Distribution of Thermokarst Lakes and Ponds at Three Yedoma Sites in Siberia

Guido Grosse
Geophysical Institute, University of Alaska Fairbanks, USA

Vladimir Romanovsky
Geophysical Institute, University of Alaska Fairbanks, USA

Katey Walter
Institute of Northern Engineering / International Arctic Research Center, University of Alaska Fairbanks, USA

Anne Morgenstern
Alfred Wegener Institute for Polar and Marine Research, Potsdam, Germany

Hugues Lantuit
Alfred Wegener Institute for Polar and Marine Research, Potsdam, Germany

Sergei Zimov
Northeast Science Station, Cherskii, Russia

Abstract

Thermokarst lake formation in ice-rich yedoma deposits in north Siberia has a major impact on regional landscape morphology, hydrology, and biogeochemistry. Detailed assessment of lake distribution characteristics is critical for understanding spatial and temporal lake dynamics and quantifying their impacts. The distribution of thermokarst ponds and lakes at three different sites with ice-rich permafrost (Bykovsky Peninsula, SW Lena Delta, and Cherskii) in northeast Siberia was analysed using high-resolution remote sensing and geographical information system (GIS) tools. Despite similarities in geocryological characteristics, the distribution of thermokarst lakes differs strongly among the study regions and is heavily influenced by the overall hydrological and geomorphologic situation as a result of past lake-landscape dynamics. By comparing our high-resolution water body dataset with existing lake inventories, we find major discrepancies in lake distribution and total coverage. The use of low-resolution lake inventories for upscaling of thermokarst lake-related environmental processes like methane emissions would result in a strong underestimation of the environmental impacts of thermokarst lakes and ponds in Arctic lowlands.

Keywords: lake distribution; remote sensing; Siberia; thermokarst lakes; yedoma.

Introduction

Global lake inventories currently used in Earth system modeling contain only lakes larger than 10 ha (Lehner & Döll 2004, Downing et al. 2006). Most of these lakes are found in the Northern Hemisphere; that is, in permafrost-influenced or formerly glaciated regions. In permafrost regions with unconsolidated sediments most of the lakes are thermokarst lakes or ponds which formed due to the melting of massive or segregated ground ice and subsequent surface settlement. Thermokarst lakes are a major component of vast arctic and subarctic landscapes in Siberia, Alaska, and Canada. Thermokarst lakes and ponds can laterally expand by thermo-erosion and thaw slumping along shores. Usually a positive relation is established between thaw subsidence, horizontal basin extension, and water body growth, resulting in continued thawing of underlying permafrost and thermoerosion on its margins. This runaway effect is mostly dependent on ground ice content in the underlying permafrost, and it continues going until thawed-out sediments form an insulating layer preventing further thawing and subsidence. Consequently, ice-rich unconsolidated permafrost deposits like those of the widespread Late Pleistocene Yedoma Suite in northeast Siberia (Schirrmeister et al. 2008) are especially vulnerable to thermokarst development initiated by natural or anthropogenic environmental change.

Regional medium-resolution studies aimed at the classification and spatial analysis of thermokarst lakes in Arctic regions were mostly based on Landsat satellite imagery, which has proven valuable for water body detection (e.g., Frazier & Page 2000). Such studies were conducted for the North Alaska Coastal Plain (Sellmann et al. 1975, Frohn et al. 2005, Hinkel et al. 2005), the Siberian Lena Delta (Morgenstern et al. 2008), the Lena-Anabar Lowland (Grosse et al. 2006), and the Tuktoyaktuk Peninsula (Côté & Burn 2002). These and other works reveal a very large number of lakes and ponds between 1–10 ha for most permafrost lowland areas. Based on field experience and spatially limited aerial surveys, it is known that there is very likely an even larger number of thermokarst ponds smaller than 1 ha.

