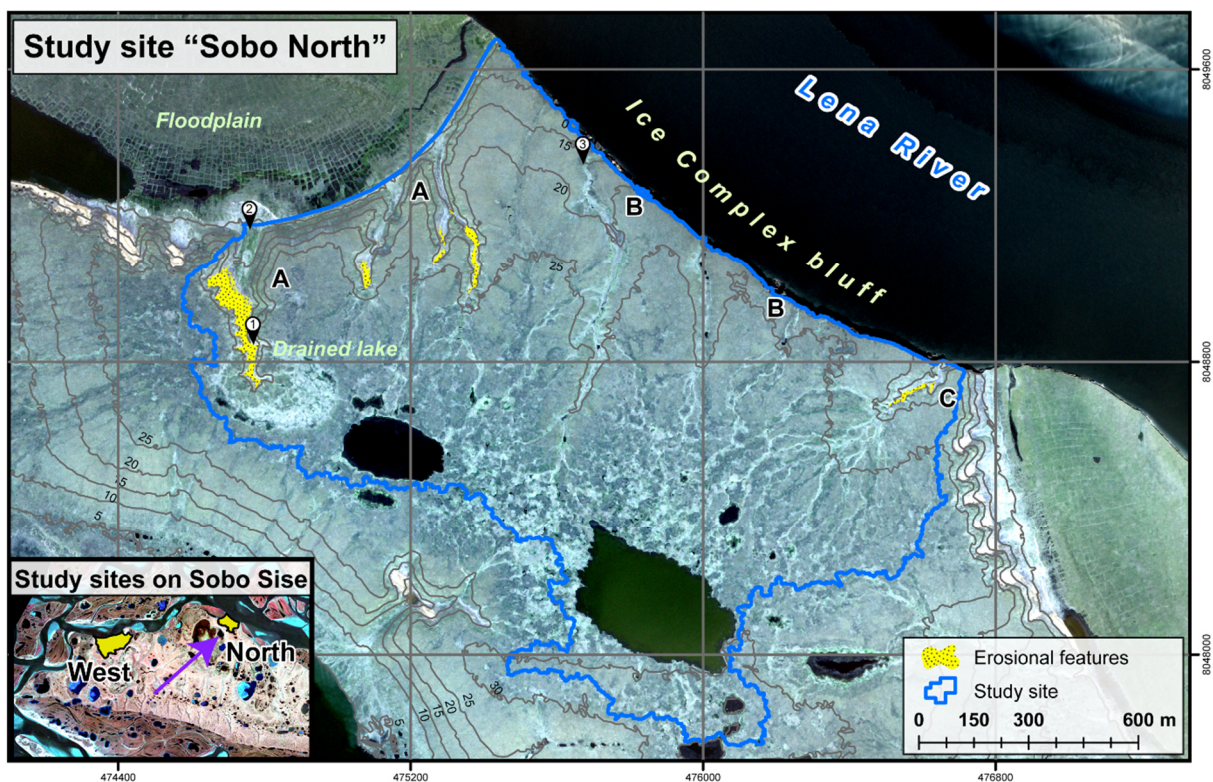

Figure 34: Study site "Sobo North". The capital letters indicate different geomorphological units as mentioned in the text. Numbers in circles indicate the location of the images in Figure 35 (Source: GeoEye-1, Natural Color Composite; contour lines were derived from the generated DEM).

sidewalls. The outlets of those valleys were relatively broad and covered with dense aquatic vegetation. The formation of the westernmost valleys was influenced by lake drainage (see the well-defined basin around the upper reaches of the valley). Its morphometry changed within 100 meters from a V-shaped valley (with active erosion in the head cut zone (Figure 35 ①) towards an U-shaped valley with a more than 50 m broad floor and only little surface water flow (Figure 35 ②).

The streams, which discharged towards the north, showed only little evidence for incision in spite of the presence of a considerable amount of water (0.96–5.95 l s^{-1}) (Figure 34 B). In some places, they dropped more than 20 m as waterfalls over the vertical Ice Complex bluff directly into the Lena River. They showed no evidence for erosion and their streambeds were covered with vital vegetation (Figure 35 ③).

The valley at the eastern end of the bluff reached the Lena River on river level and showed moderate disturbance along its course (Figure 34 C).
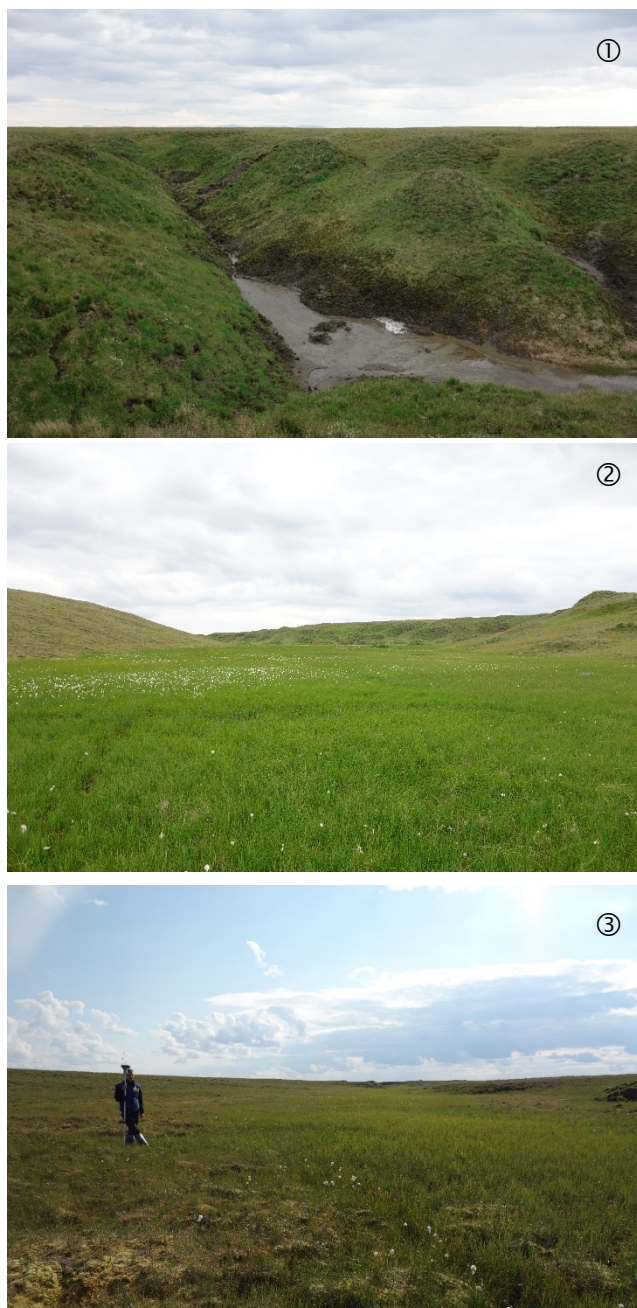
Figure 35: Examples for thermo-erosional landforms in study site Sobo North: ① deeply incised head cut zone after the drained lake; ② dense vegetated streambed of the same valley, ③ slight incision of water tracks flowing towards the Lena River. Locations are given in Figure 34.

_____

Field mapping during the expedition revealed that the key locations for thermal erosion were found

    1)   on very steep slopes along the bluffs, specifically on the protruding baydjarakhs along the bluff;

    2)   in the upper reaches of thermo-erosional valleys, where water from the upper drainage basins concentrates; and

    3)   on steep slopes and in the streambeds of V-shaped thermo-erosional valleys and thermo-erosional gullies.

5.2 Model parameters

This section evaluates the remaining model parameters after stepwise variable selection and model fitting (Sec. 4.4.3). Figure 36 illustrates the distribution of the estimated coefficients of each model parameter for all 100 model runs. The percentages at the bottom part of the figure represent the occurrence of each model parameter after stepwise AIC-based selection divided by the number of model runs. A table of the selected environmental parameters and their estimated coefficients for all 100 model runs is provided in the supplementary material of this study.
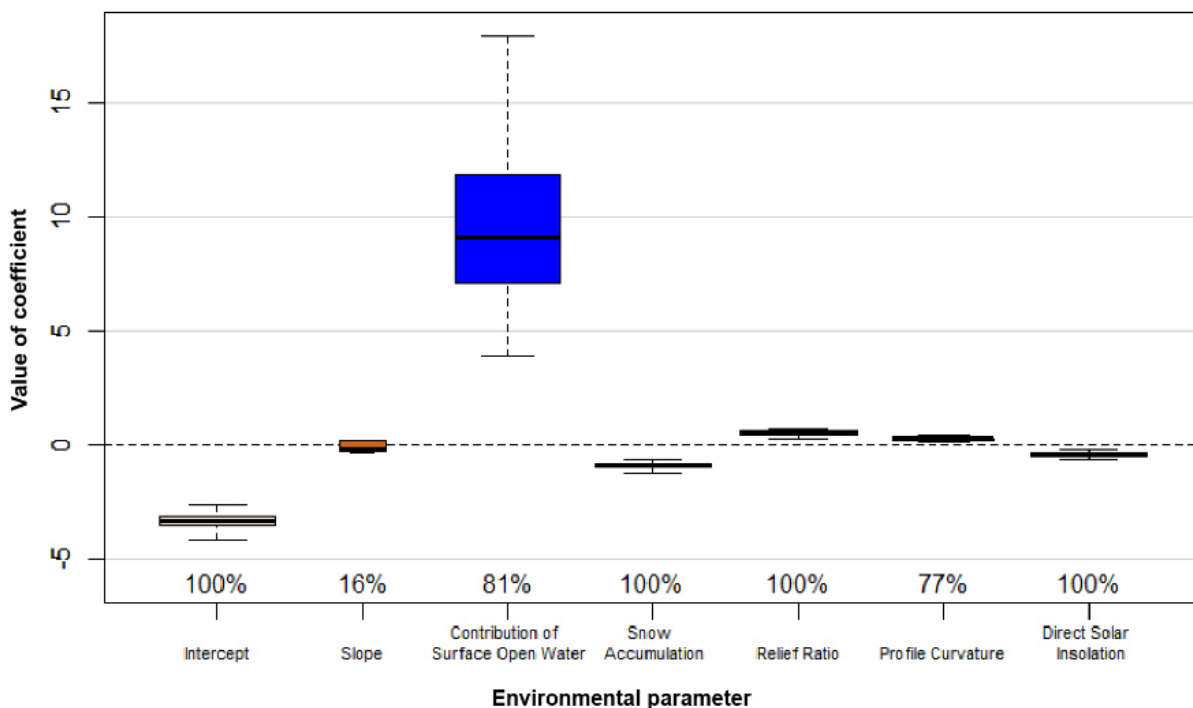
Figure 36: Distribution of parameter coefficients in 100 models. The percentage values above the parameter names indicate the percentage occurrence of the parameter in all model runs. Annotation. Boxes: interquantile range (25–75 % of the total distribution); whiskers: 1.5 * interquantile range; thick horizontal line: median.

The environmental parameters *Snow accumulation*, *Relief ratio* and *Direct solar radiation* formed part in each model run. *Contribution of surface open water* and *Profile curvature* were members in more than three fourths of all model runs, whereas *Slope* played only a minor part in the models (16 %). The range of model coefficients was very close for all model parameters except for *Contribution of surface open water*. A positive influence of the estimated coefficient on the logistic regression model was observed for the parameters *Contribution of surface open water, Relief ratio* and *Profile curvature,* whereas *Snow accumulation* and *Direct solar insolation* had negative coefficients. *Slope* takes both, positive and negative values.

5.3 Consensus and IQR90 maps

The use of stepwise AIC-based model selection reduced the number of environmental parameters to a best-fit minimum. The spatial probability for erosion in each cell within the study sites can be predicted by the cell values of the remaining parameters and their estimated coefficients. Figure 38 and Figure 39 show the spatial probability for erosion in both study sites as a consensus map of 100 model runs (Sections 4.4.3 and 4.4.4). Their uncertainties are represented by the IQR90 maps (Figure 40 and Figure 41), which comprise 90 % of the variability in all model runs. High-resolution maps are part of the supplementary material of this study.

The locations of highest erosion susceptibility in Sobo West, highlighted by reddish to purple colors, are concentrated in the head reaches and streambeds of deeply incised valleys (Figure 38 A & E), as well as in two major and several smaller thermo-erosional gullies along the bluff (Figure 38 B). Thin branches of lower probabilities propagate from major water bodies towards thermo-erosional valleys and gullies. Erosion susceptibility along the Ice Complex bluff is generally high at the toe and moderate (yellow colors) at the upper zone of the bluff. Lower probabilities of erosion (light yellow and bright yellow colors) are modeled for the two stabilized valleys (Figure 38 C) and the slope towards the alas (Figure 38 D). The flat Yedoma uplands showed very low or no susceptibility (light blue colors).

For Sobo North, the highest predicted probabilities occurred in the head reaches of the three contemporarily eroding, deeply incised valleys in the western part (Figure 39 A) as well as in the thermo-erosional valley at the eastern margin of the study site (Figure 39 C). Moderate to high erosion probabilities were assigned to the steep slope between the valleys in the western

_____

part (Figure 39 A) as well as to a valley, which drains towards the Ice Complex bluff (left B in Figure 39).

The predicted erosion susceptibility in both study sites was generally low, but the absolute predicted probabilities differed between the sites. In Sobo West, the mean predicted probability for erosion was 5.8 %. Here, 77.2 % of all cells in the consensus map were assigned a predicted erosion probability of less than 5 %. In contrast, the mean predicted probability in Sobo North was lower (3.4 %). Consequently, the percentage of cells with predicted values lower than 5 % was higher (88.2 %) in the northern study site.

