

Methane Fluxes from Arctic Lakes

Homero Alejandro Paltan Lopez
May, 2013

Methane fluxes from Arctic lakes

by

Homero Alejandro Paltan Lopez

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

Project Supervisors:

Dr. Jadu Dash
Professor Mary Edwards

UNIVERSITY OF
Southampton

Disclaimer

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

Abstract

Increases in greenhouse gases emissions, in particular methane (CH₄), have been found to be one of the drivers of global warming. Specially, arctic lakes are defined as significant but scarcely studied sources of methane. For a sensitive region as the arctic, current uncertainties related to the lack of a detailed dataset is constraining the understanding of the process. This in turn has led into misestimations of the real lake areal extension and overlook of small lakes (<10ha.), which are likely to emit more methane per unit area.

The aim of this research was to define the role of Arctic lakes in the regional methane budget. To adress this, firstly there was developed a New Arctic Lake Geodatabase (NALGDB). There were extracted lakes, from 379 Landsat 5 +TM cloud free images from the summer months between 2003 – 2012. Secondly, a new regional value for methane emissions from lakes was obtained. The NALGDB was the base for upscaling detailed field measurements conducted by previous research. Finally, a 'top-down' approach was followed to define the significance of small lakes. Point measurements of CH₄ atmospheric concentrations were used, as retrieved by the sensor TANSO-FTS onboard of GOSAT Satellite.

In this way, there were obtained some 2'000,000 lakes with a surface greater than 3,600 km². Lakes in the area are typically small as 85% of them were classified into the category of small. Moreover, the contribution of lake methane fluxes to the atmosphere was estimated at $5.65 \pm 1.24 \text{ Tg. CH}_4 \text{ y}^{-1}$. There were located hotspots of emissions in the Yedoma region in Siberia, and in Central Russia around the Gulf of Ob and the Kheta and Yenisey River basins. Lastly, it was found that areas with more small lakes are likely to present high levels of methane in the atmosphere. Methane values in these areas are approaching those measured by the same sensor in rice field in China (known for its high methane emissions).

Finally, this study asserts the significance of the usage of a detailed lake dataset. For the Eurasian Arctic in particularly, this has enriched the discussion of lakes spatial distribution. In addition, the map of the spatial distribution of methane fluxes suggests the existence of new hotspots of methane emissions that are typically not considered (Central Russia). The later alongside with the proven significance of small lakes into the regional methane flux, need further validation.

Key words: Arctic, Lakes, Methane fluxes, Remote Sensing, GIS

Acknowledgements

First, I would like to thank my family, in both Ecuador and Mexico, who have always supported me to accomplish my goals, not just in recent years but throughout my life.

I would like to thank my supervisors Prof. Mary Edwards and Dr. Jadu Dash for their critical advice and particularly for their support and academic guidance along this research.

A special acknowledgement is to my friends, those who are in Ecuador and those I met throughout this period in the Netherlands and in the UK. All those great experience will be always in my mind.

Last but not least, I would like to thank the Ecuadorian Government and the People of Ecuador for the financial support I received to complete my postgraduate studies.

*Look deep into nature, and then you will understand everything
better (Albert Einstein)*

Table of Contents

Methane Fluxes from Arctic Lakes	i
Abstract	vii
Acknowledgements	ix
Table of Contents	xiii
List of Figures	xv
List of Tables.....	xvii
List of Abbreviations.....	xix
1. Introduction	1
1.1 Arctic Lakes	2
1.1.1 Lakes Datasets.....	4
1.1.1.1 Remote sensing for mapping lakes	5
1.2 Methane emissions from Arctic Lakes	7
1.3 Measuring methane lake emissions	9
1.3.1 The Bottom-up approach.....	9
1.3.2 The top-down approach.....	11
1.4 Research Problem.....	15
1.5 Research Objectives.....	16
1.5.1 General Objective.....	16
1.5.2 Specific Objectives	16
1.5.3 Research Questions	17
2. Material and Methods	19
2.1 Study Area	19
2.1.1 Characterisation of the Area	20
2.2 Materials	23
2.2.1 Imagery Sources	24
2.2.2 Thematic Data.....	24
2.2.2 Secondary Sources	27
2.3 Methods	27
2.3.1 Developing the New Arctic Lakes Geodatabase	27
2.3.2 Analysing the arctic lakes spatial distribution.....	31
2.3.3 Estimating methane emissions from arctic lakes	32
2.3.4 Defining the role of small lakes in methane atmospheric concentrations	33
3. Results & Discussion.....	37
3.1 A new Arctic lakes geodatabase.....	37
3.1.1 Threshold sampling	37
3.1.2 The Geodatabase Development	37
3.1.2 The New Arctic Lake Geodatabase.....	38
3.1.3 Accuracy Assessment of the NALGDB	40
3.1.4 Spatial Distribution of Arctic Lakes	41
3.2 Methane emissions from arctic lakes	47
3.2.1 Spatial Pattern of CH ₄ from Arctic Lakes	49

3.2.2 Comparison with current methane estimations	51
3.2.3 Implications of a non-detailed dataset for up scaling field measured fluxes	53
3.3 Role of small lakes into CH ₄ atmospheric concentrations.....	55
3.3.1 Spatial relationship between small lakes and CH ₄ atmospheric amounts	55
3.3.2 The role of small lakes.....	59
3.3.3 Overview of factors affecting the relationship small lakes and CH ₄ column amounts	64
4. Conclusions.....	67
5. Recommendations.....	69
References.....	71

List of Figures

Figure 1: Study Area	19
Figure 2: Map of Typical Vegetation Units	21
Figure 3: Map of Permafrost Extent.....	22
Figure 4: Map of Terrain Features	26
Figure 5: Flowchart of the development of the New Arctic Lakes Geodatabase	30
Figure 6: Diagram of the approach for processing GOSAT-TNS single observation points	34
Figure 7: Map of Lake Density	44
Figure 8: Map of Lake Fraction	46
Figure 9: Map of Distribution of total CH ₄ fluxes from arctic lakes ...	50
Figure 10: Total CH ₄ fluxes from Arctic Lakes: Comparison of results using the NALGDB and GLWD.....	54
Figure 11: Correlation between lake density and lake fraction for small lakes.....	56
Figure 12: Regression for the relationship between amount of water of small lakes and CH ₄ mean total column abundance.	57
Figure 13: Regression for the relationship between total lake area and CH ₄ mean total column amounts for cluster samples that intersect with wetlands and cluster samples do not intersect.	58
Figure 14: Selected clusters for analysing the role of lakes in the abundance of CH ₄ in the atmosphere.....	60
Figure 15: Scatter plot and error bars of the distribution of lakes classes of clusters and total amounts of CH ₄	62
Figure 16: Representation of clusters with similar high CH ₄ column amount values.	63

List of Tables

Table 1: Overview of the data and materials used in this study	23
Table 2: CH ₄ fluxes values adopted in this study.....	27
Table 3: Summary of DN value statistics of the shore of the lake defined by Vegetation Class	37
Table 4: Comparison of lakes size classes between the NALGDB and the GLWD	39
Table 5: Accuracy Assessment of the NALGDB.....	41
Table 6: Number of Lakes and areal statistics for the Eurasian Arctic.	42
Table 7: Up scaling of CH ₄ gas emissions of the Eurasian Arctic lakes.	48
Table 8: Classification of CH ₄ mean total column abundances.....	60
Table 9: Atmospheric CH ₄ amounts per classes of lakes.....	61
Table 10: Predominant vegetation classes found within each class of lake area.	64

List of Abbreviations

CH₄ Methane

DN Digital Number (value of a pixel)

GOSAT Greenhouse gases Observing SATellite

GLWD Global Lake World Database

Ld Lake density

Lf Lake fraction

NALGDB New Arctic Lakes GeoDatabase

TANSO – FTS Fourier Transform Spectrometer

Tg. Teragrams (10^{12} grams)

1. Introduction

Climate is agreed to be the main external factor that controls the functioning of global ecosystems (Maxwell, 1992) . At this time there is a general consensus that in the past years the change of climate has been likely to influence average global air and ocean temperatures, melting of snow and ice layers and rising sea levels, among other natural phenomena (IPCC, 2007, IPCC, 2001). Such changes and particularly temperature increases are very likely to happen due to the increase in greenhouse gas emissions into the atmosphere (IPCC, 2007).

More specifically, methane (CH₄) has been defined as an atmospheric trace gas with significant properties which greatly influence the greenhouse phenomenon. It is, in addition, being alarmed because its concentration has doubled since the industrial revolution and is in fact increasing at a rate of approximately 1% per year (IPCC, 2001, Rhode, 1990). Of this, recent data suggest that wetlands account for 70% of natural source emissions and about 20% of total annual global emissions. (Bousquet et al., 2006, Zimov et al., 1997, Roulet et al., 1994, IPCC, 2001). Thus, the study of methane emissions in wetland regions has become an important scientific matter.

In this way, due to their high sensitive characteristic, arctic ecosystems have drawn the attention of the international community. Recent global model scenarios project, under the increase of greenhouse gas concentrations, an Arctic annual mean warming about twice that of the global mean warming (Yamanouchi, 2011, IPCC, 2007, Chapin et al., 2000) . In addition, such arctic sensitivity is also defined by their role as natural carbon reservoirs and as a significant emission source compared with temperate and tropical ecosystems (Ping et al., 1997, Dixon et al., 1994). Thus, changes in the structure, function and composition of arctic ecosystems are known to have profound implications in the carbon cycle (Oechel and Vourlitis, 1994).

These changes are known to have led to deep effects in the Arctic. For instance, the shrinking of the sea ice extent, changes in lakes freeze-up dates, decrease of mountain glaciers extent and ice caps, reduction of the areal extent of frozen ground and effects on the permafrost zone are among the most significant consequences (Lemke et al., 2007). Such effects at the same time, may in turn enhance methane dynamics in the region and therefore defining a positive feedback on global climate. Thus, the Arctic is expected to

react earlier to climate phenomena (Walter et al., 2006, Grosse et al., 2008).

In this way, particular attention must be paid when analysing the processes that explain arctic sources of methane emissions. Nevertheless, typical studies have mainly focused on understanding vegetation and land ecosystem dynamics when measuring the methane budget across this region. Recent research suggests that production in terrestrial plants can account for up to 50% of modern methane sources (Keppler et al., 2006). Therefore, this is calling for a reconsideration of the role of other natural sources in the carbon cycle.

However, the role of aquatic systems, and more specifically the role of lakes into the methane budget, is still poorly understood. This is despite their important presence in the regional landscape and more importantly, despite the significant source of methane that these lakes may represent (Walter et al., 2006). This highlights that there is still more to study about arctic lakes, their processes associated with gas production and their connection with regional methane fluxes.

1.1 Arctic Lakes

Representing an important part of the arctic landscape, lakes occupy about 30% of their land surface (Hinkel et al., 2007, Hinkel et al., 2003). Such distribution is not spatially homogeneous and therefore responds to different landscapes characteristics. For instance, Smith et al. (2007) identified climate, geomorphology, substrate permeability, glaciation history and particularly the presence of permafrost as the most important factors which define lake distribution patterns in northern regions. The results of their research is presented in the following.

They accounted for ~ 2'00.000 lakes in regions northwards 45 °N. Additionally, they denoted that lake densities and area fraction averages are between 300-350% greater in glaciated terrain versus unglaciated, and 100–170% greater in permafrost areas versus permafrost-free terrains. In addition, the presence of peatlands is associated with an increase of about 40–80% in lake density and 10–50% increases in area fraction. Surprisingly, lake statistics were found to be similar across continuous, discontinuous and sporadic permafrost zones, whereas a modest decrease was seen in isolated permafrost, and sharp drop in the absence of permafrost. Finally, they determined that lakes are most abundant in glaciated, permafrost peatlands with a current rate of approximately 14.4

lakes/1000 km², and least abundant in unglaciated, permafrost-free terrain, 1.2 lakes/1000 km². It is remarkable then to observe the effect by which permafrost is defining the abundance of lakes along arctic landscapes.

Permafrost is defined as any subsurface material (soil or rock) that remains below 0 °C for at least two consecutive years (Brown et al., 2001b). Typically, it is classified as continuous (90–100%), discontinuous (50–90%), sporadic (10–50%), or isolated patches (0–10%). The thickness of permafrost has been found to vary from centimetres to even a meter (Anisimov and Reneva, 2006). Warm temperature and therefore thawing leading to the degradation of these sub-surfaces can significantly affect the hydrological and ecological functioning of arctic areas (Frey and McClelland, 2009). This in turn can cause thickening of the active layer (the permanent seasonally thawed uppermost layer), thermokarst development, expansion or creation of thaw lakes, among others (e.g. Zhang *et al.*, 2005). The role of permafrost in prompting lake creation is then essentially by reducing infiltration from surface into the subsurface and by its role in the thermokarst process. Therefore, the thawing of large areas of continuous permafrost is of particular interest, as it in turn may redefine the arctic landscape's functioning.

Thawing lake cycles have been extensively discussed by authors. The principal point of difference has been both the explanation by which ice aggradation defines lake formation, and the treatments of the return of surfaces to near-original conditions to complete the cycle. In response to this, recently Jorgenson and Shur (2007) developed a conceptual model which integrates and reviews previous proposed models. Their revision model is based on 6 stages of lake development: (1) initial flooding of primary lakes, (2) lateral expansion (3) lake drainage, (4) differential ice aggradation in silty centers and margins, (5) secondary development of thaw lakes and infilling ponds along the ice-poor margins; and (6) lake stabilisation.

These linked processes, are then likely to be continuously re-defining arctic landscapes structures and the composition of lacustrine ecosystems. Principally, it is recognised that sediments are being redistributed during the stage of lake expansion. This may in turn, affect the composition of lake water and thereby impacting on the geochemical cycle, as will be reviewed in the coming section. Moreover, in the stage of secondary development of thaw lakes, it is formed of numerous small, shallow ponds around the margins of the drained basins. This is primarily caused by the water collected in the lowest portions of the basins that can form large ponds, and then

over time, organic accumulation and ice aggradation in the land can cause the ponds to become subdivided into small ones. Additionally, during this stage organic matter is added to the benthic layers of the pond and to the vegetation that surrounds it. In this way, the representation of an arctic lake as a spatial frame is a dynamic process, the understanding of which must include several variables.

1.1.1 Lakes Datasets

The veracity of those studies that seek to comprehend lake dynamics then relies on an accurate land cover classification. More specifically, the importance of counting with a detailed lake geographic dataset becomes substantial. For water bodies in general, the most important available source for medium scale and regional studies has been found to be the Global Lake World Database (GLWD). Considered as the finest scale geographic database available, it was developed by Lehner and Döll (2004) aiming to integrate and combine a GIS approach global digital and analog datasets of lakes and wetlands. This remarkable cartographical effort was the first milestone for the assessment of global lakes and other water bodies based on their area, shape, and location, and specially it offered the most extensive advance in the inventory of world lakes since Halbfass's compilation of 1914 (Downing and Duarte, 2009).

The GLWD comprises three levels of data. Level 1 covers polygons of long lakes ($\geq 50 \text{ Km}^2$). Level 2 contains small lakes ($\geq 0.01 \text{ Km}^2$). Finally, Level 3 includes reservoirs, rivers and wetlands types. However, due to their spatial variation and methodological procedures of the GLWD, a large number of arctic water bodies are omitted or underestimated on their area. This is particularly true for small lakes (less than 10Ha). (Grosse et al., 2008, Walter et al., 2006). The effective omission of lakes of this database has never been globally or regionally quantified.

However, despite its drawbacks, for regional and global studies the GLWD has been widely used among researchers. Typical hydrological applications include, among other, river discharge modeling, hydraulic characterisations and continental carbon assessments. Despite the fact that most known studies suggest that this is not an accurate database, their final results still rely on its surface values. A rough comparison between the GLWD and Aster scenes developed by Walter et al. (2007b), defined that about 50% of lakes are omitted in the mentioned dataset. In this particular case, they apply such value to correct the total lake area estimations.

Finally, such aerial uncertainties have led, in many cases, to the dependence on national sources, high resolution imagery or aerial photography for mapping water bodies. These approaches have been used particularly for local studies where the availability of these types of dataset is more feasible. However, in a regional scale, remote sensing techniques should be considered as an efficient way to map lakes as it is discussed hereby.

1.1.1.1 Remote sensing for mapping lakes

Traditional field techniques for mapping lakes, as it is with other water bodies, have been considered difficult due to their logistic problems and high costs. In order to overcome such shortcoming, remote sensing represents a cost-effective tool which provides the required information within a temporal framework (Carter, 1982; Töyrä, Pietroniro, & Martz, 2001). This technique draws upon the spectral characteristics of water by which it absorbs and reflects the light measured by a sensor.

In this way, common methods base their approach on the high absorption values by which water has in the visible and near infrared regions of the electromagnetic spectrum (Fu, Wang, & Li, 2007; Jiahang, Currit, & Xuelian, 2010; McFeeters, 1996). However, detection of water bodies should not rely merely on the optical characteristics of the water. Several approaches have been implemented to classify them due to the fact that some other features could be erroneously classified as water. For this reason, extraction of lakes from remote sensing imagery should consider additional properties as Jiahang et al. (2010) extensively discussed and is hereby presented.

Such factors are principally the changing spectral characteristics of water bodies, the amount – frequency of them and their not static shape. The first relates with the risk of using merely one band to detect water. The risk to do so is the known movement of the peak of water towards longer wavelength as the concentrations of mud and sand increase. This constraint, in addition, comes along with the risk of misclassifying a cloud, a shadow, a mountain slope as a water body. Secondly, the frequency of water bodies is related to the expectation of detects lots of small water bodies which make the workload more sensible. Lastly, the no static shape dictates that there can be found a water body at any geometric feature. This property rises from the physical characteristic that determines that water will take the shape of its container.

In this way, based on the spectral characteristics of water and on its additional properties, there are available several remote sensing techniques for their detection. Existent literature mainly divides them into two main categories: threshold sampling and machine learning approaches (Wang & Zhu, 2009). The first approach uses statistical information about the bands defines a suitable threshold to detect water pixels. The second approach is related to the use of algorithms which build knowledge based on rules for the analysis of remote sensing imagery (Huang & Jensen, 1997).

The application of techniques related to both categories has been extensively used for mapping lakes in particular (Wang & Zhu, 2009). Firstly, threshold sampling studies seizes the light which is absorbed in the near and middle infrared portion of the spectrum (0.70 – 1.75 μm) to differentiate water pixels from land pixels. This technique mainly depends on the appropriate threshold value. Secondly, typical machine learning techniques include Tasseled Cap Transformation (TCT), neural network, decision trees and maximum likelihood. The first one tends to convert individual band pixels into wetness values so that water bodies can be distinguished. Neural network techniques develop an algorithm where seed pixels are selected where water values are recorded. Water bodies are then extracted based upon finding similar patterns than those initially recorded. Decision trees, are build up upon series of decisions that a region of pixels have to meet in order to be considered as water (Fu et al., 2007). Maximum likelihood, finally, is based on the pattern recognition of pixels assuming a Gaussian distribution of them (Duong, 2012).

The presented extensive availability of techniques to recognize water bodies from remote sensing however has opened the discussion towards finding the most efficient extraction method. For instance, maximum likelihood has been criticized for its long exhaustive computational cost which makes it unviable for large region studies (Kumar, 2010). Conversely, the study conducted by Roach, Griffith, and Verbyla (2012) demonstrated that threshold extraction or density slicing is the most efficient method to detect water bodies. The study compared this method with a decision tree and an object – based classification. Such finding was previously asserted by Wang and Zhu (2009) who compared a threshold method with TCT and a LVQ neural network. Due to its accurate results and its efficiency for resources demanding, arguably that density classification (a threshold technique) is the best way to detect water bodies.

Finally, besides the appropriate technique selection it is important the consideration of the sensor. An accurate mapping of water bodies

need a sensor which cover a range of hydrological conditions with a high temporal resolution (Frazier, 2000). For its revision time, historical archive and spatial coverage Landsat sensors usually represents an efficient tool for these studies. Pioneer water detection's application, dates back to 1972 when the satellite was first launched (Smith, 1997). Typical recent applications have successfully applied a density slicing approach for mapping wetland/lakes extent as Frazier (2000) reports. Specifically, the same study conducted by Roach et al. (2012) demonstrated that Landsat band 4 is the most suitable band for water bodies detections. In addition, there wasn't found any significant increase in accuracy by adding band 5 to the analysis as previous research typically did.