The lacking representation of small lakes and thus bias towards large lakes in existing global lake inventories might result in a strong underestimation of the environmental impact of thermokarst lakes, which would seriously hamper the understanding of their role in the Arctic hydrological cycle and global biogeochemical cycles. First qualitative and quantitative studies on the environmental impact of thermokarst lakes over both geological and historical time frames suggest a potentially large role of such lakes in the...
Arctic carbon cycle by unlocking vast amounts of permafrost-
stored organic carbon and thus also for global climate
dynamics (Zimov et al. 1997, Walter et al. 2006, Walter et
al. 2007). Other studies indicate e.g., the impact of arctic
wetland and lake distribution on atmospheric circulation
patterns (Gutowski et al. 2007). Many studies also show that
small water bodies in permafrost lowlands, e.g., polygonal
ponds, are an important component for methane emissions
from tundra wetlands (e.g., Wagner et al. 2003, Schneider
et al. in review). To quantify the environmental impact of
such arctic lakes and ponds it is necessary to have detailed
information on their distribution and extent.

Several studies suggest a direct connection between
thermokarst lake distribution and underlying permafrost
characteristics, as well as between lake dynamics and
permafrost dynamics. Widespread lake drainage in the
discontinuous permafrost zone is related to a beginning
disappearance of permafrost due to current climate warming
(Yoshikawa & Hinzman 2003, Smith et al. 2005, Riordan et
al. 2006). On the contrary, in the continuous permafrost zone
an increase in lake area was observed in Western (Smith et
al. 2005) and Eastern Siberia (Walter et al. 2006), but not
on the Alaskan North Slope (Riordan et al. 2006, Hinkel et
al. 2007). Most of these studies are based on the medium-
resolution (30–80 m) Landsat satellite data record spanning
about 35 years and some additionally involve historical
aerial photography. However, although Landsat data may be
sufficient to detect complete drainage and drying of some
large lakes or the formation of new large lakes, it has serious
limitations in detecting changes in lake extent due to thaw
slumping and thermo-erosion. A high-resolution dataset
of recent lake and pond distribution in combination with
historical imagery would form an excellent base for studying
change in lake extent and link these changes to permafrost or
climate dynamics.

In this study a dataset of recent thermokarst lake and pond
distribution on ice-rich yedoma deposits is developed based
on high-resolution (1–2.5 m) satellite remote sensing data
for three sites in Northeast Siberia and then compared with
data from the Global Lake and Wetland Database (GLWD:
Lehner & Döll 2004).

**Study Region**

All three study sites are situated in the continuous
permafrost region of northeast Siberia (Fig. 1). The size of
the study areas, a set of basic environmental characteristics,
and information about the geocryology are provided in Table
1. The Bykovsky Peninsula (BYK) is situated southeast of
the Lena River Delta. The peninsula is surrounded on three
sides by large bays and the Laptev Sea. It is an erosional
remnant of a Late Pleistocene accumulation plain consisting
predominantly of silty to sandy ice-rich permafrost deposits
of the yedoma (Grosse et al. 2007). Lakes, predominantly of
early Holocene thermokarstic origin, are abundant at BYK
(Grosse et al. 2005). They are located either on the yedoma
uplands or as lake remnants and polygonal ponds in drained
lake basins. These basins were formed by large thermokarst
lakes during the early Holocene climate warming. They
subsequently shrunk or drained after the middle Holocene
due to climate deterioration and coastal erosion, leaving behind
lake remnants and drained lake basins, which upon freezing

---

![Figure 1. Relief map showing the location of the three study sites: Olenek Channel (OLE), Bykovsky Peninsula (BYK), and Cherskii (CHE) in NE Siberia including maps showing the lake distribution. Small inset in upper right indicates the location of the large map (black rectangle) in Siberia and the distribution of continuous (dark grey) and discontinuous permafrost (medium grey) after Brown et al. (1998).](image-url)
could support formation of polygonal ponds (Grosse et al. 2007). The land surface at BYK ranges from 0–45 m a.s.l. could support formation of polygonal ponds (Grosse et al. 2007). The land surface at BYK ranges from 0–45 m a.s.l.