The spatial variability of predicted erosion probabilities per cell, represented by the IQR90 maps, was also low for both study sites. The data values in both IQR90 maps were strongly positively skewed with an average of 0.0423 for Sobo West and 0.0377 for Sobo North, respectively (Figure 37). In other words, the range of predicted probabilities per cell varied on average by about 4 percentage points from the cell mean value in the consensus map. Uncertainties greater than 10 % occurred only rarely. The highest uncertainties were found in the modeled channels from the *Contribution of surface open water* dataset (Figure 40 and Figure 41).
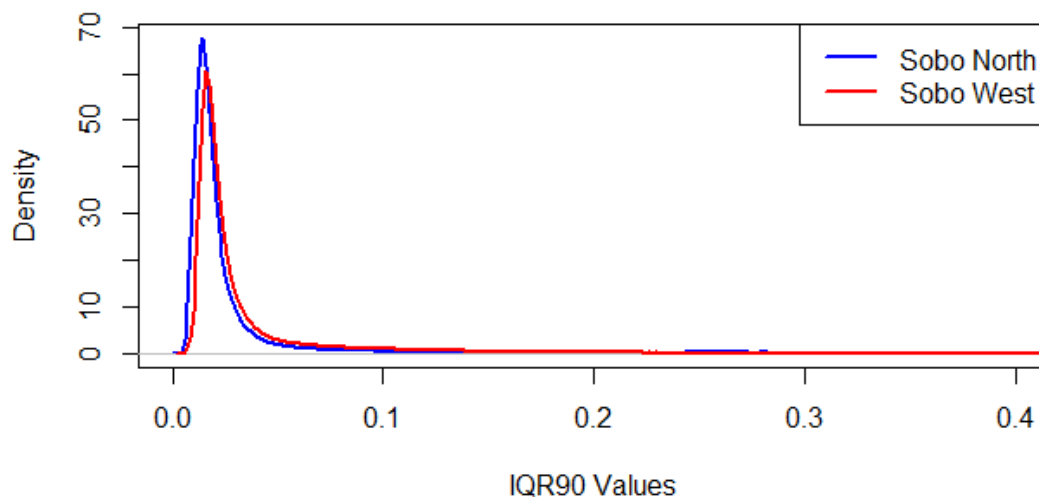
Figure 37: Probability density distribution of the IQR90 values for the study sites Sobo West and Sobo North.
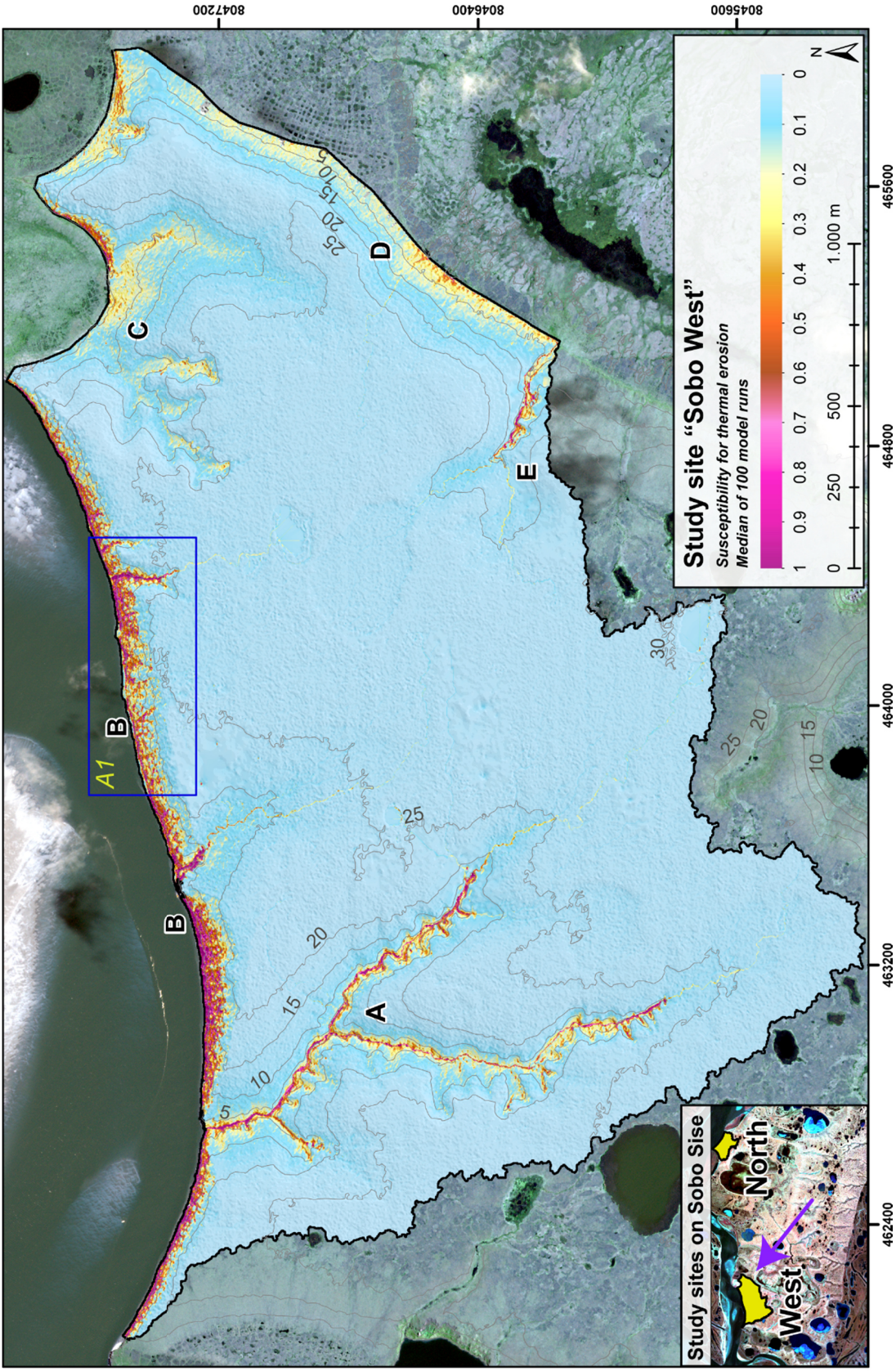
Figure 38: Consensus map for Sobo West. The capital letters A-E refer to locations specified in the text. The inset A1 refers to case study in Sec. 5.3.2.

Study site "Sobo West"
Susceptibility for thermal erosion
Median of 100 model runs
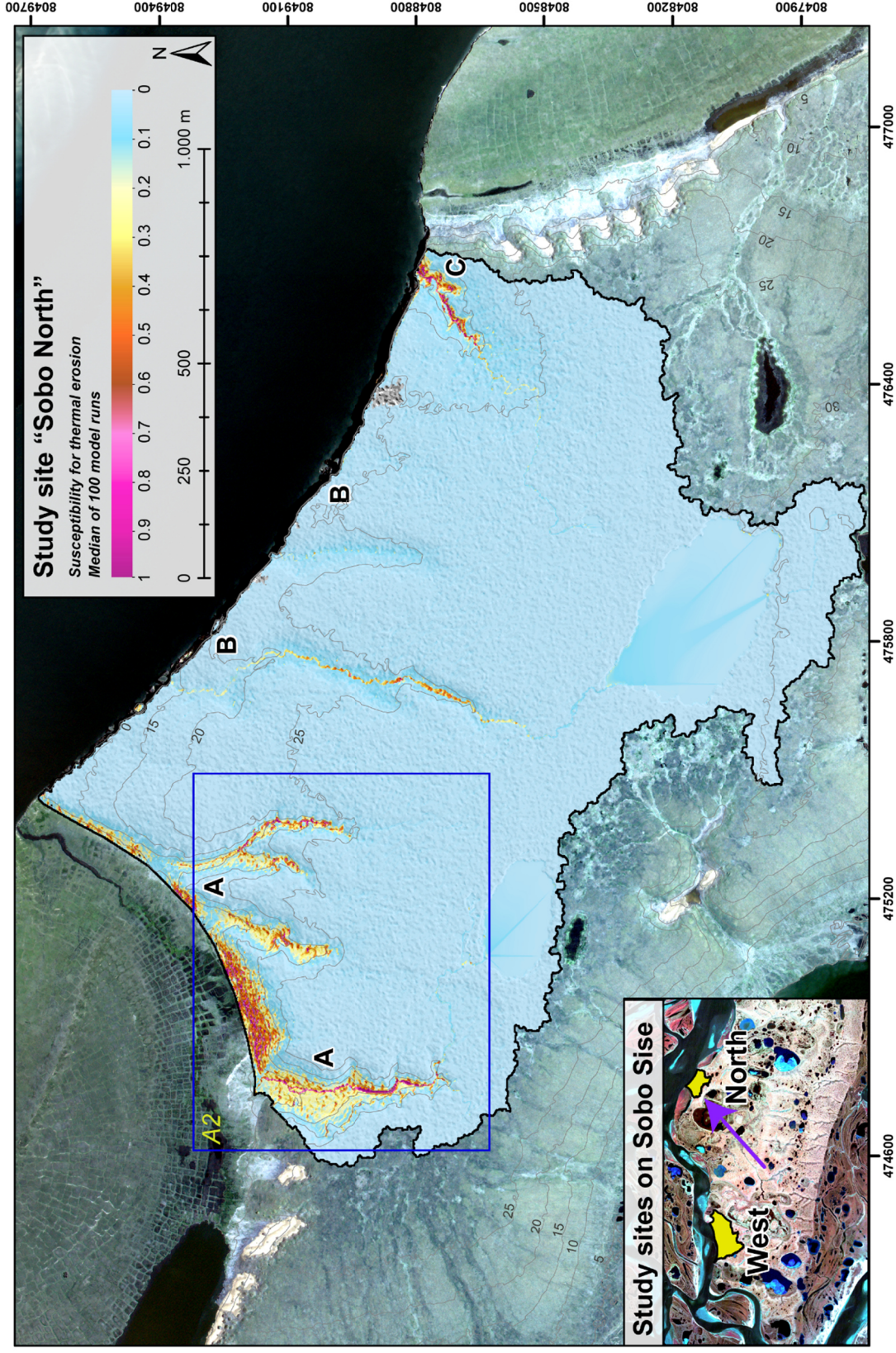
Figure 39: Consensus map for Sobo North. The capital letters A–C refer to locations specified in the text. The inset A2 refers to case study in Sec 5.3.3).
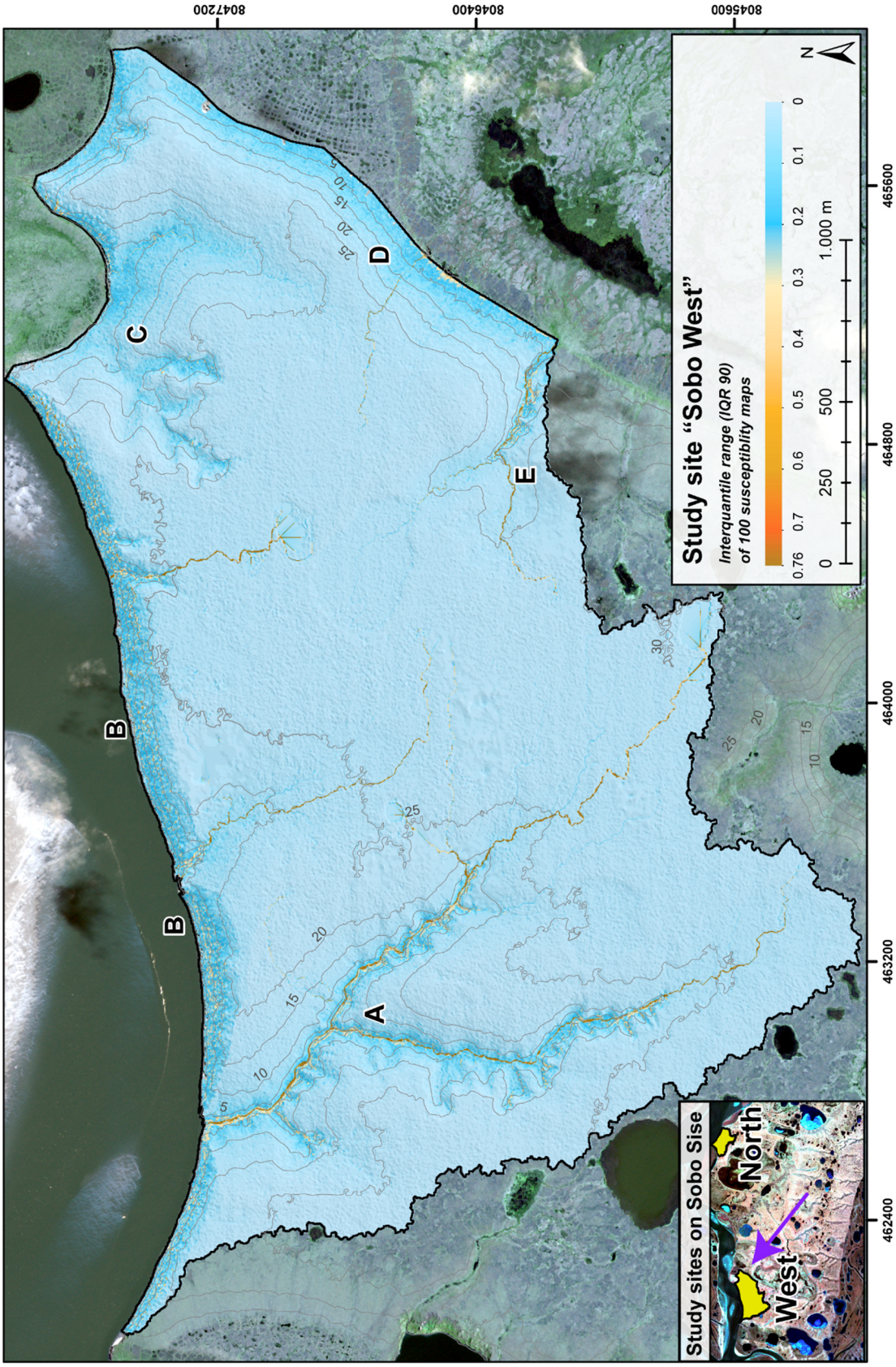
Figure 40: IQR90 map for Sobo West. The capital letters A-E refer to locations specified in the text.
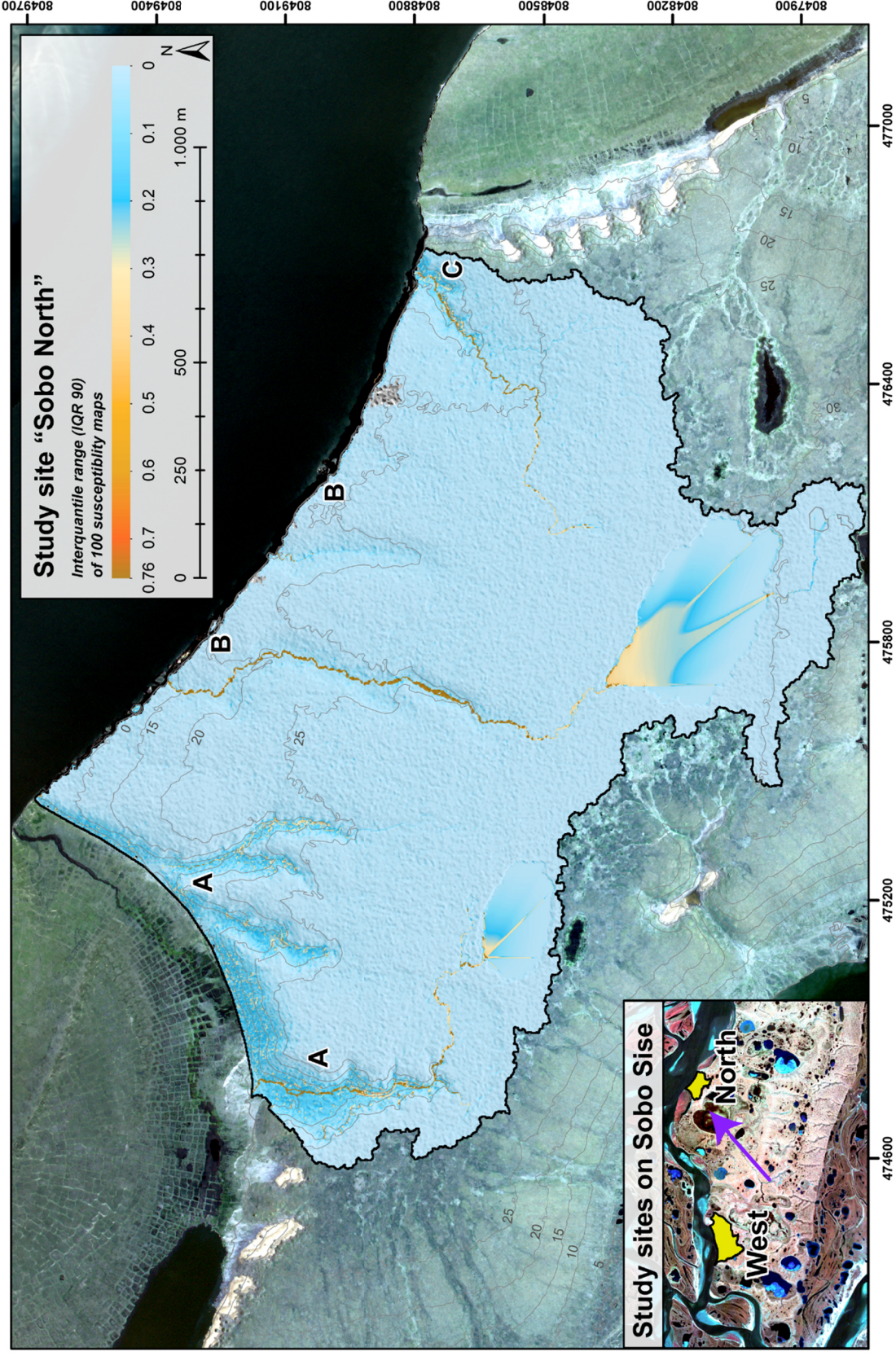
Figure 41: IQR90 map for Sobo North. The capital letters A–C refer to locations specified in the text.