1.2 Methane emissions from Arctic Lakes

As expressed in the above section, the cycle of arctic lakes is accompanied by dynamic natural processes. For instance, one of the most significant processes of lakes is the mechanism by which they release methane from the water into the atmosphere. There are several natural factors that are controlling the methane release process in northern lakes. At least four principal pathways have been identified: ebullition flux, storage flux, diffusive flux, and flux through aquatic vegetation (Bastviken et al., 2004). Such pathways have different underlying explanations and affect in differing magnitude to the methane flux process.

In this way, firstly, ebullition accounts for between 40 and 60% of total emissions from a lake (Walter et al., 2006). It has been defined as particularly effective, as sediments bypass oxidation that can occur in the water column or in adjacent oxygenated soil, and therefore directly reach the atmosphere (Chanton, 2005). Recent field observation has identified ebullition points as specific classes of bubble cluster or open holes in the lake ice. Secondly, the storage component flux which accounts for about 45% of total emissions, is mainly associated with stratified lakes. The availability of degradable organic matter increases decomposition and oxygen consumption which leads to oxygen depletion and consequently triggers the increase of gas emissions (Huttunen et al., 2003) Thirdly, diffusion occurs when methane transported from the sediments is not affected by oxidation and then reaches the upper mixed layer of the water column. The process has been found to be more significant in larger lakes, accounting for as much as 50% of their total emissions. The final release pathway is the plant-mediated flux. In this mechanism, vegetation influences the microenvironment, the decomposition rates

and therefore, regulates net emissions. (Morrissey and Livingston, 1992). This last process has been mainly found in wetlands.

Moreover, the location of a certain lake has been found to determine magnitude of its gas emission. Specifically, thermokarst is agreed as being a process that significantly boosts emissions from arctic lakes (Shakhova et al., 2009, Walter et al., 2006, Grosse et al., 2008). Defined as the melt of permafrost, thermokarst is known to be one of the most significant processes in the arctic. Firstly, as reviewed previously, permafrost is likely to determine the formation of new lakes rich in organic matter. It affects an existent lake in that this thermal erosion process can occur along lake margins and at the edges of thaw bulbs beneath lakes (Walter et al., 2007b). The ground surface collapses and releases ancient frozen organic carbon. Finally these materials are entering and affecting the decomposition process and boosting production of methane.

Furthermore, especially in northern Siberia, many lakes are underlain by yedoma terrain types. The characteristics of yedoma, also called 'ice complex', are the accumulation of high ice contents in the frozen ground and the occurrence of synergic ice wedges. This permafrost type was formed in the Pleistocene due to the accumulation of fine grained permafrost deposited under continental, cold climate conditions in tundra arctic environments (Wetterich et al., 2011). The principal characteristic of such soils is their current organic carbon richness with typical contents ranging from 2% to 5% of content (Zimov et al., 2006). This yedoma beneath thermokarst lakes is likely to be decomposed and therefore substantially fuel the methane production process in lakes. It has been observed that during the stage of lake migration, about 30% of yedoma carbon is converted to methane.

In this way, CH₄ release dynamic studies must include the definition of a lake's characteristics and its surrounding environment. Thus, Bastviken et al. (2004) & Walter et al. (2007b) have demonstrated that lake emissions are mainly associated with their surface area. A negative relation between gas concentration and lake area can be reflected by more extensive oxidation in the mixed layer of larger lakes due to their higher piston velocity and longer residence time. Furthermore, lakes with low area are more likely to be influenced by their shoreline processes. These processes, such as thermokarst and microphyte production, are estimated to enhance gases production and emission. Therefore, small, recently formed lakes are expected to have the largest fluxes per unit area.

In Addition, Zhu et al. (2010) have defined environmental variables that affect lakes emission in glaciated covered areas. A significant negative correlation between gas emissions and Daily Total Radiation (TDR) and air temperature was mainly found. This negative correlation suggests that these environmental variables may play an important role in the release of trace gases. This happens firstly by controlling algae photosynthesis and secondly by mineralising organic matter presented in the water. These two processes are considered mainly to boost the production of carbon (CO₂). Conversely, CH₄ emissions have been demonstrated to be significantly correlated with local air temperature, water table and total dissolved solids. Thus, such lake characteristics are mainly influenced by their surrounding ecosystem.

For all of these reasons, it is essential to understand lakes' surrounding ecosystems when studying gas dynamics. In such a way, Huttunen et al. (2003) have defined that the main influence of a surrounding ecosystem of a lake is by enriching it with nutrients. More specifically, Bastviken et al. (2004) analysis demonstrate that along lake's size, the lake productivity and the load of allocthonous organic carbon are influencing such dynamics. In addition, Boereboom et al. (2012) have strengthened this spatial connection by asserting that CH₄, as explained above, occurs mainly in permafrost rich areas where there are effects of thermokarst erosion, whereas CO₂ is released from gas diffusions near the forest treeline and in northern boreal regions. Therefore, the analysis of lakes fluxes dynamics, besides incorporating lakes' morphological characteristics, must include a characterisation of the surrounding ecosystem.

1.3 Measuring methane lake emissions

1.3.1 The Bottom-up approach

Typical methods to estimate the amount of gas emitted in lakes, are based on field flux measurements along the study site (Stow et al., 1998, Takeuchi et al., 2011, Roulet et al., 1994). Common techniques include: eddy covariance flux towers, headspace equilibration, automated chambers and static closed chambers (Denmead, 2008). This is followed by an up scaling and extrapolation of those local gas field measurements in order to determine regional or global estimations. This approach, traditionally used for land emission, has been successfully applied for the study of lakes fluxes as well. These procedures are commonly referred as 'bottom - up' approach (Frankenberg et al., 2006).

Specifically for the estimations of arctic lake emissions, most of the studies have primarily focused on the estimation of diffusive flux. For instance, Kling et al. (1992) firstly estimated carbon and methane diffusive fluxes from lakes in Alaska. The principal finding was a warning regarding the significant contribution that carbon released from aquatic systems may have in regional gas budget estimations. Additional research then asserted such conclusions by field measurements of concentrations profiles of CH₄ and CO₂ in lakes and peatland reservoirs in Canada (Duchemin et al., 1995, Hamilton et al., 1994). However, such studies did not account for the most important pathway of methane release into the atmosphere: ebullition.

The principal constraint found to measuring this type of emission lies in the high randomness in both the temporal and spatial occurrence of bubbles across lake surfaces (Walter et al., 2006). To overcome this, firstly research by Zimov et al. (1997) and Zimov et al. (2001) defined a strong positive relationship between active thermokarst and ebullition rates. Such observations were then addressed by Bastviken et al. (2004) and Walter et al. (2006) to enhance the understanding of CH₄ release by redefining field sampling techniques along different types of lakes. Such studies, mainly using floating or submerged chambers, redefined the role of arctic lakes into the regional and global gas budget.

In this way, Bastviken et al. (2004) conducted a study to define local regression equations for surface CH₄ concentrations in boreal lakes. Their research related: areal extent of a lake, concentration of dissolved organic carbon, and concentration of total potassium in order to determine the storage per m² and the anoxic volume fraction. Their equations yielded an emission average of 12 g C m⁻² yr⁻¹ for boreal lakes. They then up scaled those measurements to define gases emissions in a global scale. However since the equations are based on a boreal basis that does not include tropical and temperate particular variables, the results may be biased towards an arctic reality and therefore represent a risk for its interpretation.

Moreover, research conducted by Walter et al. (2007b) effectively upscaled direct field CH₄ measurements in order to estimate the role of methane released from lakes into the regional budget. Results obtained from the study significantly redefined the role of lakes as sources of methane to the atmosphere. Such research were a continuation of initial field samples conducted along Siberian lakes (Walter et al., 2006). The principal strength of those findings lies in

the clear measurement and inclusion of methane point source bubbling.

They assumed that 90% of lakes in the continuous permafrost zone are thermokarst. The ebullition flux rate for these lakes was fixed at $34.5 \pm 9.5 \text{ g CH}_4 \text{ m}^{-2} \text{ y}^{-1}$ (The numbers following a \pm are standard deviation). Thus, the thawing of permafrost along lake margins is accounting for most of the methane released from lakes (Walter et al., 2006). For the rest of the areas, i.e. non-thermokarst and non-continuous permafrost, the rate was defined at $17.9 \pm 12.1 \text{ g CH}_4 \text{ m}^{-2} \text{ y}^{-1}$. Lastly, the diffusive flux was assumed as constant for the whole study area at $1.0 \pm 0.2 \text{ Tg CH}_4 \text{ m}^{-2} \text{ y}^{-1}$ based on the rates measured by Kling et al. (1992).

Such results redefined the importance of the contribution of lakes' ebullition to the Arctic carbon budget. The total amount of gases released by this mechanism was defined as $24.2 \pm 10.5 \text{ Tg CH}_4 \text{ m}^{-2} \text{ y}^{-1}$ whereas the amount by diffusion was of $24.2 \pm 10.5 \text{ Tg CH}_4 \text{ m}^{-2} \text{ y}^{-1}$. In addition, it was found that nearly 50% of all the emissions occur in the continuous permafrost zone. Discontinuous, sporadic and isolated areas of permafrost accounted for about 10% of the total emissions each one. Finally, it was found that the remaining 20% of the emissions was from non-permafrost zones. It is important to note that these values were calculated for lakes situated from 45° northwards.

Finally, recent research by Walter et al. (2010) refined the mentioned field measured rates by measuring ebullition flux according to their type. Four types of ebullitions seeps by CH_4 were classified. Firstly, isolated bubbles in multiple ice layers. Secondly, merged bubbles in multiple ice layers. Thirdly, single gas pockets stacked in ice. And finally, relatively open holes in winter lake ice or hotspots. Long term measurements were conducted accounting variability in CH_4 concentrations and by using records of atmospheric and including hydrostatic pressure to calculate the moles of CH_4 contained within measured volumes of bubble. Compared with the previous sampling, gas fluxes obtained from this study were about 35% lower, whereas for the later types of emissions the rate increased by 18% and 48% respectively. So far this is the most detailed known characterisation of arctic lakes methane fluxes

1.3.2 The top-down approach

As explained previously, the most common method to estimate gas emissions, is the regional extrapolation of field measurements, or so

called 'bottom-up' approach. However, arctic field studies can be criticised for their considerable financial costs and the difficult accessibility of the areas. Furthermore, studying gas emissions in such regions involves additional constraints due to their high spatial and temporal variability (Walter et al., 2007b, Frankenberg et al., 2006). To overcome such difficulties, the combination of remote sensing techniques with Geographic Information Systems (GIS) is often the most cost-effective tool (Grosse et al., 2006). Firstly, medium spatial resolution sensors (30 – 80m.) and more specifically Landsat satellite imagery, may present support for mapping the arctic. This has been proven valuable for detecting water bodies and can be adapted for its wide usage in arctic regions (Grosse et al., 2008, Frohn et al., 2005, Bolch et al., 2010, Roach et al., 2012).

Nevertheless, Landsat satellite imagery is not sufficient to detect arctic lake processes such as energy fluxes. This limitation is defined by the characteristics of its instrument's sensor. The spectral resolution, measurement frequency and radiometric characteristics of TM-sensor, and other high spatial resolution instruments, are not optimal for detecting surface water processes (Kalio, 2012). As a consequence, in order to estimate and trace geochemical processes, such as gas fluxes, from arctic lakes, it is necessary to combine and alternate different types of remote sensing tools.

Frequently, these procedures are known as 'top-down' approach (Frankenberg et al., 2006). The main characteristic of such approach is the use of point measurements and atmospheric transport models in order to estimate source and sinks of CH₄ (Hein et al., 1997, Houweling et al., 1999, Fletcher M. et al., 2004, Chen, 2003, Bergamaschi et al., 2005). Originally, this approach was mainly associated with fixed observation sites of CH₄ which were principally constrained by the limited number of atmospheric observation sites. However, recent advances in remote sensing techniques may offer series of possibilities to overcome such restrictions, as discussed below.

Initially, the application of Synthetic Aperture Radar (SAR) has been used to detect winter formed bubbles in lake upper layers associated with densely packed CH₄ (Morris et al., 1995, Duguay et al., 2002, Hall et al., 1994, Jeffries et al., 1994). These bubbles have been detected as bright areas in the imagery under the application of backscatter models. The potential of such research relies on the strong relationship found between C-band VV polarisation signature and ice thickness, snow ice layer, internal bubble structure among other features. Additionally, such findings have been recently

approached by Walter et al. (2008) and Engram et al. (2012), to associate SAR imagery with trapped bubbles associated with ebullition. Principally their results suggest that the analysis of L-band backscatter intensity could support the understanding of ebullition bubbling.

However, constraints of such methods may be associated with the limited spatial coverage and the uncertainties that still surround such predictions. Those uncertainties are mainly related to the high resolution georeferencing procedures that are necessary to effectively associate bubbles with pixel values. Additionally, interference has been found of artifacts such as snow ice, moist snow, slush beneath dry snow, among others natural constraints. Overcoming these issues and thereby improving these techniques then may represent a significant improvement in the up scaling of field gas measurements. Thus, recent progress in sensor technology allows the scientific community to trace atmospheric gases at regional or global scale with a reasonable temporal frequency. This, may also represent a significant step towards understanding the regional CH₄ dynamics.

1.3.2.1 Instruments to measure the spectroscopic properties of atmospheric gases.

In order to measure atmospheric gases, satellite instruments usually recover the spectroscopic properties of atmospheric trace gases and the underlying surface i.e., the lower troposphere (Palmer, 2008). Typically, these sensors measure the composition of the troposphere in a Sun-synchronous low-Earth orbit (between 200 and 1000km). Thus it is assured the observations are at the same local time over different areas. In addition, the majority of these sensors measure using nadir geometry in order to overcome constraints presented in the upper parts of the troposphere, such as the presence of optically thick clouds.

Furthermore, the spectral characteristic of the sensor is also known to play a significant role in tracing atmospheric gases (Palmer, 2008). Thus, sensors with nadir measurement geometry read principally the backscattered solar radiation at ultraviolet/visible (UV/Vis), short-wave infrared (SWIR) wavelengths and thermal infrared (TIR). The first two wavelengths are characterised as being sensitive to clouds, aerosols, and Rayleigh scattering. In this way, these types of sensors will be more sensitive to the lower troposphere. Conversely, TIR observations will be more sensitive to both the middle and upper troposphere, and zones of high thermal tropospheric contrast.

Recent efforts have been aimed at the development of sensors that reliably trace gas concentration data in the troposphere. The latest advances in detector technology have improved the instruments' sensitivity, the spectrometry system and spatial coverage that has led to produce more reliable measurements (Palmer, 2008). Hence, Bréon and Ciais (2010) suggest TOVS (NOAA), AIRS - IASI (NASA), OCO (NASA), A-SCOPE, SCHIAMACHY and GOSAT as the most significant current satellite sensors to study spatial and temporal distribution of CO₂ and CH₄. Thus, the use of remote sensing for directly retrieving greenhouse gas concentrations, such as CO₂ and CH₄, has been defined as a novel approach. (Butz et al., 2011). More specifically, the use of GOSAT instrument for such purposes has been commonly accepted due to its assessed accuracy.

1.3.2.2 The GOSAT Sensor.

GOSAT, developed by the Japanese Space Agency, is the first satellite dedicated to monitoring greenhouse gas densities. The strength of this sensor relies on its ability to sample the low atmospheric layers, which are connected to the surface fluxes. The instrument, launched in 2009, is the first satellite especially dedicated to measure concentrations of CO₂ and CH₄. It provides observational data to ascertain the global distribution of such gases. In addition, this data can be used to analyse how sources and sinks of these gases vary within spatial and temporal dimensions (GOSAT Project Office, 2012).

The data of GOSAT is retrieved by two types of sensor equipped on board the satellite. Firstly, the Thermal and Near Infrared Sensor for carbon Observation Fourier Transform Spectrometer (TANSO-FTS) with four bands in both the Short Infrared spectral region (SWIR) and in the Thermal Infrared spectral region (NIR). This sensor retrieves sunlight reflected from the earth's surface and light emitted from the atmosphere and the surface. It is particularly designed to detect molecular absorption for radiation with high spatial resolution. Secondly, the TANSO Cloud Aerosol Imager (TANSO-CAI) with four bands designed to detect clouds and aerosols. The latter is usually used for atmospheric corrections on the spectra obtained with FTS.

The output data of GOSAT consists of 4 level products. Level 1 product is mainly radiance spectral products. Level 2 products contain data about the column amount of CO₂ and CH₄ columns observed by FTS SWIR sensor and data of vertical mixing profile for the FTS NIR. Level 3 products are globally - monthly CO₂ and CH₄ data for: the column abundance in the SWIR and concentration at each vertical

level in the TIR. Additionally, there is data retrieved from the CAI sensor about clear sky reflectance and globally NDVI. Finally, Level 4 products refer to reports of regional gas fluxes and globally three dimensional gas models.

The total column amount measurements presented in Gosat Level 2 products is defined then as the amount of the gas in a vertical column of unit cross section extending from the Earth's surface to the top of the atmosphere (Japan Aerospace Exploration Agency, 2011, Basher, 1982). These values are expressed as number of molecules per unit area. In addition, those column measurements present the advantage that they exhibit less variability than surface data, since they are not influenced by planetary boundary layer dynamics. At the same time, the column amount of a gas retain information about surface fluxes (Palmer, 2008). For instance, previous research comparing FTS installed network with eddy covariance measurements showed that the column measurements have potential for directly observing, although this feature was constrained by its difficulty in accounting for atmospheric transport. In addition, the use of the FTS sensor to measure CO₂ concentrations has been relatively recently accepted for its accuracy (Washenfelder et al., 2006). In this way, the wide range of point measurements of GOSAT can therefore complement the existing ground network for monitoring atmospheric concentrations of greenhouse gases.

1.4 Research Problem

The reviewed limitations of the Global Lake World Database (GLWD) are halting the complete understanding of arctic phenomena. Even though the GLWD is the most detailed available dataset, its areal miss estimations and small lakes omissions have led to freshwater studies to be constrained. Firstly, this this has led to a current uncertainty about the complete understanding of the spatial distribution of arctic lakes. In this way, it is yet not clear the real spatial distribution of arctic lakes and the landscape feature that may be regulating them.

Moreover, besides the hydrological and landscape characterisation misunderstanding that this implies, the use of a non-detailed dataset may also lead to significant errors when explaining natural processes, more specifically, the understanding of the role of methane fluxes from arctic lakes in the regional and global gas budget. This in addition includes a lack of comprehension of the spatial pattern of gas emissions in a regional basis which are mainly based on upscaling of local measurements. The principal constraint that limits such

understanding then is the uncertainties surrounding the estimations of a real lake areal extent in the region. In previous research, calculations have been based on the available GLWD and thereby inherited its uncertainty. Therefore, this in turn has caused current efforts to upscale and estimate CH₄ fluxes from arctic lakes to hold an undefined accuracy.

Finally, the uncertainties of lakes frequency and distribution have affected too the understanding of the impact of small lakes on the methane budget. As reviewed previously, authors have extensively discussed the inverse relation between lake area and amount of gas released. Thus, considering that small lakes are reportedly the most frequent in the area, the regional impact of this type of lake in the CH₄ regional budget might be neglected. All this suggests that the role of arctic freshwater ecosystems, and more specifically lakes, in the regional methane process is still poorly understood.

1.5 Research Objectives

1.5.1 General Objective

The aim of the present study is to investigate the role of Arctic lakes in the regional methane regional budget.

1.5.2 Specific Objectives

In order to address the aim of the stud, three specific objectives are defined for this research:

- Map the spatial distribution of Arctic lakes.
- Estimate the methane emissions from arctic lakes in the Eurasian region.
- Investigate the role of small lakes in the regional methane budget.

1.5.3 Research Questions

Objective 1.

- What are the lakes omissions of the existent GLWD?
- How can the spatial distribution of lakes be explained?

Objective 2.

- What is the spatial pattern of CH₄ fluxes from arctic lakes?
- To what extent has the usage of non-detailed datasets affected previous estimations of lakes emissions?

Objective 3.

- Do areas with small lakes significantly emit CH₄ to the atmosphere?
- Can CH₄ atmospheric concentration be related with small lakes geographic distribution?

Introduction

2. Material and Methods

2.1 Study Area

For this study the arctic region of Europe and Asia has been selected as study area. This comprises the Arctic continental land area above 65° N (Figure 1).