Table 1. General environmental characteristics of the study areas.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>OLE</th>
<th>BYK</th>
<th>CHE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location (°N)</td>
<td>72.94°</td>
<td>71.80°</td>
<td>68.75°</td>
</tr>
<tr>
<td>Study area (ha)</td>
<td>122.90°E</td>
<td>129.30°E</td>
<td>161.33°E</td>
</tr>
<tr>
<td>Permafrost depth (m)</td>
<td>7982</td>
<td>17 009</td>
<td>28 897</td>
</tr>
<tr>
<td>Active layer depth (m)</td>
<td>200–600</td>
<td>500–600</td>
<td>400–500</td>
</tr>
<tr>
<td>Annual ground temperature (°C)</td>
<td>0.3–0.6</td>
<td>0.3–0.6</td>
<td>0.3–1.5</td>
</tr>
</tbody>
</table>

(20 m depth) -9– -11 9– -11 -3 – -11

<table>
<thead>
<tr>
<th>Climate data for closest weather station (Rivas-Martinez 2008):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Station name</td>
</tr>
<tr>
<td>Annual air temperature (°C)</td>
</tr>
<tr>
<td>Annual precipitation (mm)</td>
</tr>
<tr>
<td>Vegetation zone</td>
</tr>
<tr>
<td>Yedoma thickness (m)</td>
</tr>
<tr>
<td>Gravimetric ice content (%)</td>
</tr>
<tr>
<td>Ice wedge width / length (m)</td>
</tr>
<tr>
<td>Organic carbon content (%)</td>
</tr>
</tbody>
</table>

Table 2. Satellite imagery used for mapping lake distribution.

<table>
<thead>
<tr>
<th>Site</th>
<th>Platform</th>
<th>Date</th>
<th>Ground Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>BYK</td>
<td>Spot-5</td>
<td>2006-07-09</td>
<td>2.5m</td>
</tr>
<tr>
<td>OLE</td>
<td>Spot-5</td>
<td>2006-07-08</td>
<td>2.5m</td>
</tr>
<tr>
<td>CHE</td>
<td>Ikonos-2</td>
<td>2002-07-09</td>
<td>1.0m</td>
</tr>
</tbody>
</table>

Table 3. Main parameters of lakes and ponds in the study areas.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>OLE</th>
<th>BYK</th>
<th>CHE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of water bodies N</td>
<td>15,012</td>
<td>13,001</td>
<td>13,48</td>
</tr>
<tr>
<td>Total water body area A (ha)</td>
<td>1619.6</td>
<td>2226.1</td>
<td>242.3</td>
</tr>
<tr>
<td>Limnicty (%)</td>
<td>13.3</td>
<td>15.4</td>
<td>0.8</td>
</tr>
<tr>
<td>Largest lake size (ha)</td>
<td>196.19</td>
<td>605.00</td>
<td>16.71</td>
</tr>
<tr>
<td>Mean water body size (ha)</td>
<td>0.0706</td>
<td>0.0217</td>
<td>0.1797</td>
</tr>
<tr>
<td>Median water body size (ha)</td>
<td>0.0088</td>
<td>0.0075</td>
<td>0.0115</td>
</tr>
<tr>
<td>Normalized per 10 000 ha:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of water bodies N</td>
<td>18,808</td>
<td>7644</td>
<td>466</td>
</tr>
<tr>
<td>Total water body area A (ha)</td>
<td>1327.6</td>
<td>1541.6</td>
<td>83.8</td>
</tr>
</tbody>
</table>

Methods

A simple density slice classification was applied to the most recent images at each site to distinguish water and land in the panchromatic imagery. A threshold that best separated image pixel values (Digital Numbers, DN) of water from land was chosen. Usually, there is a strong difference in reflectance between water bodies (dark or black, low DN) and bare or vegetated land surfaces (bright, high DN). A visual comparison was conducted to verify the classification. The DN of some lakes was found to be influenced by either very shallow water levels (probably less than 1 m), resulting in higher DN due to reflectance of the lake bottom, or turbid water with high sediment suspension, resulting in higher DN from the sediment load. On some lakes remaining lake ice (highly reflective, very high DN) resulted in misclassification. Additional misclassifications occurred for pixels associated with deep thermo-erosional valleys and steep north-facing cliffs or slopes. In both cases shadows were misclassified as water. These misclassifications were corrected by applying manually generated masks in ArcGIS™ to either exclude pixel (shadows, stream water bodies, man made structures) or to include pixel (lake ice, turbid and shallow water) from the lake dataset. Based on our visual examination and manual correction the resulting datasets can be considered a conservative minimum of standing water bodies in the study areas.