5.4   Validation

The qualitative evaluation of predicted erosion probabilities in Section 5.3 revealed that the selected model parameters yield a certain degree of spatial variability in the susceptibility maps. ROCs (see Section 4.4.6) quantify and visualize the success of mapped and predicted erosion for each model run within the study site. For clarification, validation via success rates as applied in this study measures the agreement of mapped and predicted erosion within the area the model was created. Success rates differ from prediction rates, which measure the transferability of a model generated in a test site, with data from a spatially and/ or temporally independent training site.

The left panel in Figure 42 shows the success curves for the 100 models. The curves lie close to each other and show low variability. No outliers exist for the model runs, thus indicating a stable performance of all models. This is additionally supported bythe close range of the AUC values (Figure 42 right panel), ranging from a minimum of 0.939 to a maximum of 0.965. The median map as the proposed consensus of all models has an AUC value of 0.957. According to Hosmer Jr et al. (2013), success rates > 0.9 describe an outstanding discrimination between true positive and false positive rate. This demonstrates the good performance of selected environmental parameters in the model runs, which generate a precise representation of current erosion susceptibility. As an example, the false positive rate at 80 % true positive rate in the consensus map is only 5.5 %. This measurement explains the model's ability to predict correctly
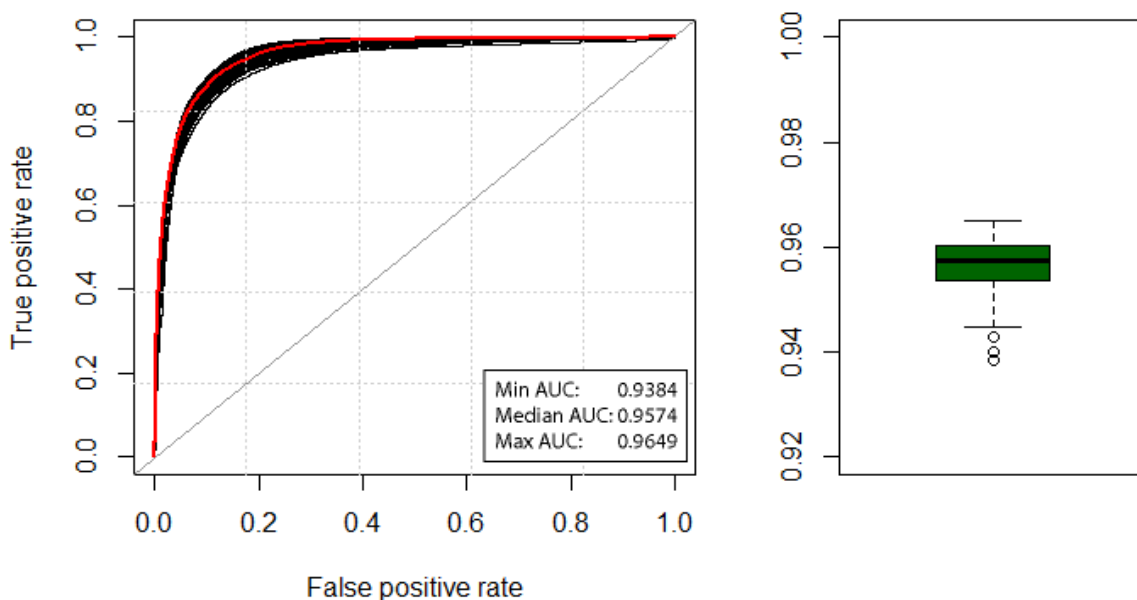
Figure 42: Evaluation of ROC curves (success rate). Left panel: ROC curves for 100 model runs. The red line indicates the Consensus model as the median of all model runs. Right: Distribution of AUC values for the ROC curves from the left panel.

61

the occurrence of mapped erosion within the study site, while it classifies most of the non-erosion cells as stable.

6   Discussion

6.1   Model parameters

The presentation of selected model parameters in Section 5.2 revealed differences in the percentage occurrence of each parameter in the susceptibility models as well as variances of the estimated parameter coefficients. The following paragraph discusses each model parameter in detail.

Slope: The positive and negative values of the coefficient values explained why the parameter *Slope* was considered least in all model runs. Erosion occurred on relief units with variable slope characteristics. Recent thermal erosion took place in steep slope sections, e.g. in the head cut zone of incised valleys and along the steep bluff, but also in the streambeds of the thermo-erosional valleys with very low slope gradients. Steep erosional slopes represent the initiation of recent erosion. In contrast, areas with lower slopes and erosion indicate terrain units where active permafrost degradation has already occurred in a former period. This conflict made it difficult to define a statistical relationship between *Slope* and the probability of erosion, because erosion was mapped in locations with both, high slopes (positive coefficients) and low slopes (negative coefficients). Furthermore, the generation of the slope map highly depends on the DEM quality, because slope represents the first derivative of elevation. A high dispersion of slope values along the valley slopes additionally displayed a critical point during parameter selection and model fitting. Consequently, this parameter was excluded in most model runs.

Contribution of surface open water: This parameter had both, the largest variability of estimated coefficients and the highest absolute coefficient values. Both effects can be explained by the value distribution in this dataset and the random sampling during model generation. The *Contribution of surface open water* dataset represents a modeled linear stream network with specifically high values at sites of open water body accumulation , and zero values at sites of no accumulation. By contrast, the occurrence of erosion represents an areal feature (Figure 43). During model generation, cells from the areas affected by erosion were randomly selected and the set of environmental parameters at these cells predicted the probability of erosion. The open water dataset did not occupy the full area of erosion cells due to its linear behavior. Thus, it was
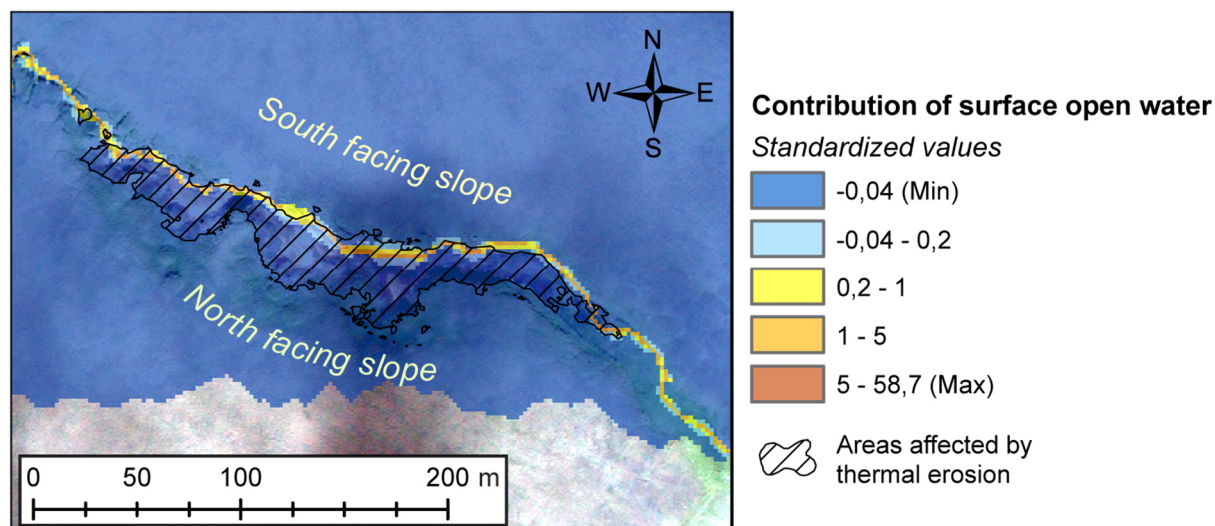
Figure 43: Example for accumulation of surface open water cells in a thermo-erosional valley (image section is the same as in Figure 28). Note that the modelled stream network occupies only parts of the mapped thermo-erosional features. Values are standardized with μ= 0 and σ= 1.

possible that an erosional cell was randomly chosen, although no flow accumulation was calculated. On the other hand, if an erosional cell represented a cell of flow accumulation, the influence of this parameter was very strong during model generation. As explained in Section 4.4.5, a one-standard-deviation change in the contribution of surface open water dataset had in this case a stronger effect than a-one-standard deviation change of any other environmental parameter. Consequently, the more erosional cells were "hit" in the random sample by the open water stream network, the higher the value of the parameter coefficient was. Conversely, if insufficient stream cells were sampled for predicting erosion, the variance of the estimated coefficient was too large and the parameter was removed from the model. This phenomenon explained the absence of this parameter in 19 % of all model runs.

Snow accumulation was considered in every model run and had the second largest absolute coefficient values. This demonstrated the strong effect of snow accumulation on thermal erosion. The sign of the parameter coefficient was negative due to the data input. Negative values indicated leeward locations and thus, high snow accumulation.

Relief ratio represented another stable parameter during all model runs. The positive values indicated that locations with a higher relief ratio were more susceptible for erosion. This was especially true along the Ice Complex bluff. Erosion occurred along the head scarps of the bluff and on the baydjarakhs. Both features had a high gravitational energy as well as a short distance to the base level.

Profile curvature showed slightly positive values. Accordingly, contemporary erosion was more likely to be observed on convex than on concave surfaces. This statement is true for erosion on baydjarakhs and for the transition from Yedoma uplands towards thermo-erosional valleys. However, erosion also occurred in streambeds with a convex profile curvature. Similar to the *Slope* parameter, the uncertainty resulted in a partial exclusion of *Profile curvature* in 23 % of the model runs.

Direct solar insolation was the third stable parameter, which was chosen in every model run. The negative values supported the assumption that erosion was more likely to occur on surfaces with reduced solar insolation. This was in agreement with field observations, which indicated that north facing slopes with lower solar radiation in summer promoted erosion (Figure 28).

In summary, the datasets and applied methods suggested *Snow accumulation*, *Relief ratio*, *Direct solar insolation* and the *Contribution of surface open water* to be the decisive parameters for thermal erosion on ice-rich permafrost. The first three parameters were the most stable variables due to their persistent occurrence in the model runs, whereas *Contribution of surface open water* had the strongest influence in the logistic regression model.

6.2 Qualitative and quantitative model evaluation

6.2.1 General model performance within the study sites

The general spatial pattern of modelled erosion susceptibility in Sobo West agreed with the mapped pattern of contemporary erosion. Recently active parts in the landscape, i.e. deeply incised valleys and the Ice Complex bluff, were very well distinguished from the stable Yedoma surface. This delineation between high and low susceptible areas was also successful for large parts of Sobo North. The different environmental settings helped to explain the discrepancy between the mean predicted erosion probabilities in the two study sites. Sobo West with its steep, elongated bluff and deeply incised valleys and gullies offered more locations for thermal erosion (based on ideal conditions according to the environmental parameters). In Sobo North, the relief had a rather uniform character and especially the highly susceptible bluff was missing. The on average low probability for thermal erosion in both study sites (5.2 %) agreed with the concept of a rare event dataset and demonstrated the localized occurrence of this phenomenon.

_____

The predicted erosional susceptibility along the bluff is as low as on the adjacent Yedoma uplands despite the high annual thermo-erosional activity along the bluff. However, the missing delineation of increased susceptibility is conceptually correct. In this study, thermal erosion was considered as a process controlled by local topography, at sites where the interaction of warm running water and topo-climatic parameters caused recent surface instabilities. Erosional features were defined as exposed soil and the geomorphological parameters at these cells were selected to model the corresponding probabilities of erosion. Given these considerations, it is not possible to delineate susceptibility for this specific type of fluvio-thermal erosion along the Lena River. The annual rates of fluvio-thermal erosion along the bluff by the Lena River exceeded the slope driven thermal erosion in thermo-erosional valleys by a multiple. The different types of relief degradation by thermal erosion express this difference. The fluvio-thermal erosion by the Lena River represents an external process affecting the bluff, whereas thermal-erosion in thermo-erosional valleys (as the prevailing type of erosion in the study sites) can be interpreted as an in-situ process. The transport capacity of the small streams originating at the Yedoma uplands and dropping as water falls into the Lena River, is insufficient to cause significant mechanical stress for erosion in the streambeds. Thus, the effective period for adjusting the local stream gradient by thermal erosion in the streambeds is too short compared to the high annual rate of bank erosion of the Lena River.