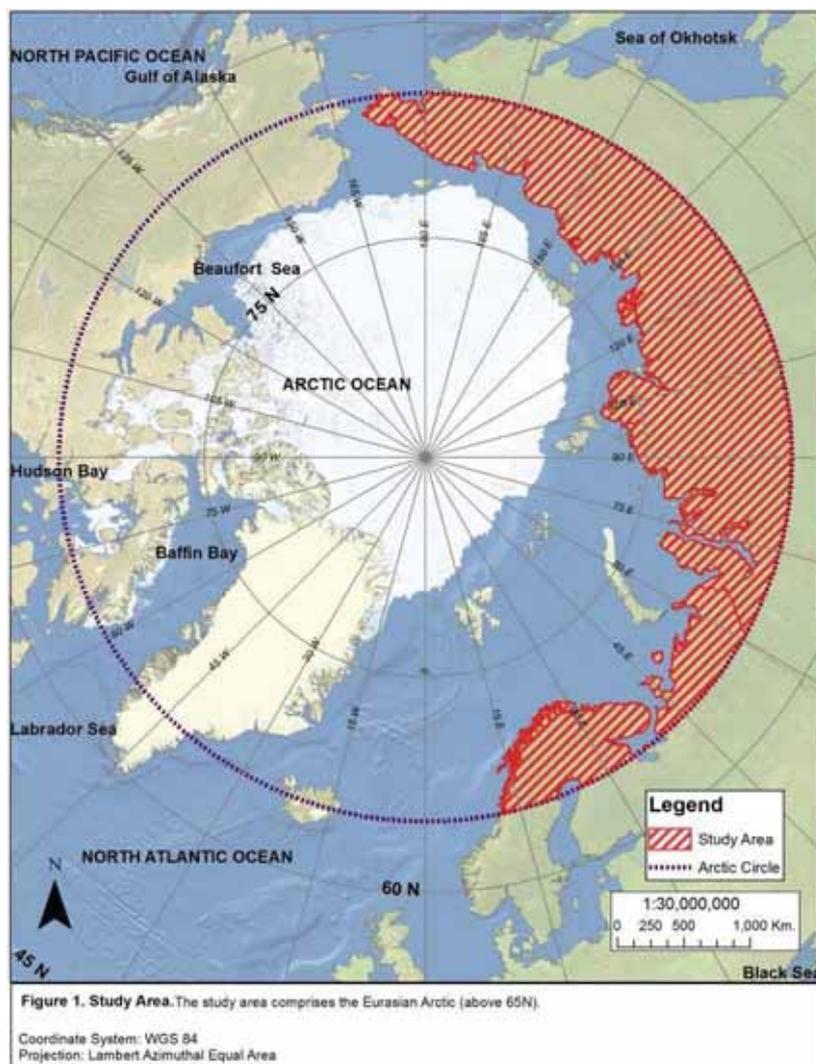


Figure 1: Study Area

2.1.1 Characterisation of the Area

According to Kaplan et al. (2003) 6 vegetation units have been identified in the study area (Figure 2). However, their spatial distribution pattern is not uniform. Firstly, cold deciduous forest and cold evergreen needleleaf forest are the land cover units with major significant presence in the area. They both occupy nearly 65% of the total study area. Their spatial extent is remarkable particularly in the central-eastern part of the study area i.e. Russia. Secondly, there are erect dwarf-shrub tundra and cold evergreen needleleaf units with individual values of 18% and 14% respectively. The former is mainly located in the northern inland areas of the region, whereas the latter can be found in central and northern areas of the European arctic. Lastly, there are prostrate dwarf-shrub tundra occupying above 4% of the area followed by cool evergreen needleleaf forest and cushion forb which extension is scarcely 1% of the total area. The first one and the last one are particularly presented in the northern shoreline of Russia while cool evergreen needleleaf forests are surrounding the Gulf of Bothnia.

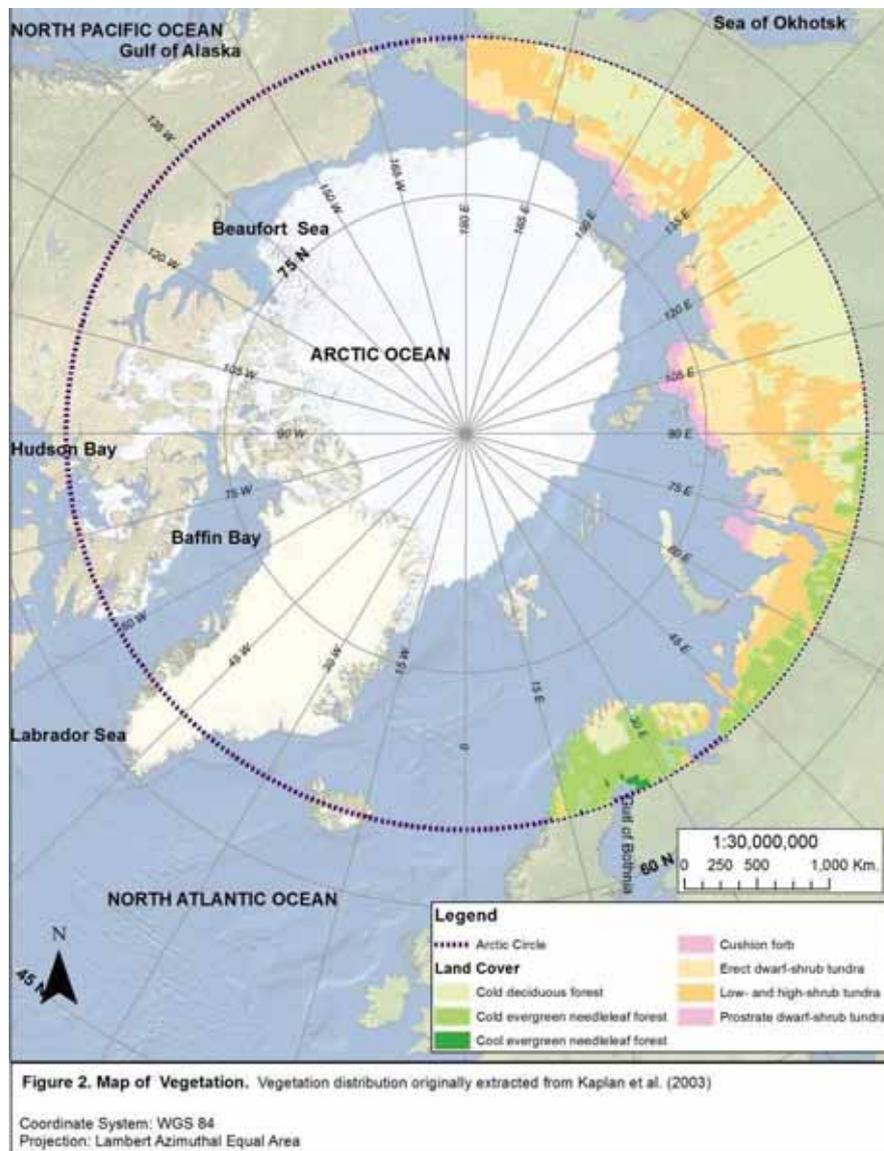


Figure 2: Map of Typical Vegetation Units

Permafrost extends throughout nearly the complete study area. Figure 3 shows the distribution of permafrost in the Eurasian Arctic. It has been classified based on the map developed by Brown et al. (2001b) into continuous, discontinuous, sporadic and isolated patches. Continuous permafrost underlies the highest percentage of the landscape than do the rest of the units (nearly 80%). This is particularly observed in the Eurasian and Asian areas of the study area. Discontinuous, sporadic and isolated patches' surface areas are

ranging from 2 to 10%. Finally, it is important to highlight that about 10% of the terrain is classified as 'Land' which is understood to be areas with no presence of permafrost.

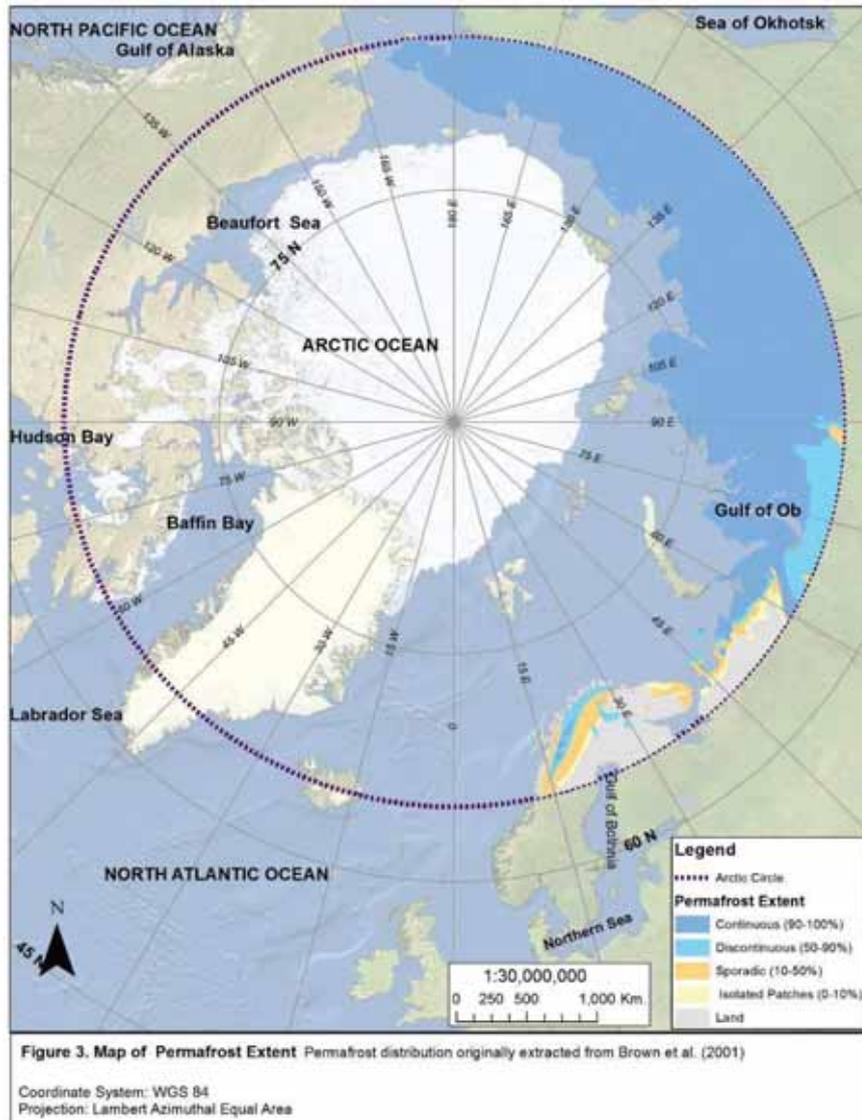


Figure 3: Map of Permafrost Extent

2.2 Materials

The following table summarises the materials used in this research:

Table 1: Overview of the data and materials used in this study

Type	Material	Temporal Frame	Spatial Coverage	Spatial resolution / Scale	Source
Images	Landsat 5 TM+	Summer Months 2006 - 2012	Eurasian Arctic	30m x 30m	USGS
	Ikonos	Summer Months 2006 - 2012	1 Images per Land Cover Class	4m x 4m	Google Earth
	GOSAT FTWS Level 2 CH ₄ atmospheric concentrations	Summer Months 2009 - 2012	Eurasian Arctic	IFOV = 10.5 Km	Japan Aerospace Exploration Agency
Spatial Thematic	Eurasian Boundary Layer	2009	Continental Eurasian Arctic (above 65 oN)	1:1'1000,000	ESRI
	Land Cover	2003	Eurasian Arctic	1:1'100,000	Kaplan et al.
	Permafrost	2001	Eurasian Arctic	1:1'100,000	Circum-arctic Map of Permafrost (Brown et al.)
	Yedoma Extension	2007	Eurasian Arctic	1:1'100,000	Walter et al.
	Terrain Altitudes	2010	Eurasian Arctic	1 sq. Km.	USGS GTOPO 30 DEM
Field Data	Diffusive fluxes from lakes	1990	Lakes in Alaska	Local	Kling et al.
	Ebullition from lakes	2006	Lakes in Alaska	Local	Walter et al.

Table 1 presented thereby the 3 main components of the data and material used in the study: images, thematic spatial data and field data. Following, there is an extended explanation of the way by which the data was obtained.

2.2.1 Imagery Sources

As base data, in this study Landsat 5 TM+ imagery was used as the source sensor for extracting water bodies. The dataset was downloaded from the United States Geological Survey (USGS) online server. It was attempted to cover the continental Eurasian arctic in its whole extension for the open water - summer months from 2009 to 2012. For areas with cloud noise, the minimum temporal threshold was diminished from 2009 to 2003. Due to cloud noise this purpose was not successfully fully achieved. In order to overcome such difficulty, the minimum temporal threshold was diminished from 2009 to 2003. In addition, for certain areas, especially in Northern Taymiria, central northern of Russia, there were no images available of the quality required i.e. less than 10% of cloud coverage. This area was excluded from the analysis. The final dataset includes 379 final images.

In addition the accuracy assessment stage of the new lake dataset it was obtained high resolution imagery. One Ikonos image per land cover class was obtained from Google Earth. The location and temporal resolution of these images relies upon availability of them.

In addition, satellite retrieved data was used to define the total column amount of methane. In this study GOSAT FTS SWIR Level 2 data products were used. The images were accessed through the portal of the Japan Aerospace Exploration Agency. Considering the satellite's launching date, and gas release connection with open water months, daytime data of the months of June, July and September from the years 2010, 2011, and 2012 was retrieved. A total of 2998 HDF5 format point measurements for the daytime hours were then acquired along the study area. It is remarkable to add as well, that the 2998 points do not represent 2998 different locations. In most of the cases, the point measurements are located in the same location as a previous measurement or within no significant distance difference.

2.2.2 Thematic Data

Thematic data embraces those datasets which were used as spatial input for the study. Such process was developed under a GIS platform and geodatabase. It was specified a Lambert Azimuthal Equal Area map projection for the study. In addition, a Eurasian boundary map was extracted from the ESRI 1:1.000.000 World Basemap. As defined above, the southern limit was defined at 65°N latitude whereas the northern limit was defined by the continental

margin. Once the basic area of interest had been defined, it was added thematic information to the geodatabase.

Firstly, a land cover characterization was added. For this purpose, the standardized circumpolar vegetation distribution proposed by Kaplan et al. (2003) was used. This dataset embraces 15 classes described at the biome level, including a detailed classification scheme for tundra vegetation. This modeled classification represents a comprehensive support to study arctic land cover patterns. Secondly, data was included about the permafrost characteristics of the zone. This data was obtained from the digital Circum-arctic Map of Permafrost developed by Brown et al. (2001a). Permafrost was categorized as continuous (90-100%), discontinuous (50-90%), isolated (10-50%) and sporadic (0-10%). This data was re-projected and added to the geodatabase.

Finally, information about the terrain characteristics was integrated. Initially, it incorporated a DEM layer to identify lowlands and highlands. This was obtained by the GTOPO30 global digital elevation model (DEM) with a horizontal grid spacing of 30 arc seconds (approximately 1 kilometer) from the U.S. Geological Survey. The boundary defined arbitrarily by Smith et al. (2007) between lowlands and highland in 300 meters above the sea level was adopted. Moreover, a map of yedoma areal extension originally presented by Walter et al. (2007a) was added, which is based on a recompilation of external sources. This map was georeferenced and manually digitalised so that it could be included in the database of the study. The resultant map of these terrain variables is presented below:

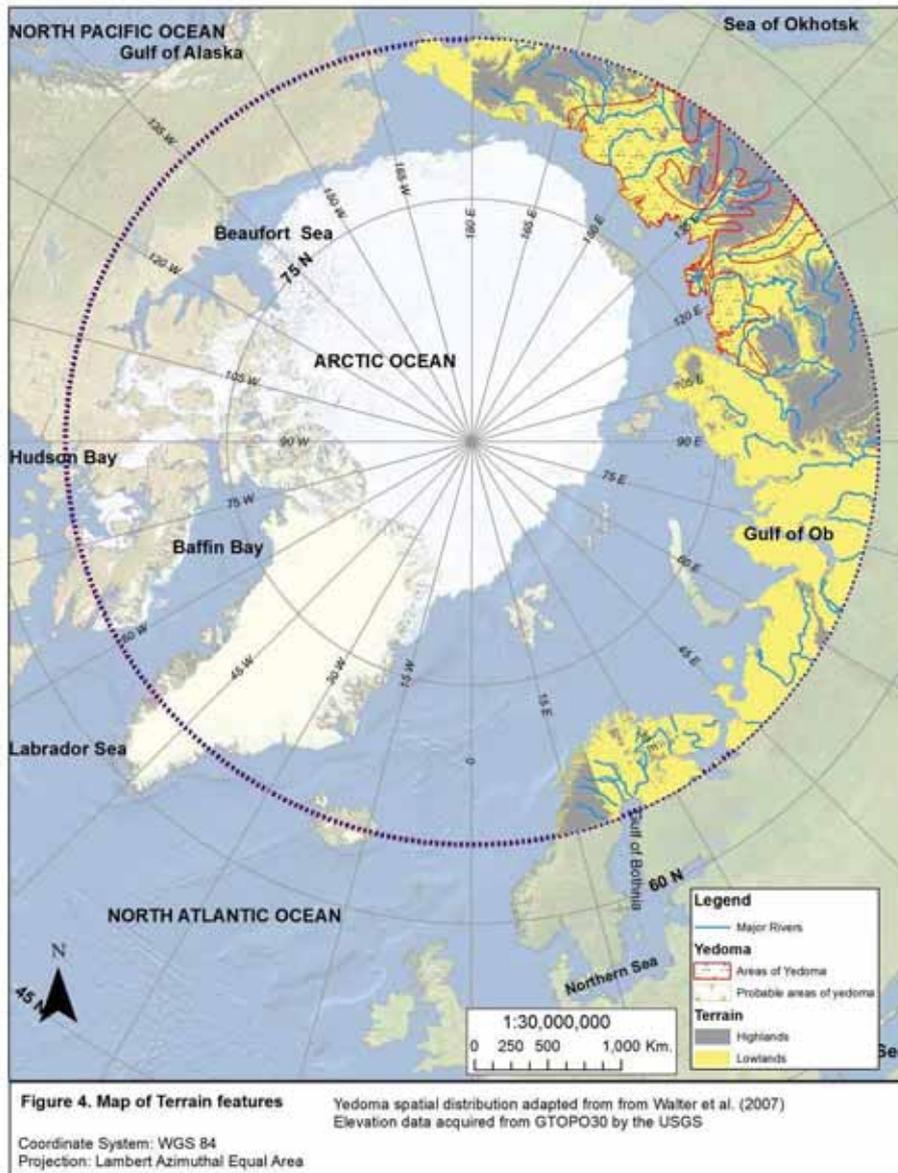


Figure 4: Map of Terrain Features

2.2.2 Secondary Sources

Finally, secondary data was used of existent field samples for methane gas fluxes in arctic lakes. For terms of this study two types of emissions were considered: ebullition and diffusion. It was not considered emissions by emergent plants, since its occurrence has not been found to be significant in total arctic lakes fluxes (Kling et al., 1992, Zimov et al., 1997). Therefore, ebullition data is obtained by the measurements done by Walter et al. (2006) in Siberian lakes whereas diffusion data is obtained by the gas fluxes measured by Zimov et al. (1997) in Alaska. An illustration of the final measured values used in this study is presented in the following table.

Table 2: CH₄ fluxes values adopted in this study.

Type of Lake	Flux by Ebullition (Walter et al., 2006)	Flux by Diffusion (Kling, et al., 1992)
Thermokarst (90% of permafrost lakes)	34.5±9.5 g CH ₄ m ⁻² y ⁻¹	1.1±0.2 g CH ₄ m ⁻² y ⁻¹
Non thermokarst (10% permafrost lakes)	17.9±12.1 g CH ₄ m ⁻² y ⁻¹	
Non thermokarst	17.9±23.2 g CH ₄ m ⁻² y ⁻¹	

Ebullition measured by bubble traps placed over water or in ice in North Siberia from April 2003 through May 2004. This value accounts for 3 ebullition types: hotspot, background and point-source. The value for thermokarst lakes embraces half of the region's lakes with modest thermokarst and the other half representing lakes with intense thermokarst erosion. Molecular diffusion measured along 25 lakes in Alaska during the open water season of 1990. Definition of type of lake follows the approach applied by Walter et al. (2007b).