Based on the ground resolution of the available imagery, a minimum of 5 pixels was considered acceptable for successful water body detection. For better comparison between all study areas we therefore only included and analyzed standing water bodies larger than 0.003 ha. Since our classification approach was conservative and aiming at open water surfaces, we assume to have missed especially small ponds overgrown by vegetation in our inventory. At BYK, three thermokarst lagoons were included in the dataset, since they morphologically belong to the peninsula. Eventually, ArcGIS™ was used to analyze the spatial distribution of water bodies in the resulting datasets. The resulting dataset was then compared with the GLWD dataset (Lehner & Doell 2004) for all study areas.

Remote Sensing Data

Spatially high-resolution recent satellite imagery was acquired from all three study sites to study thermokarst lake distribution and extent (Table 2). All images are from the snow-free early summer period.

For the BYK and the OLE sites two Spot-5 images were georeferenced to topographic maps of scale 1:100,000, for CHE several Ikonos-2 images were ortho-rectified (Table 2). All imagery used in this study is panchromatic and showed excellent contrasts for land-water separation.
Despite similar basic geocryological conditions at all three sites (ice-rich yedoma deposits), the distribution of thermokarst lakes strongly differs among them. A total number of 29,361 lakes >0.003 ha were classified in the study areas (Table 3). The highest lake cover by land area (limnicity) was found at BYK (15.4%), closely followed by OLE (13.3%). Though CHE is the largest study area, its lake portion is lower than at both other sites by more than one order of magnitude (0.8%).

Of the lakes >10 ha (49 lakes), 14 belong to OLE (609 ha), 32 to BYK (2053 ha), and 3 to CHE (44 ha) (Fig. 2). When comparing the water body distribution in various size classes, strong differences between the three study sites become even more obvious (Fig. 3). The OLE region has more than double the number of small ponds (0.003–0.01 ha) when compared to the BYK region, and almost 50-fold that of the CHE region. This disparity is also expressed in the low mean lake size for OLE. The dominance of lake numbers in the OLE area is true for all size classes except the three largest (>5 ha), where BYK dominates.

Comparing the area-normalized (per 10,000 ha) number N of water bodies larger than area A versus the area A in a logarithmic scaled diagram reveals an almost linear trend for the lake size distribution in all three study sites (Fig. 2). The possibility of describing a lake distribution with such a power law function of type $y=ax^b$ is well known from investigations of other large lake inventories (Lehner & Döll 2004, Downing et al. 2006). It seems to be also applicable to the NE Siberian dataset of comparably small-sized thermokarst lakes investigated in this study. However, obviously for the CHE water bodies the trendline is situated about one order of magnitude below BYK and OLE.

Comparison of lake density between the sites shows a generally highly dense water body population at OLE due to a large amount of small ponds, the clustering of water bodies and thus lake density at BYK, and the overall sparse lake cover at CHE. At CHE, the presence of Rodinka hill in the study area and the lack of lakes on the bedrock hill itself creates an additional cause for the scarcity of lakes. However, it appears not to be the main reason, since flat areas with unconsolidated yedoma farther away from the hill also have much lower lake densities than OLE or BYK. Image analysis and ground truthing reveal that the large lakes in the BYK area usually occur in deep thermokarst basins (subsided up to 40 m below surrounding surface) and often are only the lake remnants of previously partially drained lakes. Many polygonal ponds are situated in these drained basins. Small ponds and medium lakes are found on the yedoma uplands with poor drainage. Yedoma uplands with many thermo-erosional valleys do not contain many lakes. At OLE, most of the numerous, small ponds are found on the yedoma upland. There are only a few large, drained thermokarst basins. Remarkably, the OLE area is relatively homogeneously covered with numerous small ponds, while these small water bodies occur in irregular patterns at BYK and CHE. There are comparably few small thermokarst lakes at CHE. Many of
these were formed due to human activity in the area around the settlement of Cherskii, adding an artificial component to the natural lake distribution. Similar, at BYK human impact resulted in the drainage of several small lakes along vehicle tracks in the tundra. In contrary to CHE, the overall lake population is very high at BYK and therefore the human influence on the lake distribution characteristics is probably negligible here.