Regarding the methodological challenges, no erosion in terms of exposed soil was mapped due to the direct change within one cell from vegetated Yedoma uplands to the Lena River. The mapping of erosional features in contemporary GeoEye-1 satellite images exhibited the use of historical satellite imagery, although the multi-temporal mapping (Figure 44) revealed the high erosional activity. Additionally, none of the selected environmental parameters was capable or scientifically justifiable of demonstrating the change of elevation from ca. 20 m a.r.l. towards river level within one cell. As an example, applying the parameter *Direct solar radiation* at this location would have been unjustified, because the failure of the bluff was related to undercutting of the Lena River, not to the reduced solar radiation on the north-facing bluff. In summary, a delineation of higher susceptibility along the bluff may be desirable, but was not feasible with the approach chosen in this study.

In the following, one case study for each study site (see insets A1 in Figure 38 and A2 in Figure 39) is used to display and discuss the strengths and drawbacks of the susceptibility models.

_____

6.2.2 Case study A1 (Sobo West)

The top panel of Figure 45 illustrates a common situation along the steep Ice Complex bluff in the western study site with a very heterogeneous pattern of erosion. Erosion was mapped in i) a deeply incised valley and in short gullies; ii) on the north facing flanks of baydjarakhs; iii) on the toe of the Ice Complex bluff and; iv) on the head scarp of the bluff.

The highest probabilities of erosion (dark red to purple colors) were predicted for the valley and the gullies, as well as for the toe of the bluff. Especially for the deeply incised valley in the eastern section mapped and predicted erosion showed a very good agreement. Here, the full set of selected environmental parameters pertained to predict erosion. The valley faces towards the north and receives less solar radiation, thus promoting higher surface moisture. Snow can remain for an extended period due to the deep incision. A lake in the upper catchment feeds the valley, promoting heat transfer into the ground and thaw of underlying permafrost. The relief ratio is high, because the valley drops ca. 25 m in less than 200 m of horizontal distance. Hence, predicted probabilities ranged between 0.8 and 1 for the valley floor. However, high uncertainties existed for the modelled stream before it enters the valley. This is, as pointed out in Section 6.1, an effect of the partial absence of the parameter *Contribution of surface open water* in the 100 model runs. Yet, assuming the strong influence of this parameter on thermal erosion, this pathway provides useful information about the direction of further valley development.

Predicted probabilities along the toe of the bluff showed a similarly good agreement with mapped erosion. The decisive parameters at this location were again high *Relief ratio*, low *Incoming solar radiation* as well as high *Profile curvatures*. Lateral erosion along the bluff by the Lena River results in a permanent reshaping of the bluff through the removal of sediment. The toe forms a convex profile, thus promoting erosion. Consequently, the good spatial delineation of erosion was also indicated by the low variability in the models, as demonstrated in the IQR90 map.

The visual validation of erosion on baydjarakhs that occurred between the toe and the head scarp revealed problems. Despite its high planimetric resolution, the DEM was not capable to entirely reflect the small-scale relief differences along the slope. If the quality was sufficient, baydjarakhs were classified with very high (> 0.7) susceptibility values. Failing that, the susceptibility values rarely exceed low to moderate (< 0.4) values. Another explanation of this error is, probably to a lesser degree, a mismatch between the DEM and the satellite image. A

Figure 45: Case study A1 (Sobo West). Comparison of mapped erosion pixels (top panel), predicted erosion from the consensus map (middle panel) and the spatial variability in terms of the IQR90 map (bottom panel). Scale in the upper panel fits for all panels.

small offset between the exact location of mapped erosion on the baydjarakhs and the corresponding values of the environmental parameters results in spurious prediction values for these cells.

The head scarps along the bluff reflect the largest discrepancies between detected and predicted erosion. Many cells receive no or low susceptibility for erosion. Again, the DEM might not provide a sufficient level of detail to display the abrupt step in the relief. Head scarps differ from other observed erosional features, because the longitudinal profile of the scarp is highly concave. In the regression model, the parameter *Profile curvature* is associated with a positive coefficient (Section 5.2), which rather predicts erosion on convex relief units. The consensus map reports consistently low susceptibility values in all models for the head scarp zone. This demonstrates the low predictive power of the model for this surface feature.

6.2.3 Case study A2 (Sobo North)

Case study A2 is located in the western section of Sobo North. The environmental setting of the detailed study A2 differs from the example above, because a flood plain with elevations between 2-4 m a.r.l. represents the base level of erosion. Three thermo-erosional valleys cut deep into the Yedoma uplands. Erosion was mapped in their head reaches and partly in their middle reaches (Figure 46 top panel).

Moderate to high susceptibility values were confined exclusively to those valleys and the slope in-between facing towards the floodplain (middle panel in Figure 46). This slope is a representative example for relief entities in both study sites with high predicted probabilities, but no observed erosion. The environmental setting along the slope seems suitable for thermal erosion. Its convex longitudinal profile is combined with a high relief ratio, a northern exposure and favorable conditions for snow accumulation. The absence of erosion could be explained with the low influence of the Lena River on the slope's base level. The toe of the slope is rarely affected during extreme flood events, unlike example A1, in which the bluff is in regular interaction with the river by the annual Lena flood. However, the contemporary high slope stability does not exclude future erosion on this slope. Undercutting by several extreme flood events can destabilize the slope within a few years. Moreover, erosion might occur as slow denudation process on micro-topographic scale along the slope, which does not necessarily implies recent exposure of bare soil. For clarification, the scope of a susceptibility map is the detection of potentially vulnerable sites by a set of environmental parameters. Thus, the
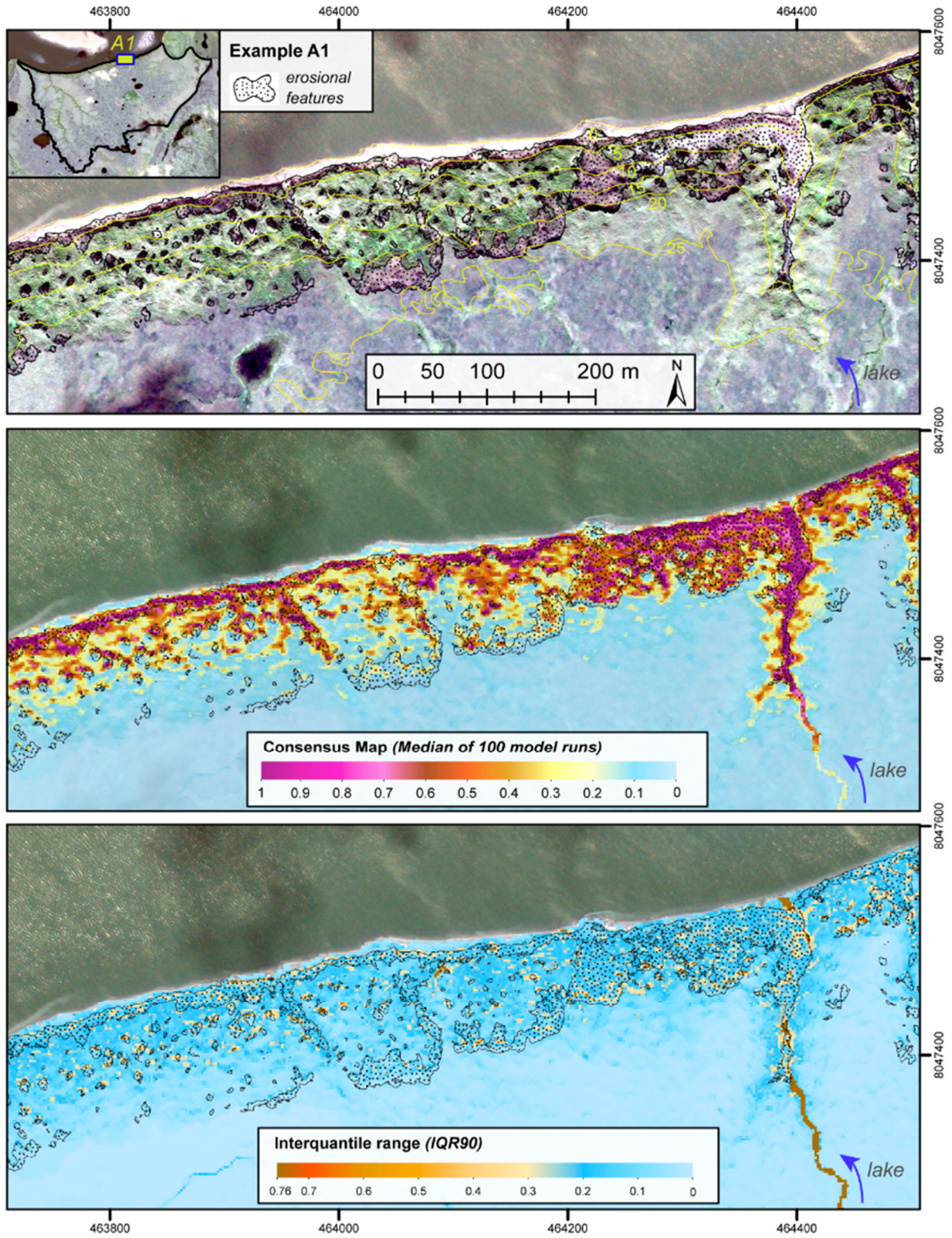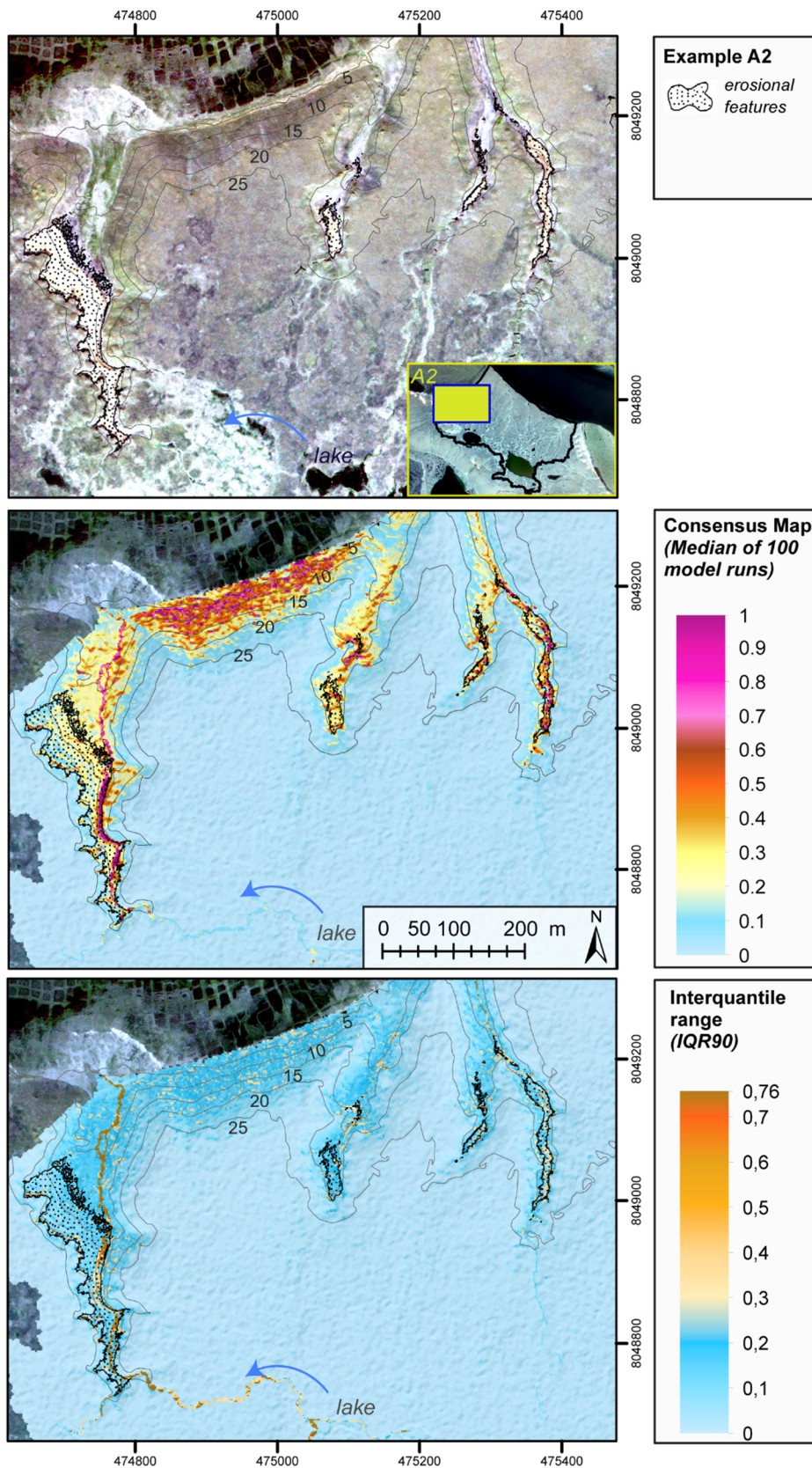
Figure 46: Detailed study A2 (Sobo North). Comparison of mapped erosion pixels (top panel), predicted erosion from the consensus map (middle panel) and the spatial variability in terms of the IQR90 map (bottom panel). Scale in the middle panel fits for all panels.

parameters can predict a high local probability for erosion, even if contemporarily no erosion is observed. Generally, discrepancies between mapping and prediction should always be regarded critically. However, as they can arise as a logical consequence of the local interaction of the model parameters, they must not necessarily represent an error in the model.