2.3 Methods

2.3.1 Developing the New Arctic Lakes Geodatabase

The first objective of this research was to develop a new arctic lake database. This was done by identifying and extracting lakes from the available Landsat 5 +ETM imagery. As it was presented in sections above, the most efficient way to detect water bodies from remote sensing imagery is by a density slicing method and if using Landsat images by merely using its band 4. Therefore, for this study it was selected and density slicing for both, band 4 of the Landsat TM scenes.

2.3.1.1 Density Slicing

This method relies on the concept that by defining a threshold value for a single spectral band, discrimination between water and non-water pixel can be effectively performed (Roach et al., 2012). In order to define a suitable threshold from the imagery dataset, a stratified random sampling technique was applied. First, a sampling population was determined at 30 images so that it will approximate a statistical normal distribution (Robinson, 2009). Secondly, the sampling population was proportionately distributed along the vegetation units according to their surface area. Consequently, with the number of images allocated per land cover unit, the images to be sampled were randomly selected among the total dataset. Then, every image selected was split into four quadrants and one was randomly selected for sampling. Finally, the process was followed by recording the maximum DN value that every lake has i.e. defining when a pixel was no longer *water* and was therefore *land*. Thus, a unique threshold band value was obtained by averaging the whole set of DN values measured per sample.

Additionally, it is important to remark that it was determined whether a single threshold value could be applied for the entire area. With the DN lakes' boundaries recorded, the statistics of each group of records per vegetation unit area were calculated. This was followed by an ANOVA analysis of the vegetation groups. The hypothesis to tests was whether there was a significant difference between the 'means' among the different groups. If there was a statistical difference, each vegetation unit would use its own mean value; otherwise a unique value would apply for the entire region.

2.3.1.2 Vectorisation

With a threshold value defined, a simple algorithm was run in order to extract the features that are within the defined threshold. The process was run for each image individually; a mosaic was not created in order to save computing resources. An automated model was built in order to compute the defined operation. Furthermore, it was defined the minimum mapping area. The minimum pixel size retrieved by the sensor is 30 m which makes areas of 90m² detectable by the algorithm. However, due to locational issues and in order to maintain a representative scale it was not considered features with an area less than 4 pixels i.e. 3,600 m². This value is thereby defined as the minimum mapping area for this study.

The process is then followed by the vectorisation of such features. Such vectors were merged together following a basic GIS procedure. Such data were subject to topological analysis following a non-overlapping. Furthermore, to complete the cleaning, a manual cleaning of water pixel in the shoreline was carried out. The usage of a 1:1000,000 scale base map in contrast with 30 pixel size imagery evidently caused a mismatch between the datasets. Finally, clouds and shadows were manually deleted from scenes that presented such noise.

Since the mentioned procedure is based on a threshold based on the extraction of water pixels, it detects lakes as well as riverine features. Therefore, it was necessary to define a function to mask out from the dataset such non-necessary water bodies. Thus, the compactness property of a feature was used. This geometric function defines the roundness of a certain feature based on a relation between its perimeter and its area. This technique has been applied for extracting both manmade objects and natural features from satellite imagery (McKeown and Denlinger, 1984). The compactness value for a feature is calculated via the following expression:

$$\text{Compactness} = \frac{4 \pi \text{Area}}{\text{Perimeter}^2}$$

Equation 1: Compactness property of a feature

Hence, by defining a suitable compactness range, rivers that present the minimum values can be masked out from the dataset. The resultant polygons therefore are then considered as lakes. It is important to remark that this is the first known study to apply such geometric function for extracting lakes in both the Arctic region and with the volume of images obtained.

Finally, a visual explanation of the whole process described hereby is expressed via the following flowchart.

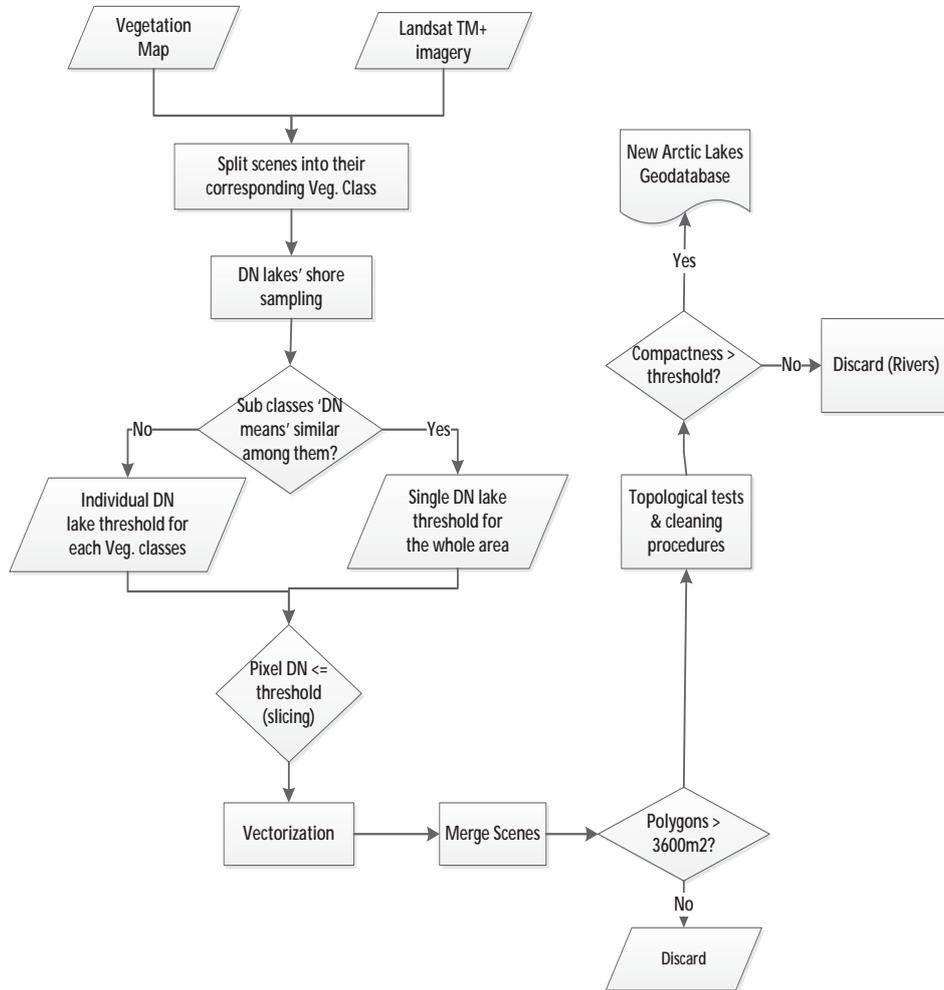


Figure 5: Flowchart of the development of the New Arctic Lakes Geodatabase

2.3.1.3 Accuracy Assessment of the NALGDB.

In order to determine the classification accuracy and therefore the NALGDB reliability, its accuracy was assessed. This was done by comparing the dataset with imagery with higher spatial resolution. Hence, a manual lake classification can be performed using Ikonos satellite imagery (4 meter of pixel size). In order to define control lakes, random images were selected following the stratified random sampling approach mentioned above. Therefore, within each selected

quadrant, the output of the NALGDB with the result of the high resolution manual classification were verified and compared. Thus, an error matrix was then built up. Such error matrix incorporates lake omissions as well as erroneous inclusions and areal differences. It is important to note that due to the temporal variations of the datasets, the areal accuracy for a defined lake was not assessed.

2.3.2 Analysing the arctic lakes spatial distribution

As part of the first objective it was attempted to define the spatial distribution of lakes in the Eurasian Arctic. Then, in order to spatially characterise water bodies this was followed by a lake density and fractional area approach. Such an approach consists of firstly slice the study area into different units of analysis i.e. rectangular cells containing spatial information. This grid scale is calculated following the same geometric size as the GOSAT dataset spatial resolution i.e. 10km². Therefore the spatial resolution match is assured for further analysis.

Furthermore, it is integrated to the grid the Lake Geodatabase. This was accompanied by a GIS straightforward analysis. This was based on intersections of the database with the thematic information of permafrost extent and vegetation units. Firstly, the number of lakes per sub-unit area was calculated. However, since neither the analysis units nor the lakes were equal in size, the mere quantification of lakes within a defined landscape could have led to mislead interpretations of lakes' spatial pattern. Thus, the criterion of lake density and lake area fraction was incorporated.

Such criteria has been already used previously by Smith et al. (2007) to characterise northern lakes. Lake density can be understood as the number of lakes per unit area (eq. 2). Lake areal fraction is a relation represented by percentage the *amount* of space that the lakes are occupying per unit area (eq. 3). The general formulas for lake areal fraction and density are presented here:

$$\text{Lake Density (Ld)} = \frac{\text{Number of Lakes}}{\text{Surface Area}}$$

Equation 2: Lake Density of an Area

$$\text{Lake Areal Fraction (Lf)} = \frac{\text{Total Area of Lakes}}{\text{Surface Area}} \times 100$$

Equation 3: Lake Areal Fraction of an Area

Finally, the values of these fractions were integrated with additional landscape features. Firstly, the distribution of lakes per vegetation and permafrost unit is qualitatively assessed. Secondly, the relation between lake density and fraction is qualitatively interpreted, with terrain characteristics such as the differentiation between lowlands and highlands, and the presence and absence of yedoma is added. Therefore by combining several criteria, a better understanding of lakes' spatial variability can be obtained.

2.3.3 Estimating methane emissions from arctic lakes

In order to define methane emissions from arctic lakes it a 'bottom – up' approach is followed. As explained in sections above, this method basically consists of using local flux measurements and statistical data to upscale emissions to a regional or global basis. Therefore, a similar up scaling procedure is applied to the one proposed originally by Walter et al. (2007b) to define northern gas emissions. The local flux measurements are those measured by Kling et al. (1992) and Walter et al. (2006) as introduced in sections above (table 1). The explanation of the up scaling procedure is explained below.

The up scaling procedure presented by Walter et al. (2007b) basically consists in assigning flux rates to lake areas according to their type of emissions and underlying lake characteristics. Firstly, the values of diffusive flux are defined as constant for the study area regardless of the lakes' underlain terrain characteristics. Conversely, values obtained by the research suggested different ebullition values according to their permafrost characteristic and possible thermokarst affection. To find whether lakes in the area are affected by the thermokarst process, it is assumed that 90% of lakes in the continuous permafrost zone are thermokarst. These types of lakes present a value of CH₄ emission. Moreover, the other 10% of lakes and those lakes that are not in the non – continuous permafrost zone are given a smaller value. Finally, lake areas are allocated according to their thermokarst characteristic and then linked with their correspondent value of gas flux measured, so that the regional fluxes are upscaled.

For this up scaling the resultant New Arctic Lake Geodatabase is used as the background lake dataset. For this process it is assumed that the lakes sampled by the mentioned research are representative for the whole study area. In addition, by using the data described previously, meteorological assumption and the boundary layer model defined by Kling et al. (1992) is taken into account. Additionally, the

values obtained by Walter et al. (2006) assumes a constant atmospheric pressure for the variables.

2.3.4 Defining the role of small lakes in methane atmospheric concentrations

For the third objective of this research, to define the role of small lakes in atmospheric methane concentrations, a 'top-down' approach was followed. It is then attempted to find a relationship between amounts methane measured in the atmosphere and small lakes distribution. This involves an integration of GOSAT- FTS SWIR point measurements of the column amount of CH₄ (Level 2 data products) and lake data of the NALGDB. The process by which these data are processed is explained below:

Initially small lakes are identified. This is done by extracting from the NALGDB those lakes with a total area equal to or less than 10ha. Then, with such lakes identified, the lake area fraction for the areas that contain them is defined. It is inferred then, that as the sizes of lakes are similar, changes in lake area fractions indicate changes in lake densities as well. Lastly, CH₄ column amount values are found for the areas that present the highest small lakes area fraction. Thus, the significance of small lakes to the regional methane budget is explained.

This significance is explained by finding statistical relationships between small lake areal fractions and the concentration of CH₄ along the study area. Such statistical procedure, besides the quantification of the relationship, may define a map of arctic CH₄ emissions based on the presence of lakes in an area. Then, it is tested whether lakes alone can explain the regional amounts of atmospheric CH₄. The steps that follow the definition of such spatial statistical relationship are explained below.

The procedure consists principally of defining areal clusters that integrate both, information about the lakes and information about CH₄ concentrations. Firstly, the 2998 point measurements that make up the 3 years of summer records were integrated into a single geodatabase. Then, there is definition of the possible range of view of a point measurement. This is achieved by drawing a circular buffer zone around each observation point based on their Instantaneous Field of View (IFOV) of the GOSAT-FTS SWIR sensor i.e. 10.5 km. Those buffer polygons – clusters that are overlaid - are merged and treated as a single buffer. Finally, the centroid of such buffer polygons is defined. Around it the final cluster is drawn. This final

cluster contains information about the average CH₄ atmospheric concentration value of the GOSAT single points that overlaid within it and information about the lake distribution. The process can be better visualized by the following figure.

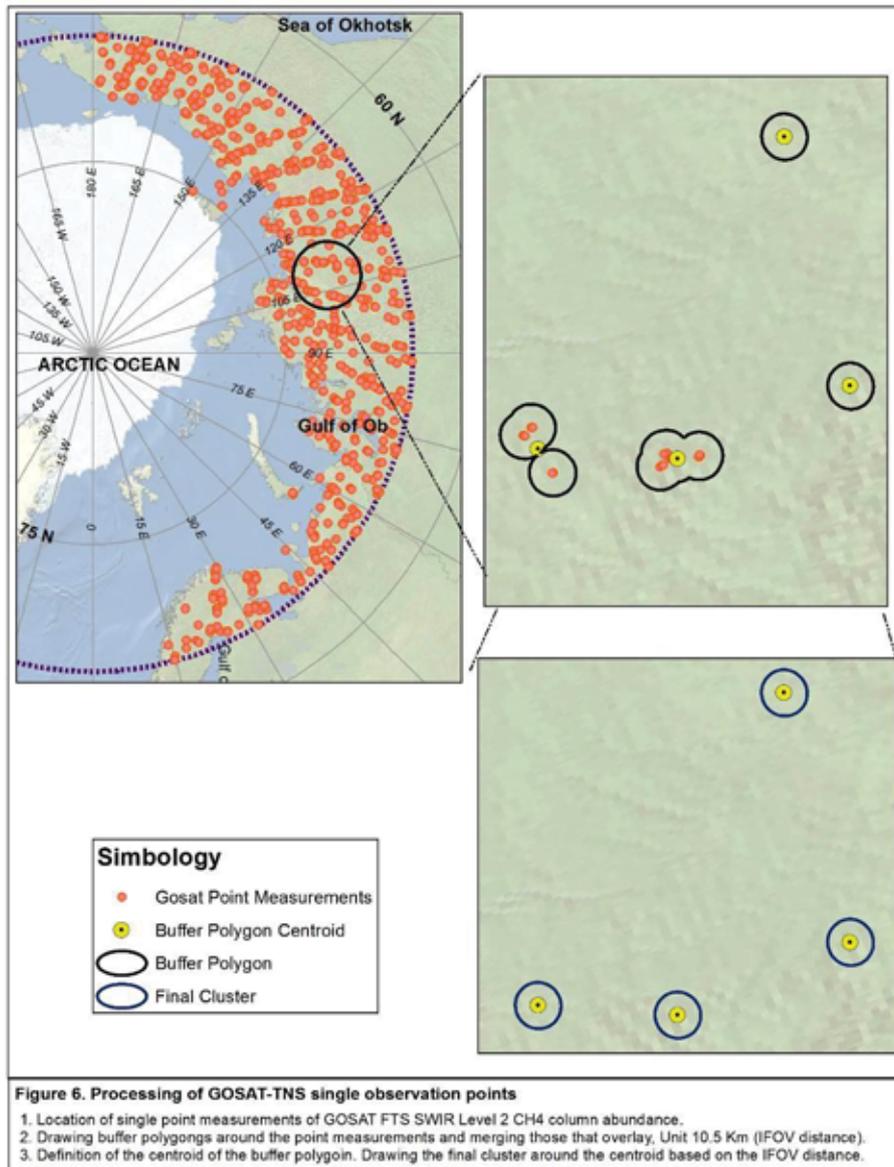


Figure 6: Diagram of the approach for processing GOSAT-TNS single observation points

2.3.4.1. Significance of Units

An analysis is conducted of the units used to express CH₄ amounts. As explained in sections above, the column amount of a gas retrieved is expressed by the total amount of gasses in a vertical column having unit cross-sections. This value, as measured by GOSAT, is expressed as molecules per cm². These values were then converted into mass units and expressed by area-cluster.

In this way, the substance units (molecules) were expressed as units of mass (grams). This is done via a simple chemical unit transformation:

Grams of CH₄

$$= \text{Number of molecules(CH}_4\text{)} \times (1 \text{ mole} / 6.022 \times 10^{23} \text{ molecules}) \times (16 \text{ g} / \text{mole})$$

Equation 4: Transformation of retrieved CH₄ into mass units

Where number of molecules is the value retrieved by the total column amounts of methane, 6.02×10^{23} molecules is the Avogadro's number; and, 16g/mole is the molar mass of methane. In this way, the value of the mass of methane is obtained, expressed as grams per square centimetre. Lastly, in order to obtain a more significant spatial unit, square centimetres are also converted into square kilometres. Then, considering that each cluster has an area of approximately 346.31Km², it is obtained the total amount of gas per cluster. Thus, the final unit is expressed simply as 'grams of CH₄' (which can be read as summer season averaged total column amount of CH₄ per cluster).

Lastly, the significance of the methane values presented in this study was determined. In this way, the values retrieved for the Arctic were compared with values of other global areas which are known to present high and low CH₄ amounts. First selected were mainland areas in China (Provinces of Jiangxi, Zhejiang and Hunan) which are known to present high fluxes of methane due to their rice cultivations (Khalil et al., 1998, Yan et al., 2003). Secondly, selected land areas in the south of Australia were selected. These areas were among those which presented the lowest CH₄ column averaged mixing ratio values as seen by the sensor SCHIAMACHY between 2003 and 2008 (Meirink et al., 2008, Bergamaschi et al., 2009). Then, by following a similar method than the one explained previously, for these areas the CH₄ amounts of 2011 and 2012 were recorded and averaged.

2.3.4.2 Assumptions

It is important to consider the assumptions used in this linear model. Firstly, a constant atmospheric pressure and a constant air mass along the area are assumed. Secondly, daylight variations may significantly affect the amount of gas measured by the sensor. Such diurnal fluctuations can be overcome by conducting intensive campaigns as Juutinen et al. (2004) proposed by studying boreal lakes. Then, it is assumed that 3 years samples obtained by GOSAT are enough to compensate for this possible constraint. In this way, such assumptions allow the study to create an average of CH₄ column quantity for a defined cluster.

3. Results & Discussion

3.1 A new Arctic lakes geodatabase

The first step of this research is the definition of a new arctic lake Geodatabase, the results, analysis and implications of such step are presented here.

3.1.1 Threshold sampling

The sampling population was defined as 30. Based on the vegetation unit surface distribution, 30 images were equitably selected along the area. Within these images a total of 352 lakes were sampled in order to obtain the DN value of their shore. The values were recorded and tabulated as presented in the table below. Furthermore, the results of the ANOVA test show an *F* value of 1.46897 and an *FCritic* of 2.24007, using a 95% of confidence. Since *F* is < *FCritic*, it is proven that there is no difference of the DN threshold means among the vegetation units. Therefore a single value can be used as threshold for the entire study area. The final averaged value, and consequently the threshold value for all the study area, is defined as 24.

Table 3: Summary of DN value statistics of the shore of the lake defined by Vegetation Class

Vegetation Class	% of Land	No. Images Samples	Lakes sampled	DN Mean	Variance
Cool evergreen needleleaf forest	0.27%	1	32	24.3	2.5
Cold evergreen needleleaf forest	13.52%	4	40	24.0	4.3
Cold deciduous forest	32.09%	10	100	24.1	3.0
Low- and high-shrub tundra	31.13%	9	90	24.7	3.9
Erect dwarf-shrub tundra	18.78%	6	60	24.3	1.6
Prostrate dwarf-shrub tundra	4.21%	1	30	24.4	3.3

3.1.2 The Geodatabase Development

The density slicing function was then run via the criterion that those pixels with a value less than or equal to 24 are defined as water. The output was individual rasters per image with values of 0 and 1. The value of 0 represents pixels for non-water areas whereas the latter is the pixel value assigned for water extracted areas. The vectorisation

operation was then executed. With the geometry known (area and perimeter), the compactness value was defined after manually testing several values. In this way, the extraction threshold was defined as 0.1. This meant that features with compactness value less than 0.1 were considered as elongated features. In this way, riverine features were extracted from the dataset.