A comparison of our detailed water body dataset with the lakes in the GLWD (Lehner & Döll 2004) reveals large discrepancies in number and area of thermokarst lakes larger than 10 ha (minimum size GLWD) (Fig. 2). As a result, GLWD limnicity at OLE (1.1%), BYK (6.8%), and CHE (0.0%) is significantly lower than in our dataset for the same lake size category (7.6%, 12.1%, and 0.2% respectively). A large percentage of the total water body area (<10 ha) that is important for hydrological and biogeochemical cycles is currently not inventoryed and used in environmental modeling (not inventoryed water body area per 10 000 ha land area at OLE: 42.7%, BYK: 21.6%, and CHE: 82.2%).

**Discussion**

Thermokarst lake distribution in our study areas seems to be strongly connected to hydrological and geomorpho-logical factors rather than to geocryology alone. At BYK, a strong thermokarst relief developed during the Late Pleistocene-Holocene transition and many first-generation thermokarst basins and thermo-erosional valleys were formed (Fig. 5A). A large number of second-generation lakes appeared during the Holocene in drained basins. OLE is dominated by a flat yedoma surface with only some thermokarst basins and valleys. Also, yedoma thickness is considerably less than at both other sites with possibly impacts on lake expansion dynamics. The poorly drained upland plain is densely and relatively homogeneously packed with a large number of first generation lakes (Fig. 5B). At CHE (Fig. 5C) the yedoma is mantling the rolling slopes around Rodinka Hill. Comparably few small water bodies occur, but proportionally many medium-sized first-generation lakes. Compared to both other study sites, lakes are less abundant, most likely due to better overall drainage. However, there are plain regions east of Rodinka Hill with the same deposits but still very low lake density.

Vegetation either growing or floating in the lake can pose a challenge for any water-land classification method, be it manual or fully automated. Careful image interpretation, field experience, and in some cases ground truthing are required. We estimate the effect of unclassified water due to vegetation on the order of <2% of the overall water body area for some lakes. Seasonal hydroclimatology can also have an effect on lake surface area especially for lakes with shallow basin topography. While many of the ponds certainly fall into this category, many of the typical thermokarst lakes in ice-rich permafrost do not. These usually have steep banks and a more pronounced basin morphology resulting predominantly in vertical lake level changes rather than lateral lake area changes during seasonal water level variations. Many of the larger lake basins have subsided tens of meters below the surrounding land surface and are surrounded by steep banks. Lake extent changes for such thermokarst lakes are more related to thermoerosion, thaw slumping, or drainage than just seasonal water level changes.

A lake change study for the study areas is still in progress. The time series of lake extent for the BYK site will range for the period from 1951–2008, providing an observational high-resolution dataset for 57 years. For OLE and CHE the time series span a period of 42 and 37 years, respectively.

**Conclusions**

High-resolution satellite imagery provides the opportunity to characterize the distribution of thermokarst lakes on yedoma deposits in high detail. Spatial analysis of thermokarst features plays an important role in understanding thermokarst dynamics in northern regions and impacts on the global hydrological and biogeochemical cycles. It was shown that thermokarst lake distributions at three yedoma sites differ greatly. Lake distribution is distinguishable for areas with first and second generation lakes. Our comparison with the GLWD lake dataset demonstrates the necessity to quantify northern lakes; that is, thermokarst lakes, in a much higher detail than currently available. Any quantification and upscaling of thermokarst lake-related parameters like methane emission might be biased due to the exclusion of a large number of small lakes and ponds not represented in current global databases. This highlights the need for more intense research on thermokarst distribution and lake dynamics.
Acknowledgments

Grosse and Walter were partially funded by University of Alaska IPY fellowships. Spot imagery for OLE and BYK was provided through the OASIS program. ikonos imagery for CHE was provided through the NSF AON project.

References


Frazier, P.S. & Page, K.J. 2000. Water Body Detection and Delineation for CHE was provided through the NSF AON project.


References


Grosse and Walter were partially funded by University of Alaska IPY fellowships. Spot imagery for OLE and BYK was provided through the OASIS program. Ikonos imagery for CHE was provided through the NSF AON project.