The western most valley revealed some difficulties in the interpretation of the consensus map. Like in example A1 (Section 6.2.2), a stream modelled from the surface open water dataset flows towards the valley, which is indicated by increasing susceptibility values towards the head cut of the valley. Predicted erosional susceptibility showed high values in the upper reaches, which was in agreement with mapped erosion (middle panel in Figure 46). The decreased probabilities towards the outlet matched with observed dense vegetation in the streambed during the expedition. Here, the vegetation stabilizes the valley floor. However, the closer examination displayed a linear feature of very high probabilities (> 0.7) along the valley floor after the zone of active erosion. The reason for this phenomenon was the influence of the surface open water parameter. This error occurred, because the linear accumulation of contributing water pixels is not capable of modeling the non-linear relationship between this parameter and thermal erosion. Conceptually, susceptibility values should slowly increase, starting with a minimum at the lake and reach a maximum, when the modelled stream enters the upper zone of an existing valley. At this location, the heat transfer of warm water into the ground is highest. When the water flows towards the middle and lower reaches, the susceptibility should decrease, since the valley already possesses a high degree of stabilization in these reaches. Here, underlying permafrost had already thawed, which makes these sections less susceptible for thermal erosion. However, the logistic regression model can only treat monotonic increasing relationships between the explanatory variable and the response variable. Consequently, an increase in the values of *Contribution of surface open water* is coherent with an increase of susceptibility. The strong model's dependence on this environmental parameter explained the inappropriate high susceptibility values in the lower reaches of this valley. Conversely, the absence of this parameter in 19% of all model runs caused a very high variability in predicting the susceptibility for the same cells (see bottom panel in Figure 46). The co-existence of both, high susceptibility and high uncertainty, raised the question about the validity of this parameter in this specific case. On the contrary, the effect of increasing susceptibility from the head cut until the outlet is desired for recently developing valleys and gullies along the bluff (see example A1). As long as bare ice is directly exposed to warm water,

_____

it is assumed that an increasing amount of warm water also results in an increased rate of erosion. Generally, the consideration of contributing warm water as a driving agent for thermal erosion has produced good model results at different locations within the study sites. However, improvements regarding the generation of this parameter could also enhance the performance and validity of the model.

A further problem addresses mapping errors. The semi-automated mapping approach classified a vast part in the middle reach of the same valley as an erosional feature (Figure 46 top panel). The identification as bare soil was correct, but the comparison of classified soil pixels with in situ photographs revealed that these pixels reflect deposited sediments from the head cut zone. Thus, they rather represent the consequence of erosion, not its cause. As an indication of the good predictive quality of the model, only low to very moderate susceptibility values were predicted for the spurious mapped erosional cells. In this particular case, the model represents the reality better than the mapped erosional features.

6.2.4   Summary of qualitative model fit

The visual interpretation of model success allowed the delineation of five categories in terms of agreement between mapped and predicted erosion.

Good model results were found in the headcut zone and on the steep slopes in the upper reaches of incised thermo-erosional valleys and in thermo-erosional gullies. Here, the interaction of warm contributing surface water, low incoming solar radiation, high relief energy and long residence time of accumulated snow reflected the pattern of contemporary erosion in a realistic way.

Medium model fit was observed for the baydjarakhs along the steep thermo-erosional bluff in Sobo West. The partial discrepancies between mapped and predicted erosion could be attributed to inaccuracies in the raw data, mainly DEM quality and/ or a mismatch between the DEM and the orthorectified satellite imagery.

Poor model results were identified along the head scarps of the thermo-erosional bluff. The low success of recognition depended on the specific morphological characteristic of those erosional features and the DEM quality.

High susceptibility without mapped erosion was predicted for steep slope sections along the alasses. The combination of environmental parameters supported the assumption that slopes along alasses are highly susceptible for future erosion. However, erosion might occur on a

longer time scale with a lower degree of surface disturbance exhibiting the observation of recent erosion.

High susceptibility values and high uncertainties appeared upstream and downstream of warm surface open water passing the active zones of thermo-erosional valleys. The paths of increased susceptibility upstream of thermo-erosional valleys indicate potential trajectories for future valley development. Assuming the strong influence of surface open water on thermal erosion, the uncertainties related to this parameter show additional requirements in terms of model and parameter improvement.

6.2.5 Validation

The validation of the consensus model showed an outstanding AUC value of 0.957, which demonstrated a very good discrimination between stable and active landscape units within the study sites. Similar good success rates for logistic regression models were reported in other studies (e.g. Pradhan and Lee, 2010; Regmi et al., 2014; Van Den Eeckhaut et al., 2010). However, the high success rates in this study also resulted from the specific characteristics of the Yedoma landscape. The success rates reflected better the small proportion of high susceptibility values along the bluff and in deeply incised valleys compared to the high proportion of low susceptibility values on the uniform Yedoma uplands. Thus, success rates in the presented landscape type will always perform better than in e.g. steep alpine terrain with generally higher predicted probabilities for the occurrence of the phenomenon (Frattini et al., 2010). Moreover, because the success rate compares the goodness of fit of predicted erosion probabilities with observed erosion in the same study site, it will always show better results than a prediction rate, which validates the model performance with data from a spatially independent study area (Chung and Fabbri, 2003). Consequently, if a model possesses already a poor success rate, it will perform even worse in another study site. Thus, most reported prediction rates rarely exceed AUC values of > 0.9, e.g. 0.83 in Heckmann et al. (2014), 0.836 in Ayalew and Yamagishi (2005), 0.842 in Yilmaz (2009) or 0.86 in Meusburger and Alewell (2009). However, the presented success rates seem promising as they indicate that the selected model parameters will also perform well in predicting thermal erosion in other study sites with comparable environmental conditions. Additionally, the IQR90 map provided valuable information about the spatial model variability, which the AUC values are unable to represent.

The combination of both methods are found to be a good approach for qualitative and quantitative validation of model uncertainty and to give further suggestions for model improvements.

6.3 Outlook

The results of this study provide a basis for further work in terms of data collection, processing and modeling of thermal erosion in order to refine model results.

Godin et al. (2014) observed increasing hydrological connectivity to promote the development of thermo-erosional gullies. In this study, the susceptibility model supported the assumption that coupled streams of warm surface water promote thermal erosion. However, this theory could not be verified empirically due to insufficient measurements in the field. Discharge and water temperature measurements over an entire summer season in both stable and unstable valleys would help to support this hypothesis. In the same way, Fortier et al. (2007) demonstrated the strong impact of snowmelt on thermal erosion. Since this effect was also accounted high priority in this model, fieldwork is supposed to generally improve the process understanding.

The review on predicted susceptibility on the steep slopes along the Ice Complex bluff in Sobo West stressed the influence of DEM noise on prediction accuracy. The environmental parameters are derivatives of the DEM and react very sensitively on DEM resolution. Noise in the explanatory parameters can distort the parameter selection process in stepwise logistic regression and the estimation of model coefficients. For future studies, affected areas in the DEM could be replaced by using other methods for DEM generation. As a first trial, a survey along an Ice Complex bluff with terrestrial structure from motion technique showed a very high level of detail and a lesser degree of spatial uncertainty in the DEM. However, a larger areal coverage using this method might be difficult to realize. If noise continues to exist in the DEM, a cell resolution of 2 m can limit the stability of the regression model. Some authors resampled the fine raster resolution to a coarser resolution, resulting in a smoothed, generalized terrain (e.g. Heckmann et al., 2014). This method has proved to affect the spatial delineation of susceptibility classes in the susceptibility map (e.g. Lucà et al., 2011). The advantages and drawbacks of different mapping units for hazard assessment besides the raster cell (terrain units, unique condition units, slope units, topographic units) have been proposed and discussed in

_____

literature (e.g. Guzzetti et al., 1999). Some studies preferred slope units over cell units due to better accuracies between observed and predicted cases (Baeza et al., 2010; Frattini et al., 2010). However, it has to be kept in mind that as the cell size is increased or transformed to a major mapping unit, information is lost and spatially the precision of modeling erosion susceptibility is partly affected. This can be a problem concerning the small-scale occurrence of thermal erosion in this study. Thus, comparing and finding the ideal mapping unit for predicting thermal erosion might be a challenging task.

With respect to the environmental parameters, improvements are especially necessary for the *Contribution of open water surfaces* parameter. Currently, this parameter represents a quasi-two-dimensional stream network with linear behavior. However, warm water causes a lateral heat transfer into the ground and, hence, the formation of a thermo-erosional niche (Fortier et al., 2007; Godin et al., 2014). The undercutting causes instabilities and failure not only at the water ice interface, but also at elevated and distal parts of the slope. The modelled stream network denies this effect since it assigns high susceptibility values only to cells within the streambed. An improved version of this dataset would also consider the flow depth and the cross section of the valley along the flow path. Ideally, flow accumulation values should increase and spread laterally in deeply incised, V-shaped valleys. When the stream flows towards broad, stabilized cross sections, accumulation values should decrease again, because these valley sections are less prone to thermal erosion.

As noted in Section 6.2.3, logistic regression models can only handle linear relationships between explanatory variables and predicted probabilities. This was challenging for the linearly increasing *Contribution of open water surfaces* parameter used in this study. However, it was supposed that erosion probability decreased when the parameter had passed a specific tipping point in the relief. Hjort and Luoto (2011) suggested using Generalized Additive Models (GAMs; a non-parametric extension of GLMs) for explanatory variables with a nonlinear relationship of the dependent variable. The authors highlighted the greater flexibility and capacity of GAMs, which provide more accurate models than GLMs. A broad variety of statistical approaches exists in predictive geomorphological and susceptibility mapping besides GLMs and GAMs. Several studies compared their applicability, degree of model fit and prediction skills (e.g. see in Hjort and Marmion, 2008; Luoto and Hjort, 2005; Rossi et al., 2010; Yilmaz, 2009). However, the choice of the best modeling method differed in each study, because the performance of the respective method depends on numerous factors as data

resolution, environmental setting, selection of explanatory variables, etc. The application of these models on the present data would reveal the relative validity of logistic regression for susceptibility mapping.

Finally, external validation in the sense of prediction rates (Section 5.4) has highest priority for further analysis. The Ice Complex islands in the Lena River Delta represent potential test sites due to their climatic and geological similarities. However, Morgenstern (2012) observed a great variability in the valley and stream morphology as well as in the stream density on these islands. Furthermore, Pollozek (2015) reported different changes in the spatial extent of thermo-erosional valleys on the four major Ice Complex islands within the delta. Rates of changes varied between 0.4 and 2.1 % in the period from 1975 to 2011, whereas the greatest increase was observed on Sobo-Sise. Thus, validation with data from these study sites will reveal whether the decisive parameters found on Sobo-Sise will also perform well in susceptibility modeling on other Ice Complex landscapes. Other studies used temporal validation, i.e. testing the predictive power of a model generated from events from a former time period with cases observed at a later period (e.g. Guzzetti et al., 2006; Meusburger and Alewell, 2009; Rossi et al., 2010). As Pollozek (2015) stated, the temporal changes of thermo-erosional valleys indicate a slow process of landscape degradation. Prior satellite images do not command the same information (resolution, bands, etc.), which were used to map erosion in the present dataset As thermal erosion represents a small-scale and locally confined phenomenon, the detection of actually active parts in the landscape between different time steps is assumed to be problematic. Despite its difficult application, temporal validation would be another option for assessing the model transferability.

6.4 Future degradation of ice-rich permafrost by thermal erosion

The benefit of susceptibility maps consists not only in the delineation of present-day vulnerability, but also in the identification of future sites of landscape degradation. The results of this study showed that the Ice Complex is affected by in-situ thermo-erosional processes. The efficiency was largest at the rapid transition from the uplands to thermo-erosional valleys and gullies. These locations (e.g. valleys Figure 38 A & E; Figure 39 A & C) are expected to be the most susceptible locations for future thermal erosion. External influences like the lateral fluvio-thermal erosion of the Lena River (Figure 44) additionally destabilize the base level of

the Ice Complex and promote a long-term disequilibrium in the hydrological system of the island.

The question of how thermal erosion will contribute to the degradation of ice-rich permafrost in the future is difficult to answer in times of global climate change. At most sites in the Arctic region, permafrost temperatures have increased during the past three decades, which was referred to increased air temperatures and changing snow cover in this period (Vaughan et al., 2013). Rising air temperatures are associated with the thawing of frozen ground layers and active layer deepening (Anisimov and Reneva, 2006; Jorgenson et al., 2006). However, the interaction between climate and permafrost is complex, since permafrost degradation has also been observed in regions with an observed decrease in air temperature (Osterkamp et al., 2009). However, air temperatures in the Polar Arctic are expected to rise by 2.2 to 8.3°C until the end of the 21st century (Collins et al., 2013), leading to a substantial loss in permafrost extent (Koven et al., 2013).

Ice-rich permafrost is regarded as the permafrost type most vulnerable type to climate warming (Shur and Jorgenson, 2007). The increasing thaw of underlying frozen sediments is expected to provide an additional amount of surface water. Drainage networks are likely to increase in extent, because the greater surface water runoff exceeds the storage capacity of the active layer (Rowland et al., 2010). Consequently, thermal erosion will occur as a widespread phenomenon in Arctic environments and will be additionally promoted by the growth and subsequent drainage of thermokarst lakes (Grosse et al., 2011). However, the uncertainty about the future development of thermo-erosional landforms is subject of a recent scientific debate. Parameters like topography, surface water, groundwater, snow, soil, vegetation and active layer thickness will form a complex system with positive and negative feedbacks on the resilience and vulnerability of permafrost to climate change (Jorgenson et al., 2010). New thermo-erosional landforms were reported to develop abruptly within a very short period of time (Godin et al., 2014; Jorgenson et al., 2006; Toniolo et al., 2009). However, their impact on the geomorphology is assumed to form long-term relief units in the permafrost landscape (Gooseff et al., 2009). Climate-change induced thermal erosion is supposed to alter terrestrial and aquatic Arctic ecosystems. Disturbances in the vegetation cover will likely result in a growing influence on plant species composition, e.g. by a rapid colonization of exposed soils and a movement of species beyond their present geographic distribution (Lantz et al., 2009). Arctic streams already showed an increased rate of sediment and nutrient delivery from thermo-erosional features,

_____

which will furthermore affect aquatic habitats (Bowden et al., 2008). On a global scale, thermal erosion represents a decisive process in the release of previously bound organic carbon and methane from the frozen ground (Schuur et al., 2008). Despite uncertainties in climate scenarios, Schuur et al. (2015) estimated the carbon emissions from permafrost to increase global temperatures by 0.13–0.27°C until the end of the 21st century.

Concluding, the recent and future development of thermo-erosional landforms is attributed major importance for future research. A precise determination and quantification of this process is necessary, since thermal erosion causes physical, chemical and biological changes across both spatial and temporal scales. As thermal erosion is a so far little investigated phenomenon, the decisive factors found in this study represent a first step to understand the complex interactions between process components and landscape evolution. The generation of susceptibility maps proved to be a useful tool to identify the locations with the highest vulnerability to thermal erosion. A further improvement of the datasets and modeling techniques is recommended to extent the findings of this study on a larger spatial scale and thus, to assess the overall impact of thermal erosion on the degradation of ice-rich permafrost.