Moreover, in some parts of the study area some riverine features were not completely masked by the process. Since in these river areas' DN values presented pixels that were within and outside the defined threshold, small water islands were created along the stream. In addition, due to their small extension, they were not deleted by the compactness function either. The result of this was features that were erroneously considered as lakes.

To overcome such constraint, rivers and wetlands were also extracted from the Landsat TM+ collection. Following a similar process as the one used for lakes, the threshold value as 42 was defined. For this purpose, there the features with a low compactness were accepted. In that way, a new mask was created containing polygons of rivers and wetlands. The latter was then used to extract the *false lakes*. It is important to remark that such extracted features cannot build up an Arctic river database. Many wetlands represent a considerable source of noise. Additionally, in some areas streams were not detected by both algorithms.

The following cleaning and merging procedures presented some unexpected constraints. Since an imagery mosaic was not created, images' borders were also included in the final database. Additionally some wetlands were also erroneously considered as lakes. Furthermore, despite the river extraction that was realised, small stream pixels, that were not elongated enough to be discarded by compactness, still remained in the dataset. This led to extra time being spent manually correcting such database's drawback. Finally, in line with this, in terms of computing, the automated processes for 379 image, rasters, and features took less time than the originally expected. A considerable amount of time, nevertheless, was devoted to cleaning scene borders.

3.1.2 The New Arctic Lake Geodatabase

A new geodatabase of lakes for the Eurasian Arctic was created. For the Eurasian Arctic, some 2,000,000 lakes with a surface greater than 3,600 km² were identified (table 4). This significantly differs from the around 30,000 lakes that the actual GLWD comprises. Such

difference was expected from previous field observations and photogrammetry-remote sensing coarse assessments (Smith et al., 2007, Walter et al., 2006, Grosse et al., 2008). However, this difference has not been previously quantified. In addition, these numbers state that the percentage of omission in the GLWD clearly exceeds the 50% previously estimated by Walter et al. (2007b).

However, such omissions are more disturbing when analysed by areal class. In the GLWD there is practically a null inclusion of lakes with areas less than 50 Ha. The GLWD therefore, cannot be used when it is aimed to upscale local-field lakes' measurements to a regional basis. Moreover, this study indicates that about half of arctic lakes have a surface between 1 and 10 Ha. Lakes of less than 1 Ha constitute nearly 35%, and a relatively small proportion (11%) is consists of lakes between 10 – 50 Ha. Surprisingly, less than 1% of the lakes have an areal extension of over 50 Ha, converse to the 96% calculated by the GLWD. Consequently, the numbers obtained by the NALGDB clearly indicate that lakes in this region are typically 'small'.

Table 4: Comparison of lakes size classes between the NALGDB and the GLWD

Size Class	NALGDB		GLWD	
	No. Lakes	%	No. Lakes	%
<=1ha	742,618	34.21%	6	0.02%
1 - 10	1,120,802	51.63%	26	0.09%
10 - 50	242,581	11.17%	993	3.54%
> 50	64,828	2.99%	27,029	96.35%
Total	2,170,829		27,782	

Specific areal comparisons reveal remarkable findings as well. In total, it was calculated that lakes in the Eurasian Arctic make up about 5% of the total land area. That means that there is approximately 225,000 km² of lake water in the region. In contrast, the GLWD have estimated some 105,000 km² of lake water for the same study area. Surprisingly, such difference is not as noteworthy as with the number of lakes. These results then, suggest that the GLWD has omitted small lakes which summed together can represent a similar total lake area than accounting for *big* lakes alone.

Additionally, the numbers obtained from the NALGDB can help to roughly estimate the total number of lakes and the amount of water that they represent for the entire Arctic region. By performing a

simple arithmetic relation the comparison between the two areas can be achieved. It is assumed that lakes in Eurasia cover a similar land extension to lakes in North America. Then, it is ascertained that in Alaska and Canada there are some 1,400,000 lakes which make up a total of 145,000 km² of lake water. Therefore, this coarse approximation firstly indicates that in the entire Arctic there are some 370,000 km² of lake water which possibly comprise some 3500000 lakes.

This estimation can then be compared with previous calculations for the amount of freshwater in the Arctic. For instance, In a study conducted by Bastviken et al. (2011) such value was defined as 300,000 km² of water in the region. This number is based on the estimations of Downing and Duarte (2009). Then, the value calculated in this study exceeds such number for about 7000km². It is important to note however that this study accounts for lakes only, and unlike the aforementioned research, does not include rivers.

3.1.3 Accuracy Assessment of the NALGDB

An accuracy assessment of the Database was conducted by splitting it into land cover units. One Landsat image was randomly selected per unit, and then compared it with Ikonos imagery obtained by Google Earth. Errors of omission were then defined (real lakes not identified in this database) and commission (defined as lakes by the dataset, but actually respond to a different land class) presented in the dataset. The general accuracy of the database is 78%. The individual numbers per land cover unit are presented below:

Table 5: Accuracy Assessment of the NALGDB

	<i>Land cover</i>	<i>Criterion</i>	<i>Lake</i>	<i>No Lake</i>	
NALGDB	Cool evergreen needleleaf forest	Lake	85	16	
	<i>Class Accuracy: 0.78</i>	No Lake	8	-	
	Cold evergreen needleleaf forest	Lake	72	8	
	<i>Class Accuracy: 0.75</i>	No Lake	16	-	
	Cold deciduous forest	Lake	63	16	
	<i>Class Accuracy: 0.74</i>	No Lake	7	-	
	Low- and high-shrub tundra	Lake	80	4	
	<i>Class Accuracy: 0.82</i>	No Lake	13	-	
	Erect dwarf-shrub tundra	Lake	81	12	
	<i>Class Accuracy: 0.82</i>	No Lake	5	-	
	Total averaged accuracy: 0.78				

The generated database was overlaid and compared with high resolution imagery obtained from Google Earth.

In general, the most common mistake was found to be the erroneous identification of lakes. This was particularly found in areas with a long extension of wetlands. For instance, by a visual inspection with the GLWD, the cold deciduous forest unit was observed to present the highest extension of wetlands, which resulted in having the lowest accuracy. Thus, in some areas of the study area the algorithm could not clearly differentiate between lakes and other wetland objects such as palsas, mires and fens. By visually inspecting, the general accuracy of this dataset is likely to increase as the minimum mapping size (3,600 m²) takes a greater value.

Finally, omission mistakes were found particularly in adjacent areas to rivers. The procedure that was used effectively extracted rivers from the database, although in some cases, the algorithm identified lakes along river banks as part of the river, thereby they were extracted from the database.

3.1.4 Spatial Distribution of Arctic Lakes

The obtained values of the New Arctic Lake Geodatabase were then tabulated and its statistics were calculated as is presented in table 5. It is found then that the major number of lakes is located in tundra vegetation units. Over 60% of the lakes are located in this vegetation unit. In more detail, data shows that low and high-shrub tundra is the

vegetation class that presents the greatest number of lakes. Nearly 40% of Eurasian arctic lakes are located along this vegetation unit. This includes the 'erect dwarf shrub tundra' class with 25% of lakes located there. The percentage of lakes then decreases to 16% for cold deciduous forest and cold evergreen needleleaf forest. Finally, the remaining classes i.e. prostrate dwarf-shrub tundra and cool evergreen needleleaf forest, present the lowest values of less than 5%.

Table 6: Number of Lakes and areal statistics for the Eurasian Arctic.

Vegetation Class	Permafrost	Number of lakes	% Lakes	Lake Area (Km ²)	Ld (10km ²)	Lf (%)
Cold deciduous forest	<i>Continuous (90-100%)</i>	233,301	13	39,644	2	3
	<i>Discontinuous (50-90%)</i>	42,017	2	3,133	7	5
	<i>Isolated Patches (0-10%)</i>	5,803	0	330	6	3
	<i>No permafrost</i>	4,317	0	449	4	5
	<i>Sporadic (10-50%)</i>	8,011	0	523	5	3
Cold evergreen needleleaf forest	<i>Continuous (90-100%)</i>	17,339	1	1691	6	6
	<i>Discontinuous (50-90%)</i>	87,823	5	6,755	6	5
	<i>Isolated Patches (0-10%)</i>	34,138	2	4,370	8	10
	<i>No permafrost</i>	120,554	7	16,570	4	5
	<i>Sporadic (10-50%)</i>	39,845	2	7,374	5	9
Cool evergreen needleleaf forest	<i>No permafrost</i>	4,703	0	430	4	3
Erect dwarf-shrub tundra	<i>Continuous (90-100%)</i>	390,166	21	55,087	5	7
	<i>Discontinuous (50-90%)</i>	4,586	0	202	7	3
	<i>Isolated Patches (0-10%)</i>	2,771	0	230	10	8
	<i>No permafrost</i>	41,432	2	3,244	10	8
	<i>Sporadic (10-</i>	14,593	1	1,124	9	7

	50%)					
Low- and high-shrub tundra	<i>continuous (90-100%)</i>	51,1494	28	49,960	4	4
	<i>Discontinuous (50-90%)</i>	105,347	6	7,323	10	7
	<i>Isolated Patches (0-10%)</i>	12,795	1	689	6	3
	<i>No permafrost</i>	21,302	1	2,201	5	5
	<i>sporadic (10-50%)</i>	43,409	2	3,077	7	5
Prostrate dwarf-shrub tundra	<i>Continuous (90-100%)</i>	108,020	6	17,017	7	12

Ld: Lake Density: Number of lakes per 10 km².

Lf: Lake Area fraction: Percentage of lake area per land unit.

Secondly, the above table indicates that there are a major number of lakes in the continuous permafrost area (table 3). It is shown that about 68% of lakes in the region are located in this area. This value slightly differs from the value calculated by Smith et al. (2007) of 73%. Furthermore, the number of lakes sharply decreases for discontinuous and no permafrost areas with values of about 10%. A small number of lakes were found in areas of sporadic and isolated patches of permafrost, with percentages values of 5% and 3% respectively. In this way, the characterisation of a zone as continuous permafrost will suggest a presence of high number of lakes.

The regional and general numbers are reasonably in agreement with those previously defined by Smith et al. (2007). The highest and lowest densities may be associated with the extent of permafrost. However, at the same time, Table 3 shows that taking general conclusions and not taking in consideration specially vegetation sub classes patterns, may lead to overlook specific singularities of the lake distribution.

Individual analysis of the resultant sub-units presents important results as well. Nearly half of Eurasian arctic lakes can be found in the continuous permafrost zones of areas of erect dwarf shrub tundra and low-high-shrub tundra. A significant amount of lakes can be found in continuous permafrost areas of cold deciduous forests (about 13%). Surprisingly, a considerable amount of lakes can also be found in the non-permafrost areas of cold evergreen needleleaf forests. The remaining units present percentages of less than 6%. However, the

mere recon of lakes per landscape unit may lead to erroneous interpretations of Arctic lake spatial distribution. It is important then, to analyse the lake density (figure 7) and areal fraction pattern (figure 8).

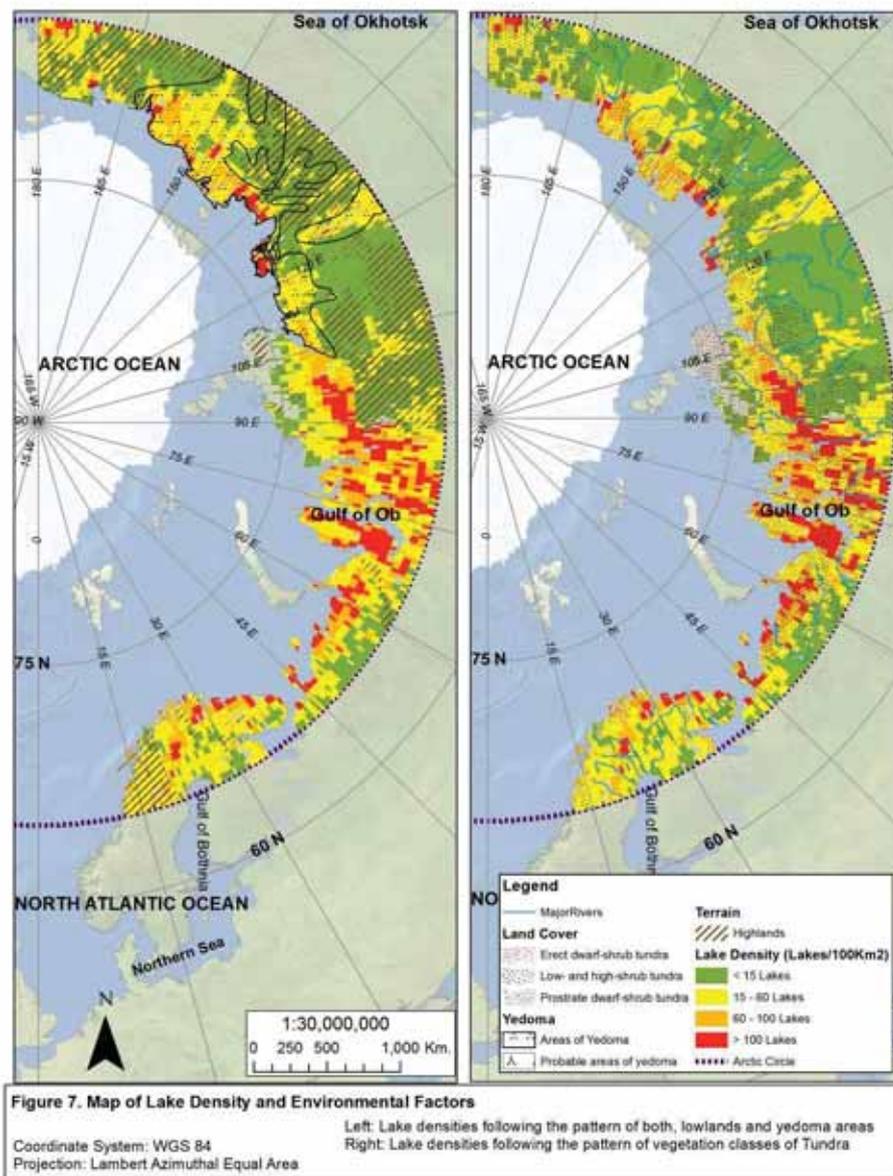


Figure 7: Map of Lake Density

3.1.4.1. Lake Density in the Arctic

The map of Lake Densities shows that the greatest densities were found in zones of continuous permafrost between parallel 60E and 90E around the Gulf of Ob. This asserts what was explained above that there more lakes were found in areas of erect dwarf shrub tundra and low-high-shrub tundra. In addition, areas in Western Siberia and northern Scandinavia also presented patches of relatively high density of lakes compared to the southern part of the study area. The first is presented in the areas of continuous permafrost within patches of prostrate dwarf-shrub tundra. For the latter, the pattern follows the extension of yedoma presence areas. Lastly, it is interesting to note that high lakes densities are found in lowlands areas. In most of the study area the boundary between lowlands and highlands also clearly delineates, the boundaries between lake density classes. In this way, the presence of lakes shows a spatial match with the presence of tundra units, yedoma areas and lowland terrains (Figure 4).

Moreover, areas with the lowest densities are found principally located in the southern Arctic. For the European area, the dominant vegetation class is cold evergreen needleleaf forest. For southern Russia a particular circumstance is presented. The predominant class is 'cold deciduous forest' and despite presenting a rather high amount of lakes (12.5% of the total), the lake density is not as high as in other parts of the area. This, consequently, indicates that lakes are scattered over a large area. It is important to remark that such area presents a continuous permafrost extent.

3.1.4.2 Lake Fraction in the Arctic

Secondly, the Lake Area Fraction map (figure 8) presents an overview of the areal extension of such lakes. Typical landscapes with higher amounts of lake water can be found between the parallel 135E and 165E in Siberia. This means permafrost areas of Northern Siberia, Russia, this area comprises vegetation classes of prostrate dwarf-shrub tundra, erect dwarf-shrub tundra and Low- and high-shrub tundra. Furthermore, a considerable area of medium and big size lakes are found in areas of Western Russia (parallel between the parallel 60E and 105E around the Gulf of Ob) where previously there was defined a high density of lakes. In the same way, a patch of larger lakes are also found and in western Scandinavia. Lastly, small lakes are principally distributed across the southern part of the Eurasian Arctic.

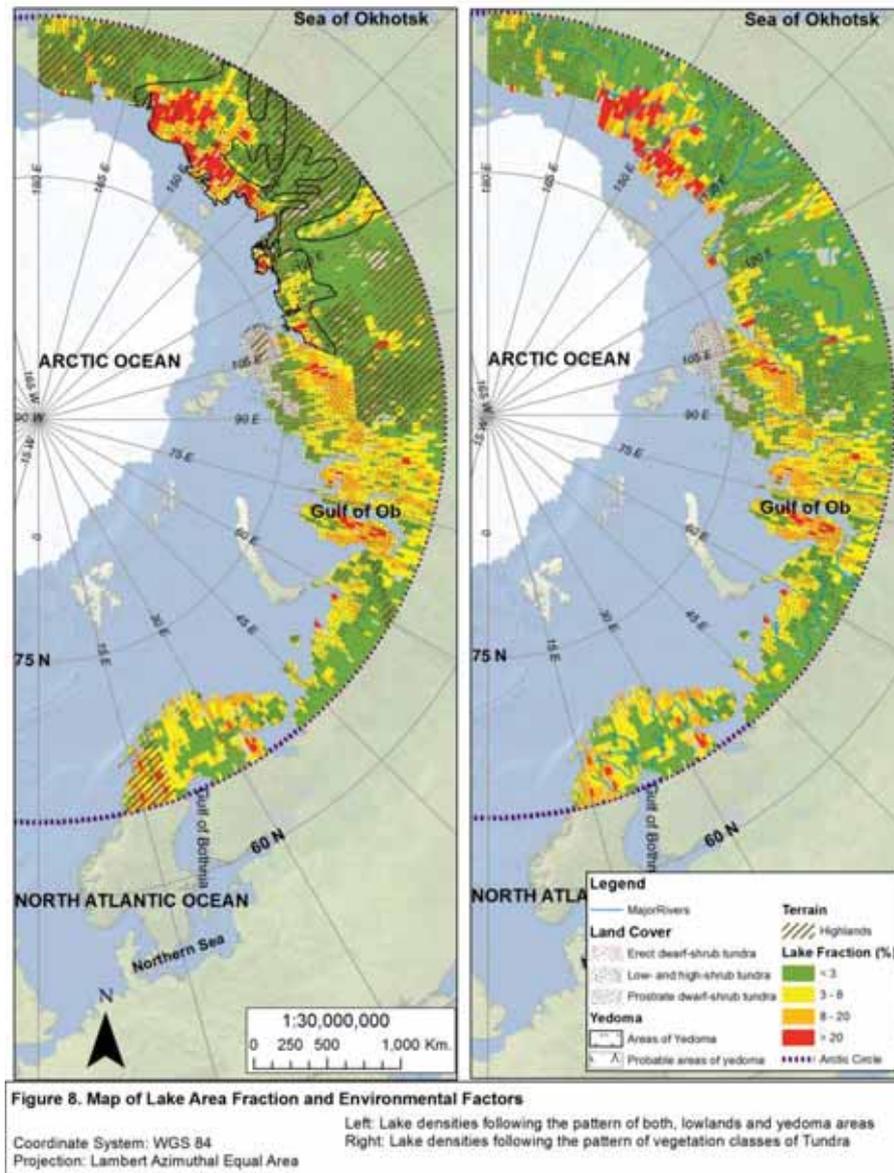


Figure 8: Map of Lake Fraction

However, unlike lake density, a predominant spatial connecting between the lake area function distribution and both, permafrost extension and vegetation classes cannot be found. For instance, for the area of northern Siberia, four vegetation units can be distinguished: Prostrate dwarf-shrub tundra, Erect dwarf-shrub tundra, Low- and high-shrub tundra and Cold deciduous forest. Conversely, large lakes are scarcely found along the same vegetation

units along different areas of the study area. A similar pattern is found for the continuous permafrost extent. Different values of lake area function are found indifferently along the units. Conversely, it can be found that areas with Yedoma and Lowlands are likely to present higher values of Lake Fraction. This result suggests that yedoma characteristics and terrain features may be better variables to explain lakes' size distribution pattern.