7   Conclusions

Although thermal erosion accounts for a large geomorphological impact on ice-rich permafrost landscapes, few studies have addressed the spatial occurrence and the decisive factors for this type of erosion. The present study investigated these research gaps using satellite imagery and DEM analysis for deriving potentially influencing parameters and logistic regression for susceptibility modeling on an island in the Arctic Lena River Delta. The environmental parameters were reduced stepwise in the model and the modeling procedure was repeated 100 times to avoid the risk of accidentally choosing a model with a poor fit.

Contribution of surface open water, *Snow accumulation*, *Relief ratio* and *Potential incoming solar radiation* were found to be the most significant and stable parameters in the logistic regression models. These components were in agreement with the conceptual understanding of thermo-erosional processes and were reported in previous studies. The consensus map (the median of all model runs) exhibited a satisfying match with the mapped erosional features. Accordingly, the upper sections of deeply incised valleys and gullies and steep bluffs at the margins of the study site received the highest susceptibilities. These relief units are expected to show high vulnerability also in other sites of ice-rich permafrost, although an external validation was not possible in this study.

The generation of a consensus map proved to be an adequate solution to account for the variability in the model building process. Furthermore, the analysis of the IQR90 map, which encompasses 90 % of the modelled susceptibility values, was an efficient way for quantifying the spatial uncertainty of the consensus map. In doing so, the predictive power of the final susceptibility map turned out to be depending on a) DEM resolution, b) the method of producing the explanatory parameters and c) the modeling approach.

Future work will include the prediction for erosion susceptibility using other terrain mapping units besides the raster cell resolution to avoid the inclusion of DEM noise in the model. Furthermore, the environmental parameters, especially the *Contribution of the open water* parameter will receive improvements to better reflect their impact on contemporary erosion. Different modeling approaches besides the logistic regression model will be applied to include non-linear relationships between environmental parameters and erosion in the susceptibility model. Finally, model validation in a spatially independent dataset has highest priority. This step would verify the transferability of evidences concerning the decisive parameters and predicted erosional susceptibility found in the study site.

_____

8 References

Agafonov, L., Strunk, H., Nuber, T., 2004. Thermokarst dynamics in Western Siberia: insights from dendrochronological research. Palaeogeography, Palaeoclimatology, Palaeoecology 209, 183–196. doi:10.1016/j.palaeo.2004.02.024

Agresti, A., 2007. An introduction to categorical data analysis. 3rd ed. Wiley-Interscience, Hoboken, NJ.

Akaike, H., 1974. A new look at the statistical model identification. Automatic Control, IEEE Transactions on 19, 716–723.

Akgün, A., Türk, N., 2011. Mapping erosion susceptibility by a multivariate statistical method: A case study from the Ayvalık region, NW Turkey. Computers & Geosciences 37, 1515–1524. doi:10.1016/j.cageo.2010.09.006

Anisimov, O., Reneva, S., 2006. Permafrost and changing climate: the Russian perspective. AMBIO: A Journal of the Human Environment 35, 169–175.

Arctic Monitoring and Assessment Programme (AMAP), 2011. Changes in Arctic Snow, Water, Ice and Permafrost. SWIPA 2011 Overview Report, Oslo.

Are, F., Reimnitz, E., 2000. An overview of the Lena River Delta setting: geology, tectonics, geomorphology, and hydrology. Journal of Coastal Research, 1083–1093.

Ayalew, L., Yamagishi, H., 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology 65, 15–31. doi:10.1016/j.geomorph.2004.06.010

Baeza, C., Lantada, N., Moya, J., 2010. Influence of sample and terrain unit on landslide susceptibility assessment at La Pobla de Lillet, Eastern Pyrenees, Spain. Environmental Earth Sciences 60, 155–167. doi:10.1007/s12665-009-0176-4

Barcaroli, G., 2014. Optimization of sampling strata with the SamplingStrata package.

Bauch, H.A., Mueller-Lupp, T., Taldenkova, E., Spielhagen, R.F., Kassens, H., Grootes, P.M., Thiede, J., Heinemeier, J., Petryashov, V.V., 2001. Chronology of the Holocene transgression at the North Siberian margin. Global and Planetary Change 31, 125–139.

Beguería, S., 2006. Validation and Evaluation of Predictive Models in Hazard Assessment and Risk Management. Natural Hazards 37, 315–329. doi:10.1007/s11069-005-5182-6

Black, R.F., 1976. Periglacial features indicative of permafrost: ice and soil wedges. Quaternary Research 6, 3–26.

Böhner, J., Antonić, O., 2009. Land-surface parameters specific to topo-climatology. Developments in Soil Science 33, 195–226.

Boike, J., Kattenstroth, B., Abramova, K., Bornemann, N., Chetverova, A., Fedorova, I., Fröb, K., Grigoriev, M., Grüber, M., Kutzbach, L., Langer, M., Minke, M., Muster, S., Piel, K., Pfeiffer, E.-M., Stoof, G., Westermann, S., Wischnewski, K., Wille, C., Hubberten, H.-W., 2013. Baseline characteristics of climate, permafrost and land cover from a new permafrost observatory in the Lena River Delta, Siberia (1998-2011). Biogeosciences 10, 2105–2128. doi:10.5194/bg-10-2105-2013

Boike, J., Wille, C., Abnizova, A., 2008. Climatology and summer energy and water balance of polygonal tundra in the Lena River Delta, Siberia. Journal of Geophysical Research 113. doi:10.1029/2007JG000540

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, 251–271. doi:10.1657/1938-4246-46.1.251

Bou Kheir, R., Abdallah, C., Khawlie, M., 2008. Assessing soil erosion in Mediterranean karst landscapes of Lebanon using remote sensing and GIS. Engineering Geology 99, 239–254. doi:10.1016/j.enggeo.2007.11.012

Bowden, W.B., Gooseff, M.N., Balser, A., Green, A., Peterson, B.J., Bradford, J., 2008. Sediment and nutrient delivery from thermokarst features in the foothills of the North Slope, Alaska: Potential impacts on headwater stream ecosystems. Journal of Geophysical Research 113. doi:10.1029/2007JG000470

Brenning, A., 2009. Benchmarking classifiers to optimally integrate terrain analysis and multispectral remote sensing in automatic rock glacier detection. Remote Sensing of Environment 113, 239–247. doi:10.1016/j.rse.2008.09.005

Brenning, A., 2005. Spatial prediction models for landslide hazards: review, comparison and evaluation. Natural Hazards and Earth System Science 5, 853–862.

Brown, J., Ferrians, O.J., Heginbottom, J., Melnikov, E., 1997. Circum-Arctic map of permafrost and ground-ice conditions. Boulder, CO: National Snow and Ice Data Center.

Chung, C.-J.F., Fabbri, A.G., 2003. Validation of Spatial Prediction Models for Landslide Hazard Mapping. Natural Hazards 30, 451–472. doi:10.1023/B:NHAZ.0000007172.62651.2b

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W., Johns, T., Krinner, G., 2013. Long-term climate change: projections, commitments and irreversibility, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1029–1136.

Conforti, M., Aucelli, P.P.C., Robustelli, G., Scarciglia, F., 2011. Geomorphology and GIS analysis for mapping gully erosion susceptibility in the Turbolo stream catchment (Northern Calabria, Italy). Natural Hazards 56, 881–898. doi:10.1007/s11069-010-9598-2

Conoscenti, C., Angileri, S., Cappadonia, C., Rotigliano, E., Agnesi, V., Märker, M., 2014. Gully erosion susceptibility assessment by means of GIS-based logistic regression: A case of Sicily (Italy). Geomorphology 204, 399–411.

Costard, F., Dupeyrat, L., Gautier, E., Carey-Gailhardis, E., 2003. Fluvial thermal erosion investigations along a rapidly eroding river bank: Application to the Lena River (Central Siberia). Earth Surface Processes and Landforms 28, 1349–1359. doi:10.1002/esp.592

Costard, F., Gautier, E., Brunstein, D., Hammadi, J., Fedorov, A., Yang, D., Dupeyrat, L., 2007. Impact of the global warming on the fluvial thermal erosion over the Lena River in Central Siberia. Geophysical Research Letters 34. doi:10.1029/2007GL030212

Czudek, T., Demek, J., 1970. Thermokarst in Siberia and its influence on the development of lowland relief. Quaternary Research 1, 103–120.

Dai, F.C., Lee, C.F., 2002. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. Geomorphology 42, 213–228.

Dallimore, S.R., Wolfe, S.A., Solomon, S.M., 1996. Influence of ground ice and permafrost on coastal evolution, Richards Island, Beaufort Sea coast, NWT. Canadian Journal of Earth Sciences 33, 664–675.

Drachev, S.S., 2000. Laptev Sea rifted continental margin: modern knowledge and unsolved questions. Polarforschung 68, 41–50.

Dupeyrat, L., Costard, F., Randriamazaoro, R., Gailhardis, E., Gautier, E., Fedorov, A., 2011. Effects of Ice Content on the Thermal Erosion of Permafrost: Implications for Coastal

and Fluvial Erosion: Effects of Ice Content on the Thermal Erosion of Permafrost. Permafrost and Periglacial Processes 22, 179–187. doi:10.1002/ppp.722

Embleton, C., King, C.A.M., 1975. Periglacial geomorphology, 2nd ed. Arnold, London.

Erskine, R.H., Green, T.R., Ramirez, J.A., MacDonald, L.H., 2006. Comparison of grid-based algorithms for computing upslope contributing area. Water Resources Research 42. doi:10.1029/2005WR004648

Essery, R., Li, L., Pomeroy, J., 1999. A distributed model of blowing snow over complex terrain. Hydrological processes 13, 2423–2438.

FAO, 2014. World reference base for soil resources 2014 international soil classification system for naming soils and creating legends for soil maps. FAO, Rome.

Fortier, D., Allard, M., Shur, Y., 2007. Observation of rapid drainage system development by thermal erosion of ice wedges on Bylot Island, Canadian Arctic Archipelago. Permafrost and Periglacial Processes 18, 229–243. doi:10.1002/ppp.595

Fox, J., Weisberg, S., 2010. An R Companion to Applied Regression, 2nd ed. Sage.

Franke, D., Hinz, K., Block, M., Drachev, S.S., Neben, S., Kosko, M.K., Reichert, C., Roeser, H.A., 2000. Tectonic of the Laptev Sea Region in North-Eastern Siberia. Polarforschung 68, 51–58.

Frattini, P., Crosta, G., Carrara, A., 2010. Techniques for evaluating the performance of landslide susceptibility models. Engineering Geology 111, 62–72. doi:10.1016/j.enggeo.2009.12.004

Freeman, T.G., 1991. Calculating catchment area with divergent flow based on a regular grid. Computers & Geosciences 17, 413–422.

French, H.M., 2007. The periglacial environment. John Wiley & Sons, Chichester.

French, H.M., 1971. Slope Asymmetry of the Beaufort Plain, Northwest Banks Island, N.W.T., Canada. Can. J. Earth Sci. 8, 717–731. doi:10.1139/e71-070

Gautier, E., Brunstein, D., Costard, F., Lodina, R., 2003. Fluvial dynamics in a deep permafrost zone: the case of the Middle Lena river (central Siberia), in: Eighth International Conference on Permafrost, Zurich. pp. 271–275.

Godin, E., Fortier, D., Coulombe, S., 2014. Effects of thermo-erosion gullying on hydrologic flow networks, discharge and soil loss. Environmental Research Letters 9, 105010. doi:10.1088/1748-9326/9/10/105010

Goetz, J.N., Brenning, A., Petschko, H., Leopold, P., 2015. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Computers & Geosciences 81, 1–11. doi:10.1016/j.cageo.2015.04.007

Gooseff, M.N., Balser, A., Bowden, W.B., Jones, J.B., 2009. Effects of hillslope thermokarst in northern Alaska. Eos, Transactions American Geophysical Union 90, 29–30.

Gordeev, V.V., Sidorov, I., 1993. Concentrations of major elements and their outflow into the Laptev Sea by the Lena River. Marine Chemistry 43, 33–45.

Gorsevski, P.V., Gessler, P.E., Foltz, R.B., Elliot, W.J., 2006. Spatial prediction of landslide hazard using logistic regression and ROC analysis. Transactions in GIS 10, 395–415.

Green, S.B., 1991. How many subjects does it take to do a regression analysis. Multivariate behavioral research 26, 499–510. doi:10.1207/s15327906mbr2603_7

Grosse, G., Robinson, J.E., Bryant, R., Taylor, M.D., Harper, W., DeMasi, A., Kyker-Snowman, E., Veremeeva, A., Schirrmeister, L., Harden, J., 2013. Distribution of late Pleistocene ice-rich syngenetic permafrost of the Yedoma Suite in east and central Siberia, Russia. US Geological Survey Open File Report 2013, 1–37.

_____

Grosse, G., Romanovsky, V.E., Jorgenson, T., Walter Anthony, K., Brown, J., Woods, H., Overduin, P.P., 2011. Vulnerability and Feedbacks of Permafrost to Climate Change 92, 73–74.

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

Grosse, G., Schirrmeister, L., Siegert, C., Kunitsky, V.V., Slagoda, E.A., Andreev, A.A., Dereviagyn, A.Y., 2007. Geological and geomorphological evolution of a sedimentary periglacial landscape in Northeast Siberia during the Late Quaternary. Geomorphology 86, 25–51. doi:10.1016/j.geomorph.2006.08.005

Günther, F., 2015. DEM generation for Sobo-Sise Island, Lena Delta, Siberia, using PCI Geomatica 2014.

Günther, F., Overduin, P.P., Sandakov, A.V., Grosse, G., Grigoriev, M.N., 2013. Short- and long-term thermo-erosion of ice-rich permafrost coasts in the Laptev Sea region. Biogeosciences 10, 4297–4318. doi:10.5194/bg-10-4297-2013

Guzzetti, F., Carrara, A., Cardinali, M., Reichenbach, P., 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31, 181–216. doi:10.1016/S0169-555X(99)00078-1

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., Galli, M., 2006. Estimating the quality of landslide susceptibility models. Geomorphology 81, 166–184. doi:10.1016/j.geomorph.2006.04.007

Harry, D.G., Gozdzik, J.S., 1988. Ice wedges: growth, thaw transformation, and palaeoenvironmental significance. Journal of Quaternary Science 3, 39–55.

Heckmann, T., Gegg, K., Gegg, A., Becht, M., 2014. Sample size matters: investigating the effect of sample size on a logistic regression susceptibility model for debris flows. Natural Hazards and Earth System Science 14, 259–278. doi:10.5194/nhess-14-259-2014

Hijmans, R.J., van Etten, J., 2012. raster: Geographic analysis and modeling with raster data. R package version 2.0-12.