3.1.4.3 Additional considerations of the NALGDB

In addition, the lake area cover in the yedoma areas was quantified. The exact area of lakes in this sub-region are covering has remained unclear, previous research has estimated that lakes in this sub-region account from 7% up to 30% of the total land area (Zimov et al., 1997, Walter et al., 2006, Czudek and Demek, 1970) . The NALGDB was processed and tabulated for this particular region and the result showed that there are some 400,000 lakes in this area (about 20% of the total lakes). They make up a lake extension of nearly 70000 km², representing about 8% of the total land surface of yedoma terrains. This late calculation differs with the one estimated by Walter et al. (2006) where it is defined that lakes in the yedoma region of Russia occupy 11% of the area (some 110000km² of lake water).

Furthermore, it was calculated the predominant classes of lakes size in this zone. It was found that nearly 40% of lakes in the yedoma zone of the Eurasian arctic are of a size between 10 – 50 Ha, being the most predominant class. Then, lakes are relatively equally distributed throughout both size classes: less than 10Ha and more than 50 Ha. They both have percentages of lake areas of about 30%, both. In this way, the results suggest that in this in this sub region a dominance of neither small nor big lakes can be found.

3.2 Methane emissions from arctic lakes

The second objective of this study is to determine to what extent the use of a non-detailed lake database may interfere with the correct interpretation of lake greenhouse gases emissions. As described in the above sections, the regional total CH₄ flux was calculated based on a 'bottom-up' approach using existent field data of gas fluxes. Accordingly, the NALGDB was used to calculate the total amount of lake water that each class represented. The results of the upscaled emissions, per land cover and permafrost zone, are presented here:

Table 7: Up scaling of CH₄ gas emissions of the Eurasian Arctic lakes.

Land Cover	Permafrost zone	Lake Area (Km ²)	Lake Ebullition (Tg CH ₄ y ⁻¹)	Lake Difussion (Tg CH ₄ y ⁻¹)	% of Annual Flux
Cold deciduous forest	Continuous	3.53E+04	1.160	0.046	21.36%
	Discontinuous	2.68E+03	0.048	0.003	0.91%
	Sporadic	1.46E+03	0.026	0.002	0.50%
	No Permafrost	3.85E+02	0.007	0.001	0.13%
<i>Sub-total</i>			1.24	0.039	22.90%
Cold evergreen needleleaf forest	Continuous	1.12E+03	0.037	0.001	0.68%
	Discontinuous	5.96E+03	0.107	0.006	2.03%
	Sporadic	5.10E+03	0.091	0.005	1.74%
	Isolated Patches	9.38E+02	0.017	0.001	0.32%
	No Permafrost	1.17E+04	0.209	0.012	3.98%
<i>Sub-total</i>			0.46	0.025	8.74%
Cool evergreen needleleaf forest	No Permafrost	2.20E+02	0.004	0.000	0.07%
<i>Sub-total</i>			0.00	0.000	0.07%
Erect dwarf-shrub tundra	Continuous	4.49E+04	1.475	0.045	27.16%
	Discontinuous	8.04E+02	0.014	0.001	0.27%
	Sporadic	6.06E+02	0.011	0.001	0.21%
	Isolated Patches	5.07E+02	0.009	0.001	0.17%
	No Permafrost	1.25E+03	0.022	0.001	0.42%
<i>Sub-total</i>			1.53	0.048	28.24%
Low- and high-shrub tundra	Continuous	4.68E+04	1.536	0.047	28.28%
	Discontinuous	5.38E+03	0.096	0.005	1.80%
	Sporadic	2.18E+03	0.039	0.002	0.74%
	Isolated Patches	4.33E+02	0.008	0.000	0.14%
	No Permafrost	1.53E+03	0.027	0.002	0.52%
<i>Sub-total</i>			1.71	0.056	31.49%
Prostrate dwarf-shrub tundra	Continuous	1.42E+04	0.465	0.014	8.56%
<i>Sub-total</i>			0.46	0.014	8.56%
Total			5.409	0.183	

The values of table 7 are represented per land cover and permafrost zone. The percentage that each sub-class represents was merely calculated for the amount of fluxes by ebullition as this type of emissions is the more significant.

In this way, by summing the flux by ebullition and by diffusion, the total regional flux adds up to a total of $5.65 \pm 1.24 \text{ Tg. CH}_4 \text{ y}^{-1}$ (the variance stands for the possible error due to the NALGDB accuracy).

3.2.1 Spatial Pattern of CH₄ from Arctic Lakes

The estimation of the methane lakes flux value leads then to the understanding of the spatial pattern of those emissions. The calculated values indicate that nearly 85% of the Eurasian arctic CH₄ lake fluxes occur in zones of continuous permafrost. More specifically, there is a strong contribution to the regional total lake emissions of lakes located in tundra covered areas. Over 60% of arctic methane fluxes are found in this type of land cover. Since these values were calculated based on the total amount of water presented per each subclass, there is in agreement with the lake spatial distribution discussed in the section above. In the same way, since the highest flux rates were assigned to the ebullition release pathway, the highest emissions are presented with this type of emission.

In addition, the spatial distribution of arctic lakes' CH₄ release was mapped. As the principal criterion to estimate lakes emissions was the lake areal extension, figure 9 presented below shows a similar spatial pattern than figure 7 of lake density. Therefore, it was found that the principal *hotspot* of emissions is the north eastern region of Siberia principally underlain by yedoma soils, the continuous permafrost zone and tundra vegetation types and a small portion of continuous deciduous forest. Additional important zones of emissions are found in the transition zone of continuous permafrost around the Gulf of Ob in western Siberia and around the Kheta and Yenisey River basin in central Russia. Surprisingly, these mentioned areas are not underlain by Yedoma, the presence of which has been found to boost lake emissions, but high lake area fractions in such landscapes are likely to define small hotspots of considerable amounts of emissions. Finally, it is noteworthy that the mentioned hotspots found in central Russia are not typically investigated by similar studies which have mainly focused in the yedoma region of Siberia.

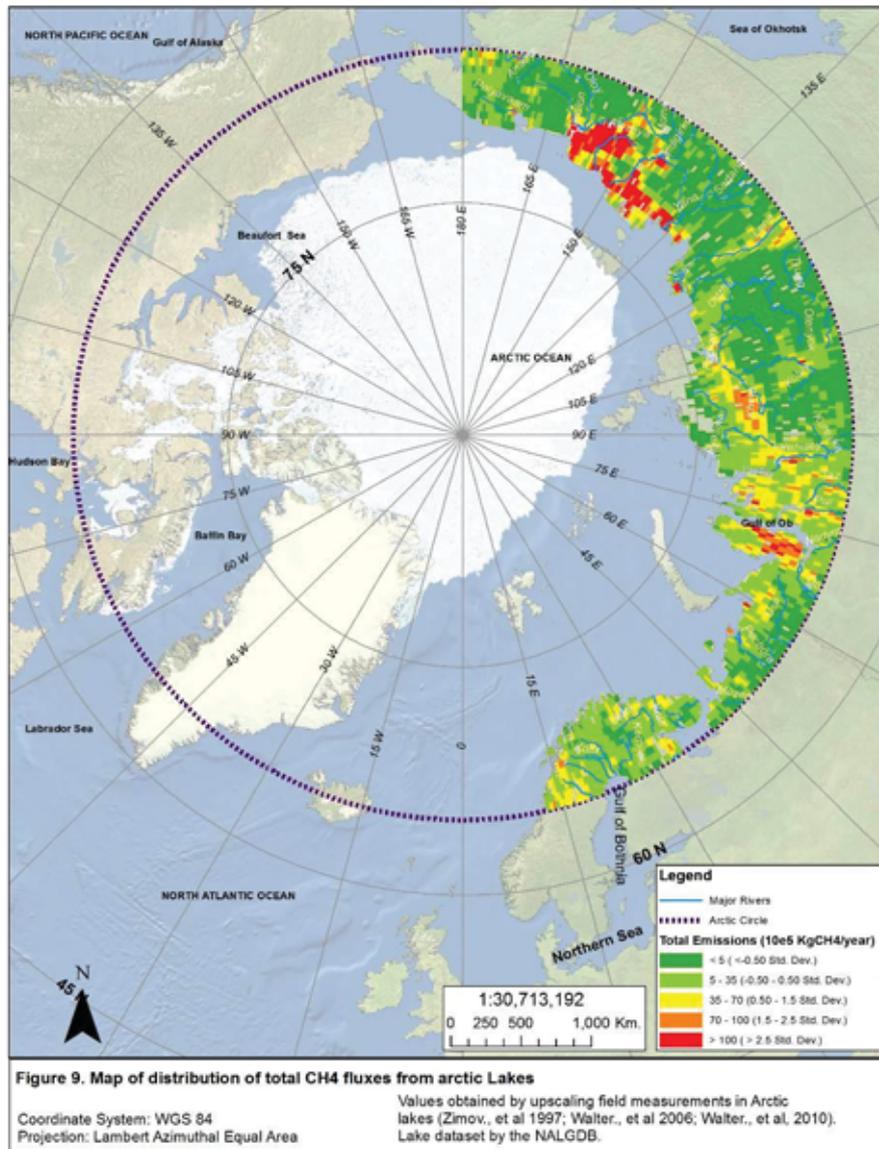


Figure 9: Map of Distribution of total CH₄ fluxes from arctic lakes

3.2.2 Comparison with current methane estimations

The total amount of CH₄ emitted by these types of fluxes in the Eurasian lakes was established at 5.65 Tg.CH₄ y⁻¹. The calculated value in this study accounts for up-scaling local fluxes based in a new detailed lake dataset. As part of the second objective of this research, this value is compared with several estimations of CH₄ in order to understand the role of lakes into the gas budget.

Firstly, If the value presented here is compared with the total global estimations of CH₄ emissions, 600 Tg. CH₄ y⁻¹ (IPCC, 2007) ; the CH₄ release from lakes in the Eurasian arctic lakes account for about 1% of global lakes emissions. In comparison with regional values, previous research estimated that the total emission from northern boreal areas ranges from 20 to 40 Tg. CH₄ yr⁻¹ (Worthy et al., 2000, Mikaloff Fletcher et al., 2004, Zhu et al., 2013). The inclusion of this estimation from boreal lakes will therefore increase these projections to about a 25%. However, the comparison of the calculated value of this study with existent regional and global lakes fluxes estimations unveils noteworthy points of discussions.

In line with this, Bastviken et al. (2011) recently estimated the global CH₄ emission from freshwater ecosystems to be 103 Tg CH₄ yr⁻¹. More specifically, the research indicates that for northern areas, i.e. over more than 66 °N, the total emission for open water ecosystems is 6.8 Tg CH₄ yr⁻¹. This value is then slightly higher than the one established in this study at 5.65 Tg. CH₄ y⁻¹. It is important to remark that both studies used a similar total area for the calculation.

As explained in the above sections, Bastviken et al. (2011) calculated a total freshwater area of nearly 300,000 km², this includes rivers, lakes and reservoirs; whereas the NALGDB presents a lake area of some 220,000 km². This similarity in areas may then explain the slight difference of the results. Yet, the NALGDB accounts for Eurasian arctic lakes alone. Thus, if the lake water area of North America was included, which has been roughly estimated at 150,000km², the total arctic lake water and therefore the methane emissions is likely to significantly increase.

Moreover, the study developed by Walter et al. (2007b) aimed to estimate the current CH₄ ebullition emissions from all northern lakes, presents some arguable results. All lakes north of 45° N were considered as northern lakes. By using similar rates and a similar method as those used in this study, the total regional flux was estimated at 24.2±10.5 Tg CH₄ y⁻¹, and therefore suggests that

northern lake emissions account for nearly 6% of total global emissions. This value clearly differs from the one presented in this study since the study areas are different. However, the most important comparison that arises between both studies is the use of different lake databases.

The study by Walter et al. (2007b) used the GLWD as source data for lake extensions. Identifying the previously discussed lack of detail of this database, the authors applied a correction factor for the total lake area. By comparing the GLWD with ASTER imagery, this factor was established as 50%. The results of the NALGDB presented in the sections above, clearly indicate that the area of lakes omitted in the GLWD exceeds 50%. In this way, with a detailed new database and application of a correction factor based on the NALGDB, the regional and global estimations of CH₄ fluxes from lakes are likely to soar dramatically.

On an arctic regional scale, the research conducted by Walter et al. (2006) estimated the total fluxes in the yedoma region of North Siberia. The estimate value for the lake area in the region was defined as 11%, yielding a regional flux of 3.8 Tg CH₄ y⁻¹. However, as was reviewed in the above sections, the area of lakes in the Eurasian arctic account for an 8% of the total land surface, and a surface of lake water of nearly 70,000 Km². Therefore applying the same equations used in both studies, the annual yield of CH₄ emissions from Eurasian arctic lakes is 2.41 Tg CH₄ y⁻¹, which is approximately half of the value that was previously found.

Finally, these presented results should however be considered as conservative for some specific reasons. For the purposes of this study, flux rates were defined as constant for all the regions regardless of the underlying processes of each specific sub-region or the characteristics of the lake itself. For instance, fluxes in thermokarst lakes are likely to be higher if there are considered specific point – source ebullition values observed by Walter et al. (2007b) in Alaskan lakes. A rate of 195 ± 8 CH₄ m⁻² y⁻¹ will significantly increase the values presented in this study. Additionally, the definition of the regional total flux in this study is not making a differentiation of lakes per size class. This could also increase the value presented here as the study area have shown to be dominated by small lakes (<10 Ha) and these type of lakes, as previously discussed, are known to emit more fluxes per unit area.

Furthermore, the flux rates used in this study are likely to differ if additional criteria are considered. More specifically, recent research

conducted by Walter et al. (2010) improved their past observations by refining long-term measurements on a large number of seeps for each type of ebullition: Kotenok (Type A), Koshka (type B), 'Kotara' (type C) and Hs (Hotspot). Such more accurate observations were 36% and 31% lower than A and B types than those measured in Walter et al. (2006), and 18% and 47% higher in the C and Hs estimations. Finally, as stated in previous sections, in the total emissions the methane release pathway related to vegetation is not quantified.

3.2.3 Implications of a non-detailed dataset for up scaling field measured fluxes

A typical way to overcome the lack of a detailed dataset is the application of a correction factor, as it was discussed in the section above. The 50% of areal omission in the GLWD calculated by Walter et al. (2006) is likely to be the most accurate estimation if the value is compared with the numbers calculated by the NALGDB. However the accurate understanding of methane fluxes patterns may not be represented by the application of a correction factor to previous areal estimations.

The possible regional misestimations of the methane flux when applying a corrected factor was tested. This was done by calculating and mapping the regional emission for the same study area applying a 50% of areal correction for the GLWD. A similar up scaling criterion, as the used previously, was then applied. Therefore, the regional methane value was estimated at $4.14 \text{ Tg CH}_4 \text{ y}^{-1}$. The value is less than the previously estimated by the NALGDB at $5.65 \text{ Tg CH}_4 \text{ y}^{-1}$. This result suggests that the usage of a non-detailed dataset is under estimating the regional methane emissions from arctic lakes.

However, besides the quantification of the flux the understanding of the spatial flux pattern may be constrained too. The comparison between the regional flux calculated by using a detailed dataset (NALGDB) and an aerial corrected non-detailed dataset (GLWD) is shown in figure 10. It is then observed, that both datasets have effectively identified the yedoma region in Siberia as a hotspot. However, the corrected GLWD fails to identify the hotspots of emissions in central Russia. This exercise therefore shows in what extent the usage of a non-detailed dataset, even when applying a correction factor, affects the understanding of the methane flux as a spatial phenomenon.

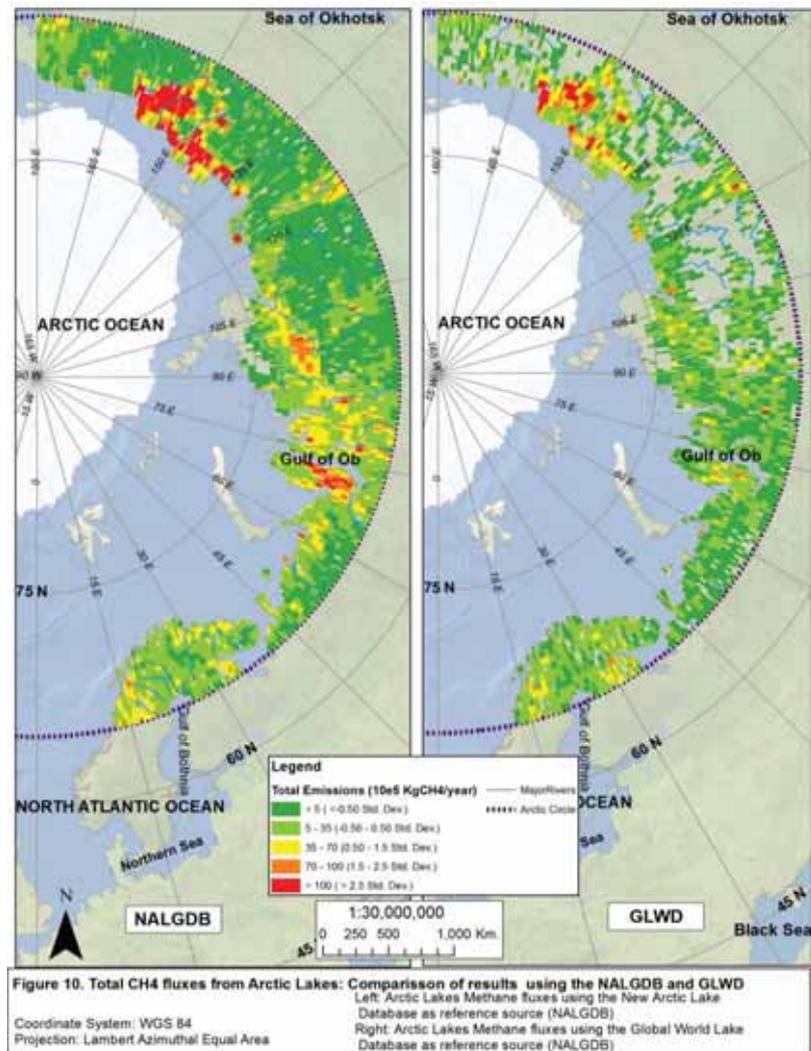


Figure 10: Total CH₄ fluxes from Arctic Lakes: Comparison of results using the NALGDB and GLWD

3.3 Role of small lakes into CH₄ atmospheric concentrations

As the final part of this research, it was also aimed to define the role of small lakes into the atmospheric concentrations of methane. The process was carried out following a 'top down-approach'. The relationship between concentrations of atmospheric CH₄ and distribution of small lakes was investigated.

In this way, final 2998 point measurements from GOSAT were obtained. They were computed and integrated into 490 final clusters. Each cluster integrates both means of total column amounts of methane and small lake areal fractions derived from the NALGDB. As defined in previous sections, table 4, the majority of lakes in the region (about 85% of them) have a surface of less than 10ha. Therefore, to test the role of arctic lakes in atmospheric methane concentration, this is the specific class to study in detail.

From the 490 clusters, lakes with an area over 10 ha were excluded. These lakes represented nearly 20% of the total lakes included in the cluster. In addition, those clusters that had no lakes were excluded from the final analysis. The final number of clusters was reduced to 440. In this way each cluster represents individual sample units. The clusters are distributed along the land cover units: 140 units in the cold deciduous class (32%), 58 in the cold evergreen needleleaf forest (13%), 98 in the erect dwarf-shrub tundra (22%), 121 in the low- and high-shrub tundra (28%) and 23 in the prostrate dwarf-shrub tundra (5%). Besides in the prostrate dwarf-shrub tundra all the units have over 30 samples. Furthermore, these percentages of distribution of the sample clusters are in agreement with the percentages of areal distribution of the vegetation classes along the study area. Thus, it is ensured that there are enough samples in the study area and per vegetation unit which allow the usage of this dataset (which makes the dataset suitable for this study).

3.3.1 Spatial relationship between small lakes and CH₄ atmospheric amounts

With the defined final clusters, it was firstly tested whether an increase in small lakes increases as well as the amounts of CH₄ in the atmosphere. Firstly, it was explored as to whether an increase in the number of small lakes represented as well an increase in the amount of water (figure 10).