Hjort, J., Luoto, M., 2011. Novel theoretical insights into geomorphic process-environment relationships using simulated response curves. Earth Surface Processes and Landforms 36, 363–371. doi:10.1002/esp.2048

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

Holmgren, P., 1994. Multiple flow direction algorithms for runoff modeling in grid based elevation models: an empirical evaluation. Hydrological processes 8, 327–334.

Hosmer Jr, D.W., Lemeshow, S., Sturdivant, R.X., 2013. Applied logistic regression, 3rd ed, Wiley Series in Probability and Statistics. John Wiley & Sons.

Hubberten, H.W., Andreev, A., Astakhov, V.I., Demidov, I., Dowdeswell, J.A., Henriksen, M., Hjort, C., Houmark-Nielsen, M., Jakobsson, M., Kuzmina, S., 2004. The periglacial climate and environment in northern Eurasia during the Last Glaciation. Quaternary Science Reviews 23, 1333–1357.

Huete, A.., 1988. A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment 25, 295–309. doi:10.1016/0034-4257(88)90106-X

Hugelius, G., Strauss, J., Zubrzycki, S., Harden, J.W., Schuur, E.A.G., Ping, C.-L., Schirrmeister, L., Grosse, G., Michaelson, G.J., Koven, C.D., O'Donnell, J.A., Elberling, B., Mishra, U., Camill, P., Yu, Z., Palmtag, J., Kuhry, P., 2014. Estimated

_____

stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps. Biogeosciences 11, 6573–6593. doi:10.5194/bg-11-6573-2014

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., Osterkamp, T.E., 2005. Response of boreal ecosystems to varying modes of permafrost degradation. Canadian Journal of Forest Research 35, 2100–2111. doi:10.1139/x05-153

Jorgenson, M.T., Romanovsky, V., Harden, J., Shur, Y., O'Donnell, J., Schuur, E.A.G., Kanevskiy, M., Marchenko, S., 2010. Resilience and vulnerability of permafrost to climate changeThis article is one of a selection of papers from The Dynamics of Change in Alaska's Boreal Forests: Resilience and Vulnerability in Response to Climate Warming. Canadian Journal of Forest Research 40, 1219–1236. doi:10.1139/X10-060

Jorgenson, M.T., Shur, Y.L., Pullman, E.R., 2006. Abrupt increase in permafrost degradation in Arctic Alaska. Geophysical Research Letters 33. doi:10.1029/2005GL024960

Karlsson, J.M., Lyon, S.W., Destouni, G., 2012. Thermokarst lake, hydrological flow and water balance indicators of permafrost change in Western Siberia. Journal of Hydrology 464-465, 459–466. doi:10.1016/j.jhydrol.2012.07.037

King, G., Zeng, L., 2001. Logistic regression in rare events data. Political analysis 9, 137–163.

Kokelj, S.V., Jorgenson, M.T., 2013. Advances in Thermokarst Research: Recent Advances in Research Investigating Thermokarst Processes. Permafrost and Periglacial Processes 24, 108–119. doi:10.1002/ppp.1779

Koven, C.D., Riley, W.J., Stern, A., 2013. Analysis of Permafrost Thermal Dynamics and Response to Climate Change in the CMIP5 Earth System Models. Journal of Climate 26, 1877–1900. doi:10.1175/JCLI-D-12-00228.1

Koven, C.D., Ringeval, B., Friedlingstein, P., Ciais, P., Cadule, P., Khvorostyanov, D., Krinner, G., Tarnocai, C., 2011. Permafrost carbon-climate feedbacks accelerate global warming. Proceedings of the National Academy of Sciences 108, 14769–14774.

Kunitsky, V., 2007. Nival lithogenesis and Ice Complex on the territory of Yakutia. Original Doctoral Dissertation Thesis, Permafrost Institute Yakutsk.

Kunitsky, V., Schirrmeister, L., Grosse, G., Kienast, F., 2002. Snow patches in nival landscapes and their role for the Ice Complex formation in the Laptev Sea coastal lowlands. Polarforschung 70, 53–67.

Kutzbach, L., Wagner, D., Pfeiffer, E.-M., 2004. Effect of microrelief and vegetation on methane emission from wet polygonal tundra, Lena Delta, Northern Siberia. Biogeochemistry 69, 341–362.

Kutzbach, L., Wille, C., Pfeiffer, E.-M., 2007. The exchange of carbon dioxide between wet arctic tundra and the atmosphere at the Lena River Delta, Northern Siberia. Biogeosciences 4, 869–890. doi:10.5194/bg-4-869-2007

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

Lacelle, D., Bjornson, J., Lauriol, B., 2010. Climatic and geomorphic factors affecting contemporary (1950-2004) activity of retrogressive thaw slumps on the Aklavik Plateau, Richardson Mountains, NWT, Canada. Permafrost and Periglacial Processes 21, 1–15. doi:10.1002/ppp.666

_____

Lachenbruch, A.H., 1966. Contraction theory of ice-wedge polygons: A qualitative discussion. Presented at the Permafrost International Conference: Proceedings, National Academy of Sciences-National Research Council, p. 63.

Lachenbruch, A.H., 1962. Mechanics of thermal contraction cracks and ice-wedge polygons in permafrost. Geological Society of America Special Papers 70, 1–66.

Lantuit, H., Overduin, P.P., Couture, N., Wetterich, S., Aré, F., Atkinson, D., Brown, J., Cherkashov, G., Drozdov, D., Forbes, D.L., Graves-Gaylord, A., Grigoriev, M., Hubberten, H.-W., Jordan, J., Jorgenson, T., Ødegård, R.S., Ogorodov, S., Pollard, W.H., Rachold, V., Sedenko, S., Solomon, S., Steenhuisen, F., Streletskaya, I., Vasiliev, A., 2012. The Arctic Coastal Dynamics Database: A New Classification Scheme and Statistics on Arctic Permafrost Coastlines. Estuaries and Coasts 35, 383–400. doi:10.1007/s12237-010-9362-6

Lantuit, H., Pollard, W.H., 2008. Fifty years of coastal erosion and retrogressive thaw slump activity on Herschel Island, southern Beaufort Sea, Yukon Territory, Canada. Geomorphology 95, 84–102. doi:10.1016/j.geomorph.2006.07.040

Lantuit, H., Pollard, W.H., 2005. Temporal stereophotogrammetric analysis of retrogressive thaw slumps on Herschel Island, Yukon Territory. Natural Hazards and Earth System Science 5, 413–423.

Lantz, T.C., Kokelj, S.V., 2008. Increasing rates of retrogressive thaw slump activity in the Mackenzie Delta region, N.W.T., Canada. Geophysical Research Letters 35. doi:10.1029/2007GL032433

Lantz, T.C., Kokelj, S.V., Gergel, S.E., Henry, G.H.R., 2009. Relative impacts of disturbance and temperature: persistent changes in microenvironment and vegetation in retrogressive thaw slumps. Global Change Biology 15, 1664–1675. doi:10.1111/j.1365-2486.2009.01917.x

Lauriol, B., Lacelle, D., Labrecque, S., Duguay, C.R., Telka, A., 2009. Holocene evolution of lakes in the Bluefish Basin, northern Yukon, Canada. Arctic 212–224.

Lee, S., 2005. Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. International Journal of Remote Sensing 26, 1477–1491. doi:10.1080/01431160412331331012

Lenz, J., Fritz, M., Schirrmeister, L., Lantuit, H., Wooller, M.J., Pollard, W.H., Wetterich, S., 2013. Periglacial landscape dynamics in the western Canadian Arctic: Results from a thermokarst lake record on a push moraine (Herschel Island, Yukon Territory). Palaeogeography, Palaeoclimatology, Palaeoecology 381-382, 15–25. doi:10.1016/j.palaeo.2013.04.009

Lepore, C., Kamal, S.A., Shanahan, P., Bras, R.L., 2012. Rainfall-induced landslide susceptibility zonation of Puerto Rico. Environmental Earth Sciences 66, 1667–1681. doi:10.1007/s12665-011-0976-1

Lewkowicz, A.G., Harris, C., 2005. Frequency and magnitude of active-layer detachment failures in discontinuous and continuous permafrost, northern Canada. Permafrost and Periglacial Processes 16, 115–130. doi:10.1002/ppp.522

Lewkowicz, A.G., Kokelj, S.V., 2002. Slope sediment yield in arid lowland continuous permafrost environments, Canadian Arctic Archipelago. Catena 46, 261–283.

Liston, G.E., Sturm, M., 2002. Winter precipitation patterns in arctic Alaska determined from a blowing-snow model and snow-depth observations. Journal of hydrometeorology 3, 646–659.

Lucà, F., Conforti, M., Robustelli, G., 2011. Comparison of GIS-based gullying susceptibility mapping using bivariate and multivariate statistics: Northern Calabria, South Italy. Geomorphology 134, 297–308. doi:10.1016/j.geomorph.2011.07.006

Luoto, M., Hjort, J., 2005. Evaluation of current statistical approaches for predictive geomorphological mapping. Geomorphology 67, 299–315. doi:10.1016/j.geomorph.2004.10.006

Mackay, J.R., 1990. Some observations on the growth and deformation of epigenetic, syngenetic and anti-syngenetic ice wedges. Permafrost and Periglacial Processes 1, 15–29.

Mackay, J.R., 1974. Ice-wedge cracks, Garry Island, Northwest Territories. Canadian Journal of Earth Sciences 11, 1366–1383.

Mackay, J.R., 1972. The World of Underground Ice. Annals of the Association of American Geographers 62, 1–22. doi:10.1111/j.1467-8306.1972.tb00839.x

Marmion, M., Hjort, J., Thuiller, W., Luoto, M., 2009. Statistical consensus methods for improving predictive geomorphology maps. Computers & Geosciences 35, 615–625. doi:10.1016/j.cageo.2008.02.024

Marsh, P., Russell, M., Pohl, S., Haywood, H., Onclin, C., 2009. Changes in thaw lake drainage in the Western Canadian Arctic from 1950 to 2000. Hydrological Processes 23, 145–158. doi:10.1002/hyp.7179

Martínez-Casasnovas, J.A., Ramos, M.C., Poesen, J., 2004. Assessment of sidewall erosion in large gullies using multi-temporal DEMs and logistic regression analysis. Geomorphology 58, 305–321. doi:10.1016/j.geomorph.2003.08.005

Menard, S., 2010. Logistic Regression: From Introductory to Advanced Concepts and Applications. SAGE Publications, Thousand Oaks, California.

Menard, S., 2004. Six approaches to calculating standardized logistic regression coefficients. The American Statistician 58, 218–223.

Meusburger, K., Alewell, C., 2009. On the influence of temporal change on the validity of landslide susceptibility maps. Natural Hazards and Earth System Science 9, 1495–1507.

Morgenstern, A., 2012. Thermokarst and thermal erosion: degradation of Siberian ice-rich permafrost. Dissertation, Universität Potsdam, Potsdam.

Morgenstern, A., Grosse, G., Günther, F., Fedorova, I., Schirrmeister, L., 2011. Spatial analyses of thermokarst lakes and basins in Yedoma landscapes of the Lena Delta. The Cryosphere 5, 849–867. doi:10.5194/tc-5-849-2011

Muster, S., Langer, M., Heim, B., Westermann, S., Boike, J., 2012. Subpixel heterogeneity of ice-wedge polygonal tundra: a multi-scale analysis of land cover and evapotranspiration in the Lena River Delta, Siberia. Tellus B 64. doi:10.3402/tellusb.v64i0.17301

Nefeslioglu, H.A., Gokceoglu, C., Sonmez, H., 2008. An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. Engineering Geology 97, 171–191. doi:10.1016/j.enggeo.2008.01.004

Nelson, F.E., Anisimov, O.A., Shiklomanov, N.I., 2002. Climate change and hazard zonation in the circum-Arctic permafrost regions. Natural Hazards 26, 203–225.

O'Brien, R.M., 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. Quality & Quantity 41, 673–690. doi:10.1007/s11135-006-9018-6

Ohlmacher, G.C., Davis, J.C., 2003. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. Engineering Geology 69, 331–343. doi:10.1016/S0013-7952(03)00069-3

_____

Osterkamp, T.E., Jorgenson, M.T., Schuur, E.A.G., Shur, Y.L., Kanevskiy, M.Z., Vogel, J.G., Tumskoy, V.E., 2009. Physical and ecological changes associated with warming permafrost and thermokarst in Interior Alaska. Permafrost and Periglacial Processes 20, 235–256. doi:10.1002/ppp.656

PCI Geomatics, 2014. Geomatica Help.

Petschko, H., Brenning, A., Bell, R., Goetz, J., Glade, T., 2014. Assessing the quality of landslide susceptibility maps – case study Lower Austria. Natural Hazards and Earth System Science 14, 95–118. doi:10.5194/nhess-14-95-2014

Pietroniro, A., Töyrö, J., Leconte, R., Kite, G., 2005. Remote sensing of surface water and soil moisture. Remote sensing in northern hydrology: measuring environmental change 119–142.

Pohl, S., Marsh, P., Onclin, C., Russell, M., 2009. The summer hydrology of a small upland tundra thaw lake: implications to lake drainage. Hydrological Processes 23, 2536–2546. doi:10.1002/hyp.7238

Pollozek, L., 2015. Fernerkundungs- und GIS-basierte Analyse der raum-zeitlichen Entwicklung von Thermoerosionstälern im Lena-Delta. Master Thesis, Universität Potsdam, Potsdam.

Pomeroy, J.W., Marsh, P., Gray, D.M., 1997. Application of a distributed blowing snow model to the Arctic. Hydrological processes 11, 1451–1464.

Pradhan, B., Lee, S., 2010. Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. Environmental Earth Sciences 60, 1037–1054. doi:10.1007/s12665-009-0245-8

Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sensing of Environment 48, 119–126. doi:10.1016/0034-4257(94)90134-1

R-ArcticNET, 2015. A Regional, Electronic, Hydrographic Data Network For the Arctic Region. Lena at Stolb, World Wide Web Access: http://www.r-arcticnet.sr.unh.edu/v4.0/ViewPoint.pl?View=ALL&Unit=mm&Point=6343

Rachold, V., Grigoriev, M.N., 1999. Russian-German Cooperation SYSTEM LAPTEV SEA 2000: The Lena Delta 1998 Expedition. Berichte zur Polarforschung (Reports on Polar Research) 315.