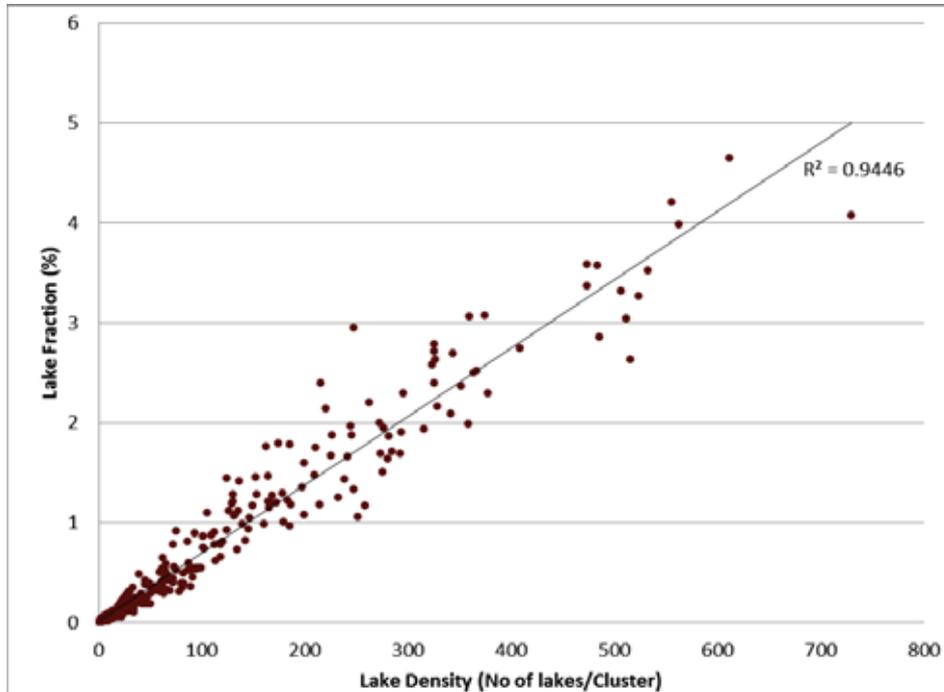


Figure 11: Correlation between lake density and lake fraction for small lakes.

The strong R2 found value suggests that as small lakes increases in number, the cluster that contains them, increases in amount of water at the same time.

As the linear relationship between these two variables is very strong ($R^2 = 0.94$), each cluster is then represented by the lake fraction. In addition, as the clusters have the same area, such lake fraction could be then merely expressed as the lake area or amount of lake water (expressed in Km^2). Secondly, it was tested whether the increase in the total area of water for a cluster may in turn increase the amounts of CH_4 . In this way, the values were plotted together to find a possible correlation, as shown in figure 11.

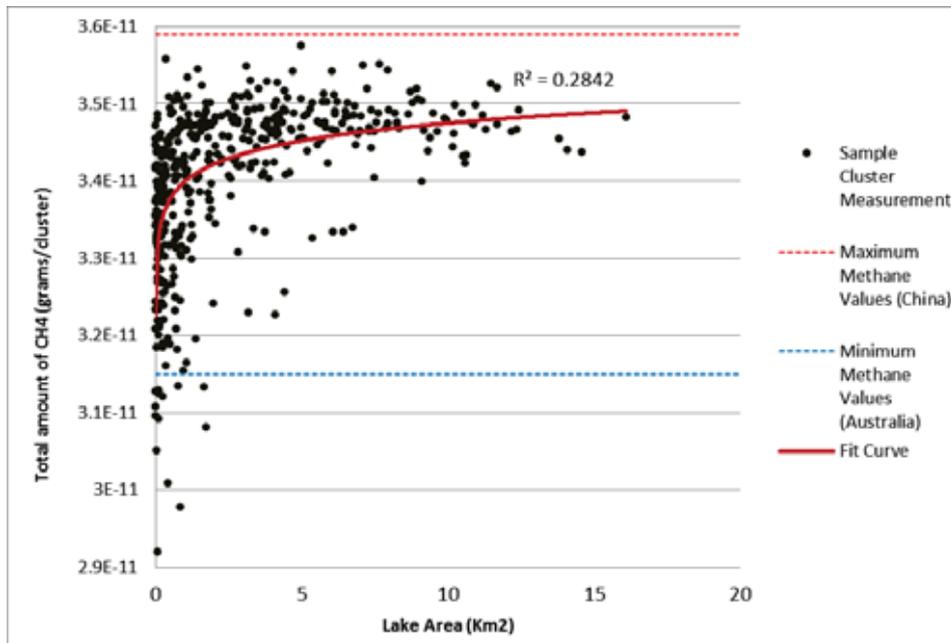


Figure 12: Regression for the relationship between amount of water of small lakes and CH₄ mean total column abundance.

The curve that better explains the relationship is an exponential function. However, the r-squared value shows a weak relationship between the two variables. The values are presented per sample cluster (346km²)

From graph 11 shown above it can be seen that methane concentrations cannot be defined for a certain cluster based merely on their amount of lake water. Therefore, unfortunately for the original aims of this research, this weak correlation represents that a relationship between amounts of methane in the atmosphere and the spatial distribution of small lakes cannot be found. However, certain relations can be extrapolated from it.

In addition, the graph presents a high variability in the range of methane values. It is observed that in the Arctic there are clusters that present low CH₄ amount values, even lower than the reference values obtained for the south of Australia. Conversely, it is found at the same time that there are clusters which present CH₄ atmospheric concentrations, similar to the averaged value for provinces in China which present high levels of emissions. Moreover, the general trend found in the graphic shows that CH₄ total column amounts are unsteady in clusters with low lake area fraction values. Contrarily, when the lake area fraction increases, the CH₄ levels tend to stabilise. These particular results worth a further discussion which is presented below.

As was discussed in previous sections, emissions from wetlands represent a significant source of atmospheric CH₄. In order to avoid this possible source of *noise* (as this study focuses on lakes) a differentiation was made between clusters that intersect with a wetland area and those that do not. The process is done based on the GLWD L3 dataset which includes a categorisation of wetlands. It is then assured that the analysis is done just for clusters which contain lakes and no other possible water features.

276 lakes were identified that do not intersect with a wetland area from those 164 clusters which do. The resultant figure is presented below:

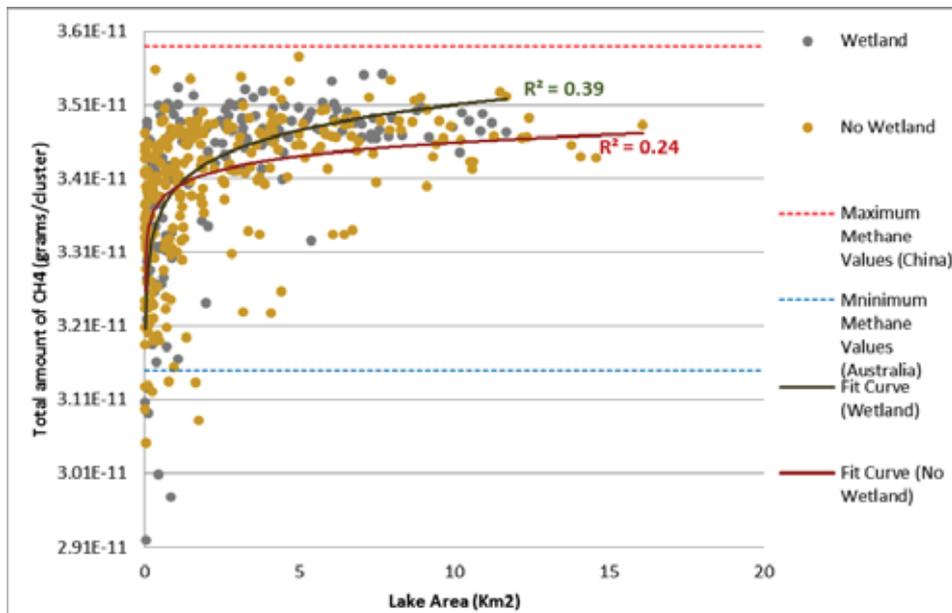


Figure 13: Regression for the relationship between total lake area and CH₄ mean total column amounts for cluster samples that intersect with wetlands and cluster samples do not intersect.

In both cases, an exponential trend is repeated, although the the r-squared value still shows a weak relationship between the two variables. The values are presented per sample cluster (346km²)

Figure 12 presents a similar exponential trend to figure 11 where the samples for the whole study area are represented. Both correlation values are weak although they present interesting differences. By just analysing those clusters which include only lakes i.e. no wetlands, the R² value remains not significant and similar as the value calculated for the entire region. Surprisingly, the R² value for the relationship

between clusters which include both wetlands and lakes, and CH_4 amounts increased considerably (by nearly 15 points). This observation then suggests that wetlands are more likely to determine the total abundance of CH_4 in the atmosphere than lakes do. However the value still remains low and insufficient to explain a strong relationship between increase of lake water and abundances of CH_4 .

3.3.2 The role of small lakes

Despite the weak spatial relationship found between small lake area and CH_4 atmospheric concentrations, it is yet necessary to analyze in detail CH_4 values in areas with small lakes. It is followed two approaches to do so. The first one is the analysis of the outstanding zone of points which are not part of the main body of point as it is observed in figure 11 and figure 12. Such zones are characterized by their low lake areas or their high lake areas. The second approach is the comparison of CH_4 values between CH_4 between clusters with mainly small lakes and clusters with mainly big lakes.

3.3.2.1 Analysis of outstanding zones of points

For this analysis it was considered merely which just include lakes and do not have wetlands i.e. noise. There are then plotted the 276 clusters that do not intersect with a wetland area. There were identified two categories of analysis: clusters that have both, low abundance of CH_4 and clusters with great lake area of CH_4 (figure 13). Each cluster then contains 10% of the points i.e. 25 points. In order to increase the fluency of the discussion henceforth the first category will be referred as 'Bottom 25' whereas the later as 'Top 25'. It is important to remark that the category 'Top 25' is likely to represent areas with more small lakes, and therefore it is of special interest for the aims of this research.

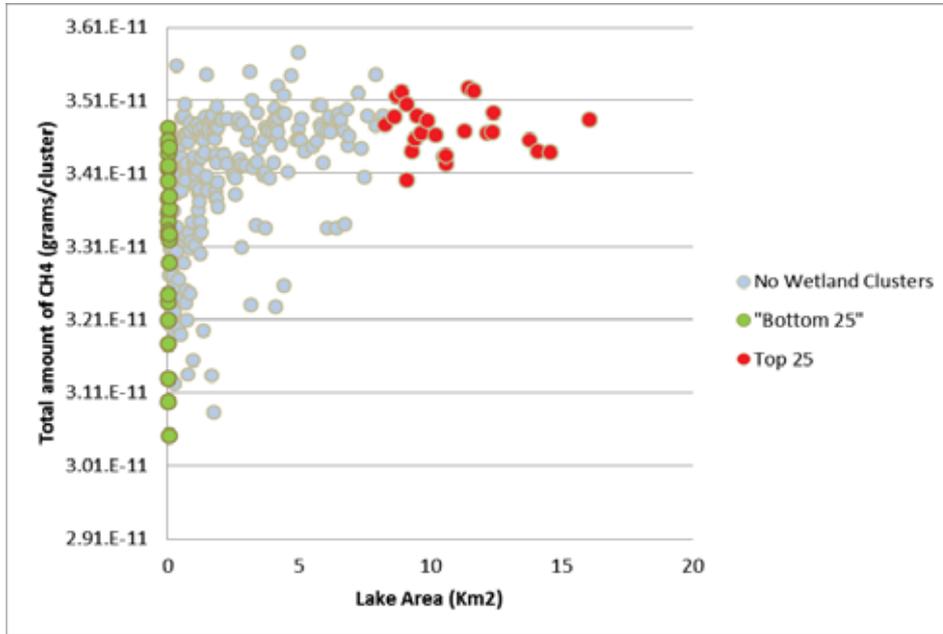


Figure 14: Selected clusters for analysing the role of lakes in the abundance of CH₄ in the atmosphere.

'Bottom 25' refers to the category of clusters with small lake area. Conversely, 'top 25' refers to the category of clusters with both large lake area and high CH₄ abundances.

In order to find whether CH₄ concentrations in these sample clusters can be considered significant, the CH₄ concentration values were classified. This was done by defining interval of classes based on the difference in standard deviations from the mean. A table summarising the classification is shown below:

Table 8: Classification of CH₄ mean total column abundances

CH ₄ mean total column abundance (grams x 10e-11)	Classes
3.051- 3.175	I
3.176 - 3.255	II
3.255 - 3.355	III
3.356 - 3.462	IV
3.462 - 3.5752	V

5 Classes were identified across the study area, the rank of the class increases as the CH₄ values increases. The difference between each class is one standard deviation from the mean.

Initially, the distribution pattern of 'Bottom 25' category can be observed. Figure 13 reveals that those clusters with the lowest CH₄ column abundances values are located in areas with scarce lakes densities and lake fractions. Typically low values of CH₄ amounts can be associated with areas with infrequent presence of lake water

Moreover, from table 7 and figure 13 it is observed that Top 25 clusters present mainly high levels of emissions. For instance, it is found that about 75% of these units are located in the class V. The remaining 25% was found in class IV. This suggests that areas with a strong presence of lakes are more likely to present high levels of CH₄ total column abundances. These observations can be understood that as the clusters increase on their amount of lake water, the total column abundances CH₄ tends to stabilise at higher values. In fact, this result suggests that there is a significant impact of gas emissions from small lakes in the total CH₄ column abundance of CH₄.

3.3.2.2 CH₄ Significance of small lakes in relation with big lakes

As an impact of small in the total CH₄ atmospheric column abundances was identified, it is then tested whether such values are significant if they are compared with areas with big lakes. Clusters of lakes were classified by their area fraction into 5 classes. The resultant classes of small lakes (< 10 ha) were then classes of big lakes (> 10 ha) as it is shown in table 9 and figure 15.

Table 9: Atmospheric CH₄ amounts per classes of lakes.

Class	Amount of Lake Water (Km ²)	Amount of CH ₄ (grams) small lakes		Amount of CH ₄ (grams) big lakes	
	Range (Lake fraction)	Mean	Std	Mean	Std
Class A	<1.30	3.35E-11	9.89E-13	3.38E-11	1.06E-12
Class B	1.30 - 4.55	3.43E-11	8.41E-13	3.44E-11	6.83E-13
Class C	4.55 - 7.79	3.46E-11	5.23E-13	3.45E-11	6.18E-13
Class D	7.79 - 11.03	3.47E-11	3.60E-13	3.46E-11	5.51E-13
Class E	> 11.03	3.48E-11	3.07E-13	3.46E-11	6.29E-13

The difference between each class is one standard deviation from the mean of lake water.

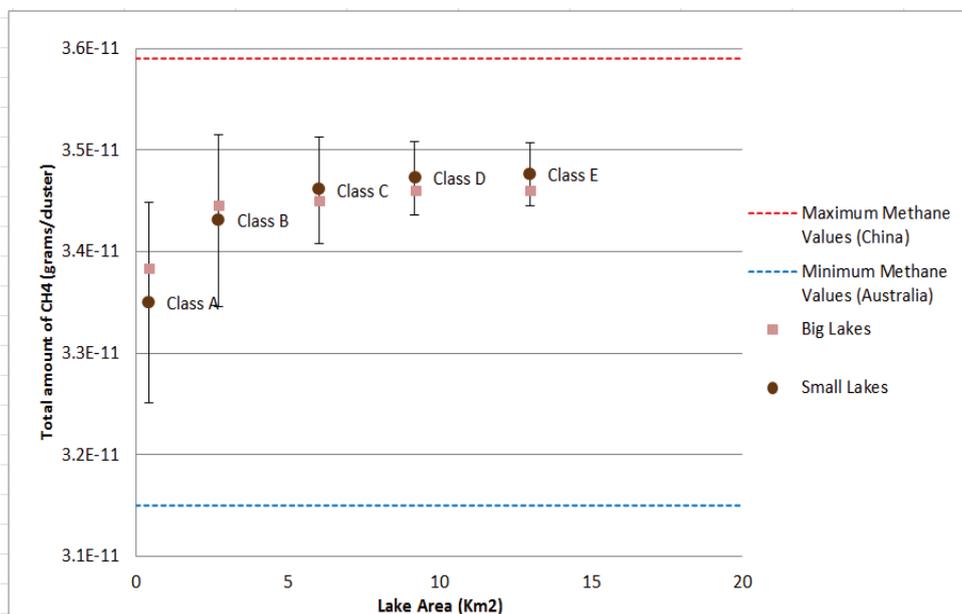


Figure 15: Scatter plot and error bars of the distribution of lakes classes of clusters and total amounts of CH₄

It is observed that atmospheric CH₄ values in classes with more small lakes are greater than those measured for clusters with large lakes. Additionally, it is observed that as the mean lake area increases for small lakes, the variability of the amount of atmospheric CH₄ retrieved tends to decrease. This finding then asserts the pattern described in the previous section. As a cluster tends to present more small lakes, its atmospheric CH₄ values tend to stabilize at high values. Surprisingly, the CH₄ values are greater in cluster of small lakes than clusters of big lakes. Therefore this result suggests that areas with small lakes are emitting more methane than areas with large lakes.

However, despite the proven impact of small lakes in the amounts of atmospheric CH₄, there cannot be defined a relationship between the two variables as it was already shown in figure 12. Such constraint certainly undermines the possibility of a better understanding of the emission pattern along areas with different fractions and densities of lakes. The weak correlation found between the two variables may be due to the wide range of CH₄ values that is presented in areas with different amount of total lake water. Such great variance, is clearly observed in classes with few amount of lakes (class A and class B). For instance, in figure 15 it is shown how within a similar class of CH₄ column abundances (Class V, see table 8), there can be found both

an area with high amount of small lakes and an area with scarce presence of lakes.

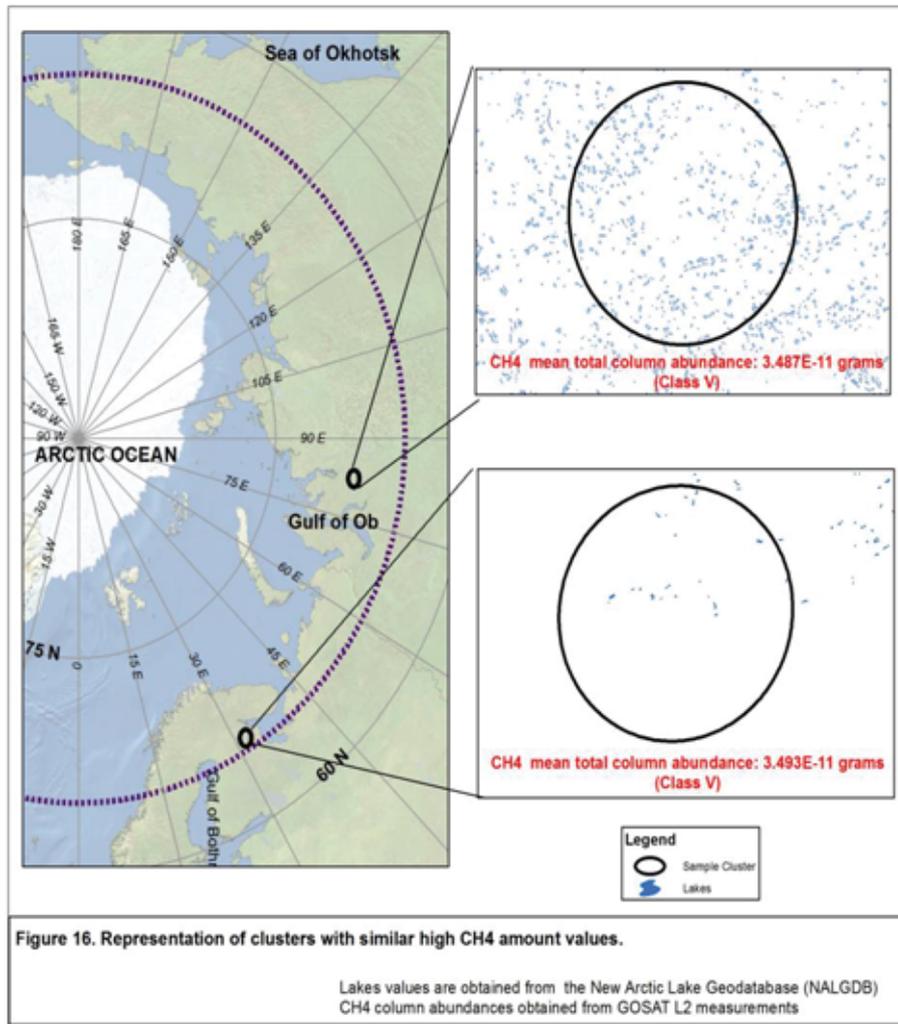


Figure 16: Representation of clusters with similar high CH₄ column amount values.