Randriamazaoro, R., Dupeyrat, L., Costard, F., Gailhardis, E.C., 2007. Fluvial thermal erosion: heat balance integral method. Earth Surface Processes and Landforms 32, 1828–1840. doi:10.1002/esp.1489

Regmi, N.R., Giardino, J.R., McDonald, E.V., Vitek, J.D., 2014. A comparison of logistic regression-based models of susceptibility to landslides in western Colorado, USA. Landslides 11, 247–262. doi:10.1007/s10346-012-0380-2

Romanovskii, N., 2004. Permafrost of the east Siberian Arctic shelf and coastal lowlands. Quaternary Science Reviews 23, 1359–1369. doi:10.1016/j.quascirev.2003.12.014

Romanovskii, N.N., Gavrilov, A.V., Tumskoy, V.E., Kholodov, A.L., Siegert, C., Hubberten, H.-W., Sher, A.V., 2000. Environmental evolution in the Laptev Sea region during Late Pleistocene and Holocene. Polarforschung 68, 237–245.

Romanovsky, V.E., Drozdov, D.S., Oberman, N.G., Malkova, G.V., Kholodov, A.L., Marchenko, S.S., Moskalenko, N.G., Sergeev, D.O., Ukraintseva, N.G., Abramov, A.A., Gilichinsky, D.A., Vasiliev, A.A., 2010. Thermal state of permafrost in Russia. Permafrost and Periglacial Processes 21, 136–155. doi:10.1002/ppp.683

_____

Romanovsky, V.E., Smith, S.L., Christiansen, H.H., 2010. Permafrost thermal state in the polar Northern Hemisphere during the international polar year 2007-2009: a synthesis. Permafrost and Periglacial Processes 21, 106–116. doi:10.1002/ppp.689

Roshydromet, 2015. Russian Federal Service For Hydrometerology and Environmental Monitoring. Climatological Information about Tiksi, Russia. World Wide Web Address: http://worldweather.wmo.int/en/city.html?cityId=1040

Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A.C., Peruccacci, S., 2010. Optimal landslide susceptibility zonation based on multiple forecasts. Geomorphology 114, 129–142. doi:10.1016/j.geomorph.2009.06.020

Rowland, J.C., Jones, C.E., Altmann, G., Bryan, R., Crosby, B.T., Hinzman, L.D., Kane, D.L., Lawrence, D.M., Mancino, A., Marsh, P., others, 2010. Arctic landscapes in transition: responses to thawing permafrost. Eos, Transactions American Geophysical Union 91, 229–230.

Rudy, A.C.A., Lamoureux, S.F., Treitz, P., Collingwood, A., 2013. Identifying permafrost slope disturbance using multi-temporal optical satellite images and change detection techniques. Cold Regions Science and Technology 88, 37–49. doi:10.1016/j.coldregions.2012.12.008

Sachs, T., Wille, C., Boike, J., Kutzbach, L., 2008. Environmental controls on ecosystem-scale CH_4 emission from polygonal tundra in the Lena River Delta, Siberia. Journal of Geophysical Research 113. doi:10.1029/2007JG000505

Schirrmeister, L., Froese, D., Tumskoy, V., Wetterich, S., 2012a. Yedoma: Late Pleistocene ice-rich syngenetic permafrost of Beringia. The Encyclopedia of Quaternary Science 542–552.

Schirrmeister, L., Siegert, C., Strauss, J., 2012b. Permafrost ein sensibles Klimaphänomen– Begriffe, Klassifikationen und Zusammenhänge. Polarforschung 81, 3–10.

Schirrmeister, L., Grosse, G., Schnelle, M., Fuchs, M., Krbetschek, M., Ulrich, M., Kunitsky, V., Grigoriev, M., Andreev, A., Kienast, F., Meyer, H., Babiy, O., Klimova, I., Bobrov, A., Wetterich, S., Schwamborn, G., 2011a. Late Quaternary paleoenvironmental records from the western Lena Delta, Arctic Siberia. Palaeogeography, Palaeoclimatology, Palaeoecology 299, 175–196. doi:10.1016/j.palaeo.2010.10.045

Schirrmeister, L., Kunitsky, V., Grosse, G., Wetterich, S., Meyer, H., Schwamborn, G., Babiy, O., Derevyagin, A., Siegert, C., 2011b. Sedimentary characteristics and origin of the Late Pleistocene Ice Complex on north-east Siberian Arctic coastal lowlands and islands – A review. Quaternary International 241, 3–25. doi:10.1016/j.quaint.2010.04.004

Schirrmeister, L., Siegert, C., Kunitzky, V.V., Grootes, P.M., Erlenkeuser, H., 2002a. Late Quaternary ice-rich permafrost sequences as a paleoenvironmental archive for the Laptev Sea Region in northern Siberia. International Journal of Earth Sciences 91, 154–167. doi:10.1007/s005310100205

Schirrmeister, L., Siegert, C., Kuznetsova, T., Kuzmina, S., Andreev, A., Kienast, F., Meyer, H., Bobrov, A., 2002b. Paleoenvironmental and paleoclimatic records from permafrost deposits in the Arctic region of Northern Siberia. Quaternary International 89, 97–118.

Schirrmeister, L., Siegert, C., Kunitsky, V., Sher, A., 1999. Paleoclimate Signals of Ice-rich Permafrost. Russian-German Cooperation System Laptev Sea 2000: the Lena Delta 1998 Expedition, Reports on Polar Research 315,145–152.

Schuur, E.A., Bockheim, J., Canadell, J.G., Euskirchen, E., Field, C.B., Goryachkin, S.V., Hagemann, S., Kuhry, P., Lafleur, P.M., Lee, H., others, 2008. Vulnerability of permafrost carbon to climate change: Implications for the global carbon cycle. BioScience 58, 701–714.

_____

Schuur, E.A.G., Crummer, K.G., Vogel, J.G., Mack, M.C., 2007. Plant Species Composition and Productivity following Permafrost Thaw and Thermokarst in Alaskan Tundra. Ecosystems 10, 280–292. doi:10.1007/s10021-007-9024-0

Schuur, E.A.G., McGuire, A.D., Schädel, C., Grosse, G., Harden, J.W., Hayes, D.J., Hugelius, G., Koven, C.D., Kuhry, P., Lawrence, D.M., Natali, S.M., Olefeldt, D., Romanovsky, V.E., Schaefer, K., Turetsky, M.R., Treat, C.C., Vonk, J.E., 2015. Climate change and the permafrost carbon feedback. Nature 520, 171–179. doi:10.1038/nature14338

Schwamborn, G., Rachold, V., Grigoriev, M.N., 2002. Late Quaternary sedimentation history of the Lena Delta. Quaternary International 89, 119–134.

Shur, Y.L., Jorgenson, M.T., 2007. Patterns of permafrost formation and degradation in relation to climate and ecosystems. Permafrost and Periglacial Processes 18, 7–19. doi:10.1002/ppp.582

Sing, T., Sander, O., Beerenwinkel, N., Lengauer, T., 2005. ROCR: visualizing classifier performance in R. Bioinformatics 21, 3940–3941. doi:10.1093/bioinformatics/bti623

Soil Survey Staff, 2014. Keys to Soil Taxonomy. 12th ed, USDA-Natural Resources Conservation Service, Washington, DC.

Soloviev, P.A., 1962. Alas relief of Central Yakutia and its formation. In: Permafrost and accompanying phenomena on the territory of the Yakutian ASSR. USSR Academy of Sciences 38–53.

Stettner, S., 2015. Characterizing thermo-erosional landforms in Siberian ice-rich permafrost. Morphometric investigations in the Lena Delta using high-resolution satellite imagery and digital elevation models. Master Thesis, Freie Universität Berlin, Berlin.

Tarnocai, C., Canadell, J.G., Schuur, E.A.G., Kuhry, P., Mazhitova, G., Zimov, S., 2009. Soil organic carbon pools in the northern circumpolar permafrost region. Global Biogeochemical Cycles 23, GB2023. doi:10.1029/2008GB003327

Toniolo, H., Kodial, P., Hinzman, L.D., Yoshikawa, K., 2009. Spatio-temporal evolution of a thermokarst in Interior Alaska. Cold Regions Science and Technology 56, 39–49. doi:10.1016/j.coldregions.2008.09.007

Ulrich, M., Morgenstern, A., Günther, F., Reiss, D., Bauch, K.E., Hauber, E., Rössler, S., Schirrmeister, L., 2010. Thermokarst in Siberian ice-rich permafrost: Comparison to asymmetric scalloped depressions on Mars. Journal of Geophysical Research 115. doi:10.1029/2010JE003640

Valentin, C., Poesen, J., Li, Y., 2005. Gully erosion: Impacts, factors and control. CATENA 63, 132–153. doi:10.1016/j.catena.2005.06.001

Van Den Eeckhaut, M., Marre, A., Poesen, J., 2010. Comparison of two landslide susceptibility assessments in the Champagne–Ardenne region (France). Geomorphology 115, 141–155. doi:10.1016/j.geomorph.2009.09.042

Van Den Eeckhaut, M., Vanwalleghem, T., Poesen, J., Govers, G., Verstraeten, G., Vandekerckhove, L., 2006. Prediction of landslide susceptibility using rare events logistic regression: A case-study in the Flemish Ardennes (Belgium). Geomorphology 76, 392–410. doi:10.1016/j.geomorph.2005.12.003

Van Everdingen, R., 2005. Multi-language glossary of permafrost and related ground-ice terms. National Snow and Ice Data Center/World Data Center for Glaciology, Boulder, CO. World Wide Web Address: http://nsidc. org/fgdc/glossary.

Vaughan, D.G., Comiso, J.C., Allison, I., Carrasco, J., Kaser, G., Kwok, R., Mote, P., Murray, T., Paul, F., Ren, J., 2013. Observations: cryosphere, in: Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the

_____

Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 317–382.

Venables, B., Ripley, B., 2015. Support Functions and Datasets for Venables and Ripley's MASS. Version 7.3-43.

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

Walker, H.J., 1998. Arctic deltas. Journal of Coastal Research 719–738.

Walter-Anthony, K.M.W., Zimov, S.A., Grosse, G., Jones, M.C., Anthony, P.M., Iii, F.S.C., Finlay, J.C., Mack, M.C., Davydov, S., Frenzel, P., Frolking, S., 2014. A shift of thermokarst lakes from carbon sources to sinks during the Holocene epoch. Nature 511, 452–456. doi:10.1038/nature13560

Walter, K., Edwards, M., Grosse, G., Zimov, S., Chapin, F., 2007. Thermokarst lakes as a source of atmospheric CH4 during the last deglaciation. science 318, 633–636.

Wang, L., Liu, H., 2006. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modeling. International Journal of Geographical Information Science 20, 193–213. doi:10.1080/13658810500433453

Wetterich, S., Kuzmina, S., Andreev, A.A., Kienast, F., Meyer, H., Schirrmeister, L., Kuznetsova, T., Sierralta, M., 2008. Palaeoenvironmental dynamics inferred from late Quaternary permafrost deposits on Kurungnakh Island, Lena Delta, Northeast Siberia, Russia. Quaternary Science Reviews 27, 1523–1540. doi:10.1016/j.quascirev.2008.04.007

Wilson, J.P., Gallant, J.C., 2000. Terrain analysis: principles and applications, 1st ed. Wiley, New York.

Wobus, C., Anderson, R., Overeem, I., Matell, N., Clow, G., Urban, F., 2011. Thermal erosion of a permafrost coastline: Improving process-based models using time-lapse photography. Arctic, Antarctic, and Alpine Research 43, 474–484.

Woo, M., Heron, R., Marsh, P., Steer, P., 1983. Comparison of weather station snowfall with winter snow accumulation in high arctic basins. Atmosphere-Ocean 21, 312–325. doi:10.1080/07055900.1983.9649171

Yalcin, A., 2008. GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): Comparisons of results and confirmations. CATENA 72, 1–12. doi:10.1016/j.catena.2007.01.003

Yang, D., Kane, D.L., Hinzman, L.D., Zhang, X., Zhang, T., Ye, H., 2002. Siberian Lena River hydrologic regime and recent change. Journal of Geophysical Research: Atmospheres (1984–2012) 107, ACL–14.

Yilmaz, I., 2009. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat—Turkey). Computers & Geosciences 35, 1125–1138. doi:10.1016/j.cageo.2008.08.007

Yoshikawa, K., Hinzman, L.D., 2003. Shrinking thermokarst ponds and groundwater dynamics in discontinuous permafrost near council, Alaska. Permafrost and Periglacial Processes 14, 151–160. doi:10.1002/ppp.451

Zevenbergen, L.W., Thorne, C.R., 1987. Quantitative analysis of land surface topography. Earth surface processes and landforms 12, 47–56.

Zhang, T., Barry, R.G., Knowles, K., Heginbottom, J., Brown, J., 1999. Statistics and characteristics of permafrost and ground-ice distribution in the Northern Hemisphere 1. Polar Geography 23, 132–154.

Zimov, S.A., Schuur, E.A., Chapin III, F.S., 2006. Permafrost and the global carbon budget. Science(Washington) 312, 1612–1613.

Zubrzycki, S., Kutzbach, L., Grosse, G., Desyatkin, A., Pfeiffer, E.-M., 2013. Organic carbon and total nitrogen stocks in soils of the Lena River Delta. Biogeosciences 10, 3507–3524. doi:10.5194/bg-10-3507-2013

Zubrzycki, S., Kutzbach, L., Pfeiffer, E.-M., 2014. Permafrost-affected soils and their carbon pools with a focus on the Russian Arctic. Solid Earth 5, 595–609. doi:10.5194/se-5-595-2014

Danksagung

In besonderer Weise bedanke ich mich bei Dr. Florian Haas, sowohl als Betreuer dieser Arbeit, als auch als langjähriger Mentor in meiner Studienzeit an der Katholischen Universität Eichstätt-Ingolstadt.

Weiterhin gilt mein tiefster Dank Dr. Anne Morgenstern für eine hervorragende und fürsorgliche fachliche Beratung während der Ausarbeitungsphase.

Ludwig Hilger danke ich für seine vielfältige Hilfe rund um Statistik und Programmierung.

Meine besondere Wertschätzung gilt allen Kollegen am Alfred Wegener Institut, mit denen ich während der Expedition und in Potsdam zusammengearbeitet habe.

Allen meinen Freunden, insbesondere meinen Eichstätter Studienkollegen, meinen Mitbewohner in Potsdam und meiner Freundin Steffi, danke ich für die Unterstützung und Begleitung durch die Zeit der Masterarbeit.

Meine größte Dankbarkeit möchte ich meinen Eltern zum Ausdruck bringen, welche mir alle Wege in meinem Studium offen gehalten haben.

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle von Autoren wörtlich übernommenen Stellen, wie auch sich an die Gedanken anderer Autoren eng anlehnende Ausführungen meiner Arbeit, sind unter Angabe der Quelle kenntlich gemacht.

Potsdam, 27. August 2015   _____

                                                                        Georg Veh