It is observed that clusters can have a similar value for CH₄ total column abundances (they are both part of the Class V), this is despite the fact they significantly differ in the amount of lake water area that it is contained within them.

The weak relationship between amount of lake water and CH₄ column amounts can possibly be explained by two non-considered factors in this 'top-down' approach. Firstly, this can be explicated by the

contribution of land ecosystems fluxes when analysing CH₄ abundances. Secondly, it can also be explained by the variability by which the gas behaves in the lower atmosphere. The former leads this study to consider the spatial distribution of the categories defined above whereas the latter leads to ponder the interaction between lake fluxes and the atmosphere.

3.3.3 Overview of factors affecting the relationship small lakes and CH₄ column amounts

Role of Terrestrial ecosystems (1st possibility). It was previously discussed that the emissions from other ecosystems may interfere in the understanding of the role of lakes alone to the quantification of CH₄ amounts in the Arctic. The land cover units which are dominant in the classes of small lakes are analyzed as it shown in table 12:

Table 10: Predominant vegetation classes found within each class of lake area.

Class / Vegetation Unit	Cold Deciduous Forest	Cool Evergreen Needleleaf Forest	Erect Dwarf-shrub Tundra	Low- and High-shrub Tundra	Prostrate dwarf-Shrub Tundra
I	66	16	11	34	5
II	23.75	17.5	43.75	8.75	6.25
III	23.5	11.75	32.35	23.5	8.8
IV	-	-	40	50	10
V	-	-	-	50	50

Values expressed as percentages (%). Permafrost extension was not included as it was found that continuous permafrost is the predominant type across the clusters (percentages of over 85%).

From table 12 it is observed that Tundra Vegetation is the dominant type among the classes. Such strong dominance of tundra then may explain the variability of methane amounts when it is related with lake presence. This can potentially be explained by the substantial source of methane emissions that arctic tundra ecosystems represents (Christensen, 1993, IPCC, 2001). In this way, these results suggest that emissions from arctic tundras are likely to affect the explanation of a relationship between lakes distribution and methane atmospheric amounts.

Finally, it is important to consider the drivers that have been found to regulate those land emissions. For instance, Olefeldt et al. (2013) empirically demonstrated the factors that are likely to strongly control those emissions. It was found that the more predominant factors were: water table, soil temperature, plant productivity, and the presence of sedges. These factors then are more likely to be regulating the emissions in areas where lakes are not the dominant features of the landscape.

Role of interactions surface – atmosphere (2nd possibility). Lastly, as defined previously, an additional factor can restrict this quantitative relation between CH₄ total column amounts and lake water. This additional possible driver is that the gas released from a lake may mix in the atmosphere and therefore its trace may be constrained.

This option can be principally explained by both the mechanism by which a released gas interacts with the lower atmosphere, and the characteristics of the TANSO-FTS sensor installed in GOSAT. Related research in the area suggest that the column of CH₄ presents variability caused by: changes in surface pressure, changes in the tropospheric CH₄ volume mixing ratio, and changes in the amount of stratospheric CH₄ due to changes in tropopause altitude (Washenfelder et al., 2006, Washenfelder et al., 2003). It is for this reason that the main key to relate trace gases concentration to underlying processes, that determine such variability, generally relies on the applicability of a model to describe such process.

Such problem has been addressed by Palmer (2008) through relating trace gas concentration measurements to surface processes by a straightforward equation, as illustrated below. Principally, the mass balance (C) of a generic gas is described by the surface production (P) and loss (L), atmospheric transport (T) and additional processes (X). All of these variables then are determining such changes in measurements of the gas in function of time.

$$dC(t) = P(t) - L(t) + T(t) + X(t)$$

Equation 5: Mass balance of a generic trace gas C

Based on this equation, the relationship between surface processes, i.e. gas fluxes from lakes, and atmospheric gas concentrations can be modelled. The equation therefore needs to be complemented by adding the correspondent transport values and additional processes that might define the relationship. Such transfer are likely to be influenced by wind speed, by turbulence at the air-water interface and boundary layer stability (Wanninkhof, 1992, Jähne, 1987) ,

whereas additional process (X), may include by the chemical gradient across the air–water interface (Bade, 2009, Palmer, 2008). In this way, the certain influence of such physical and chemical variables that are not included in this study, are also determining the presence of CH₄ in the column measurements not considered in this study.

Additionally, besides these physical and chemical interactions of a certain gas with the atmosphere, it is important to comment on characteristics of the method used to retrieve CH₄ column abundances. The possible constraints of tracing concentrations of gases in the lower atmosphere has been discussed previously by Palmer (2008). Firstly, there have been associated difficulties caused by the noise of interfering gases that can eventually overlap the target measured gas, stratospheric mixing of target gases, a poor representation in the tropospheric column of the aimed gas, among others.

Secondly, it is important to consider the mechanism by which the GOSAT – TANSO FTS sensor retrieves CH₄ column abundances from water dominant landscapes. The observation of reflected light over the surfaces of water bodies and oceans is difficult since the water absorbs sunlight. The water, however, reflects sunlight in a specular way towards different directions (De Ridder, 2000, Moore, 1980, Hilton, 1984). The instrument thus targets such points when observing large water surfaces. Therefore, the retrieved values analysed in this study may in fact represent, specular light reflected from different water bodies across the region. This may represent a particular limitation when areas with high lake densities and areal fractions are being studied.

4. Conclusions

A New Arctic Lakes Database (NALGDB) for the Eurasian region was successfully developed. With an overall accuracy of 78%, over 2.000.000 lakes were identified. This number clearly differs from the around 30.000 lakes presented in the commonly used Global Lake and Wetland Database (GLWD) for the same area. The principal omissions of the GLWD were found when mapping small lakes (less than 10Ha.).

The numbers calculated by the NALGDB estimate that about 85% of lakes in the Eurasian Arctic have an area of less or equal than 10Ha. This suggests that the lakes in the region are typically small. In total, lakes in the Eurasian Arctic occupy a total of 5% of the total land area (some 225,000 km² of lake water). A more specific regional comparison of the NALGDB defined that lakes comprise about 8% of the total land area of terrains underlain by yedoma. This number, then, clarifies the existent high uncertainty of current researches when accounting for lakes within such type of terrain (ranging from 7% up to 30%).

Furthermore, landscape drivers were identified which are likely to be regulating the presence of lakes. For instance, it was found that the major lake densities (number of lakes per unit area) occur in areas of tundra vegetation, continuous permafrost, overlaid by yedoma and located in lowlands (less than 300 m.a.s.l.). Specifically, this was found in Western Siberia, some highly dense areas in Northern Scandinavia and a considerably large area around the gulf of Ob. A spatial match for lake area was only found with tundra vegetation cover and lowlands.

Moreover, with a detailed lake dataset the study presented a new regional estimation of methane emissions from Eurasian arctic lakes. The regional flux was established at $5.65 \pm 1.24 \text{ Tg. CH}_4 \text{ y}^{-1}$. Integrating this refined value into existing regional projections, the overall regional value for land emissions increases by about 25%. The usage of a detailed dataset then is likely to significantly increase regional estimations of lakes' emissions. However, the usage of a non-detailed dataset and further corrections may also overestimate local estimations of emissions. More specifically, for the Yedoma region, the overall emissions were calculated at $2.41 \text{ Tg. CH}_4 \text{ y}^{-1}$, which is approximately half of the value calculated by previous research ($3.8 \text{ Tg. CH}_4 \text{ y}^{-1}$).

Conclusions

Additionally, hotspots of lake emissions are distributed along areas with high presence of lakes or in areas with terrain characteristics which enhance methane production. The principal emissions hotspots were identified in: the Yedoma region in Siberia, around the Gulf of Ob, Around Kheta and Yenisey River basin in central Russia. The first area has been extensively studied for their known processes related with thermokarst processes which enhance methane emissions. Conversely, when estimating regional fluxes, areas in central Russia, as the mentioned previously, are likely to be neglected. This, despite their reasonably high numbers of lakes, makes these areas important hotspots of emissions. In this way, the usage of a detailed lake dataset suggests the existence of additional significant lakes hotspots of methane emissions.

Finally, this study aimed to define the role of small lakes into the regional methane abundance. It was found that in fact, areas with more small lakes are likely to present high levels of methane amounts the atmosphere. In this way, as small lakes in an area are more frequent, methane values tends to increase (approaching those values observed for rice field areas in China, generally agreed to present high methane emissions). Moreover, the observations suggest that there was a strong interference of emissions from landscape tundra vegetation. In addition, a possible mixing of the gas emitted in the atmosphere is likely to affect the understanding of the phenomenon.

5. Recommendations

This research highlights the importance of having a detailed geographic dataset for lakes and other water body studies. The presented method used to classify lakes in this study is considered straightforward. It effectively coupled remote sensing and GIS techniques with process times and technological resources usage that were less than those that were anticipated. It is then strongly recommended to use a similar procedure not only for the North American Arctic, but for other world regions. However, it is encouraged to test other remote sensing techniques to map areas that were not included in this study due to cloud noise, or to check those areas that presented the lowest accuracy values. The usage of radar imagery coupled with Landsat ETM imagery and DEM is then expected to effectively increase the accuracy of this study.

The development of this New Arctic Lake GeoDatabase is expected to improve the understanding of arctic hydrological and natural processes. This research, at the same time, is expected to promote awareness in the scientific community of the necessity of a detailed lakes and water body dataset. This will then serve as a starting point for further solid regional or global hydrological research. For the arctic region, in particular, this research can contribute to define a first step towards long-term monitoring of arctic hydrology. This takes noteworthy importance since previous projections estimated that between 10% and 20% of the current permafrost area will thaw in the next 100 years (Bowden et al., 2008, Lunardini, 1996). Therefore, effective arctic research must include monitoring the regional water bodies and their associated processes.

In line with this, a suggested further stage is to classify the presented lakes based on their morphometric characteristics. For instance, this detailed dataset can be coupled with the straightforward morphological methodology proposed by Morgenstern et al. (2011) to identify drained thaw lake basins and thaw lakes. This in turn may define to what extent lakes are a significant indirect source of nutrients and sediments to local streams, and thereby enhance geochemical processes along other water bodies. The NALGDB is believed to have the sufficient precision so that a regional morphometric study for arctic lakes can be accomplished.

Moreover, the detailed NALGDB was used as an input for the bottom-up approach to determine the regional methane emissions. This has provided new insights of the role of lakes into the methane budget. It

is suggested to verify by field measurements the significance of the hotspots of emissions in the areas of central Russia. In addition, adding the new values of methane fluxes measured by Walter et al. (2010) will likely to enrich this approach. In line with this, future research should include refined field measurements of methane flux rates from lakes characterized by their size (with special attention on small lakes) and their landscape-vegetation class. The final values presented in this study are expected to significantly increase if these variables are considered as well.

It is also recommended that the 'top-down' approach introduced in this study be extended. The challenge for the scientific community should be to effectively model the behaviour of gas in the atmosphere. Therefore, in order to refine such approach, it should be carried out along with an understanding of the interactions between the gas fluxed from a lake and the atmosphere. In addition, land-vegetation emissions has to be considered and factored into those refined models. Moreover, it is expected that the usage of GOSAT data will open the path to remotely study gas flux processes, overcoming in this way the limited spatial coverage of the current ground networks. In addition, it is strongly recommended to use different techniques to test the results obtained here. For instance, different interpolation techniques can be used to analyse the point measurements of GOSAT. Not least is the fact that GOSAT is one sensor which directly provides CH₄ atmospheric values. However, additional methods can be applied to retrieve the optical water properties of lakes which may lead to a better understanding of lakes' gas processes.

Finally, existing models which quantify regional CH₄ fluxes from the Arctic should also be reviewed. The detailed dataset presented here should be integrated as a parameter to existent models of methane and carbon dynamics. For instance, it is strongly recommended to integrate the NALGDB and the results of gas fluxes obtained here, with the Terrestrial Ecosystem Model (TEM). The study of lakes thaw cycling dynamics should be integrated into the component already introduced by Zhuang et al. (2004) who refined the model by integrating a hydrological component. In addition, results retrieved from GOSAT can collaborate to validate procedures of the model.

References

- ANISIMOV, O. & RENEVA, S. 2006. Permafrost and Changing Climate: The Russian Perspective. *Ambio*, 35, 169-175.
- BADE, D. L. 2009. Gas Exchange at the Air–Water Interface. *In*: EDITOR-IN-CHIEF: GENE, E. L. (ed.) *Encyclopedia of Inland Waters*. Oxford: Academic Press.
- BASHER, R. E. 1982. Units For Column Amounts of Ozone and Other Atmospheric Gases. *Quarterly Journal of the Royal Meteorological Society*, 108, 460-462.
- BASTVIKEN, D., COLE, J., PACE, M. & TRANVIK, L. 2004. Methane emissions from lakes: Dependence of lake characteristics, two regional assessments, and a global estimate. *Global Biogeochem. Cycles*, 18, GB4009.
- BASTVIKEN, D., TRANVIK, L. J., DOWNING, J. A., CRILL, P. M. & ENRICH-PRAST, A. 2011. Freshwater methane emissions offset the continental carbon sink. *Science (New York)*, 331, 50.
- BERGAMASCHI, P., FRANKENBERG, C., MEIRINK, J. F., KROL, M., VILLANI, M. G., HOUWELING, S., DENTENER, F., DLUGOKENCKY, E. J., MILLER, J. B., GATTI, L. V., ENGEL, A. & LEVIN, I. 2009. Inverse modeling of global and regional CH₄ emissions using SCIAMACHY satellite retrievals. *Journal of Geophysical Research: Atmospheres*, 114, D22301.
- BERGAMASCHI, P., KROL, M. C., DENTENER, F., VERMEULEN, A., MEINHARDT, F., GRAUL, R., RAMONET, M., PETERS, W. & DLUGOKENCKY, E. J. 2005. Inverse modelling of national and European CH₄ emissions using the atmospheric zoom model TM5. *Atmospheric Chemistry and Physics*, 5, 2431-2460.
- BOLCH, T., MENOUNOS, B. & WHEATE, R. 2010. Landsat-based inventory of glaciers in western Canada, 1985–2005. *Remote Sensing of Environment*, 114, 127-137.
- BOUSQUET, P., CIAIS, P., MILLER, J. B., DLUGOKENCKY, E. J., HAUGLUSTAINE, D., A., PRIGENT, C., VEN DE RWERF, G., R., PEYLIN, P., BRUNKE, E. G., CAROUGE, C., LANGENFELDS, R. L., LATHIÈRE, J., PAPA, F., RAMONET, M., SCHMIDT, M., STEELE, L., P., TYLER, S., C. & WHITE, J. 2006. Contribution of anthropogenic and natural sources to atmospheric methane variability. *Nature*, 443, 439-443.
- 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: Biogeosciences*, 113, G02026.
- BRÉON, F.-M. & CIAIS, P. 2010. Spaceborne remote sensing of greenhouse gas concentrations. *Comptes Rendus Geoscience*, 342, 412-424.
- BROWN, J., FERRIANS, O. J., HEGINBOTTOM, J. A. & MELNIKOV, E. S. 2001a. *Circum-arctic map of permafrost and ground ice conditions*. Boulder, CO: National Snow and Ice Data Center. Boulder, USA: National Snow and Ice Data Center.
- BROWN, J., FERRIANS, O. J. & HEGINBOTTOM, J. A. M., E.S. 2001b. *Circum-arctic map of permafrost and ground ice conditions*. Boulder, CO: National Snow and Ice Data Center. Boulder, USA: National Snow and Ice Data Center.
- BUTZ, A., GUERLET, S., HASEKAMP, O., SCHEPERS, D., GALLI, A., ABEN, I., FRANKENBERG, C., HARTMANN, J. M., TRAN, H., KUZE, A., KEPPEL-ALEKS, G., TOON, G., WUNCH, D., WENNERBERG, P., DEUTSCHER, N., GRIFFITH, D., MACATANGAY, R., MESSERSCHMIDT, J., NOTHOLT, J. & WARNEKE, T. 2011. Toward accurate CO₂ and CH₄ observations from GOSAT. *Geophys. Res. Lett.*, 38, L14812.
- CHANTON, J. P. 2005. The effect of gas transport on the isotope signature of methane in wetlands. *Organic Geochemistry*, 36, 753-768.
- CHAPIN, F. S., MCGUIRE, A. D., RANDERSON, J., PIELKE, R., BALDOCCHI, D., HOBBIE, S. E., ROULET, N., EUGSTER, W., KASISCHKE, E., RASTETTER, E. B., ZIMOV, S. A. & RUNNING, S. W. 2000. Arctic and boreal ecosystems of western North America as components of the climate system. *Global Change Biology*, 6, 211-223.
- CHEN, Y.-H. 2003. Estimation of methane and carbon dioxide surface fluxes using a 3-D global atmospheric chemical transport model. In: TECHNOLOGY, M. I. O. (ed.). Cambridge, USA: Department of Earth, Atmospheric and Planetary Sciences. Center for Global Change Science.
- CHRISTENSEN, T. 1993. Methane emission from Arctic tundra. *Biogeochemistry*, 21, 117-139.
- CZUDEK, T. & DEMEK, J. 1970. Thermokarst in Siberia and its influence on the development of lowland relief. *Quaternary Research*, 1, 103-120.
- DE RIDDER, K. 2000. Remote sensing of parameters that regulate energy and water transfer at the land-atmosphere interface. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 25, 159-165.

- DENMEAD, O. T. 2008. Approaches to measuring fluxes of methane and nitrous oxide between landscapes and the atmosphere. *Plant and Soil*, 309, 5-24.
- DIXON, R. K., BROWN, S., HOUGHTON, R. A., SOLOMON, A. M., TREXLER, M. C. & WISNIEWSKI, J. 1994. Carbon Pools and Flux of Global Forest Ecosystems. *Science*, 263, 185-190
- DOWNING, J. A. & DUARTE, C. M. 2009. Abundance and Size Distribution of Lakes, Ponds and Impoundments. In: EDITOR-IN-CHIEF: GENE, E. L. (ed.) *Encyclopedia of Inland Waters*. Oxford: Academic Press.
- DUCHEMIN, E., LUCOTTE, M., CANUEL, R. & CHAMBERLAND, A. 1995. Production of the greenhouse gases CH₄ and CO₂ by hydroelectric reservoirs of the boreal region. *Global Biogeochemical Cycles*, 9, 529-540.
- DUGUAY, C. R., PULTZ, T. J., LAFLEUR, P. M. & DRAI, D. 2002. RADARSAT backscatter characteristics of ice growing on shallow sub-Arctic lakes, Churchill, Manitoba, Canada. *Hydrological Processes*, 16, 1631-1644.
- ENGRAM, M., WALTER, K. A., MEYER, F. J. & GROSSE, G. 2012. Synthetic aperture radar (SAR) backscatter response from methane ebullition bubbles trapped by thermokarst lake ice. *Canadian Journal of Remote Sensing*, 38, 667-682.
- FLETCHER M., SARA E., TANS, P. P., BRUHWILER, L. M., MILLER, J. B. & HEIMANN, M. 2004. CH₄ sources estimated from atmospheric observations of CH₄ and its ¹³C/¹²C isotopic ratios: 2. Inverse modeling of CH₄ fluxes from geographical regions. *Global Biogeochemical Cycles*, 18, GB4005.
- FRANKENBERG, C., MEIRINK, J. F., BERGAMASCHI, P., GOEDE, A. P. H., HEIMANN, M., KÖRNER, S., PLATT, U., VAN WEELE, M. & WAGNER, T. 2006. Satellite cartography of atmospheric methane from SCIAMACHY on board ENVISAT: Analysis of the years 2003 and 2004. *Journal of Geophysical Research: Atmospheres*, 111, D07303.
- FREY, K. E. & MCCLELLAND, J. W. 2009. Impacts of permafrost degradation on arctic river biogeochemistry. *Hydrological Processes*, 23, 169-182.
- FROHN, R. C., HINKEL, K. M. & EISNER, W. R. 2005. Satellite remote sensing classification of thaw lakes and drained thaw lake basins on the North Slope of Alaska. *Remote Sensing of Environment*, 97, 116-126.
- GOSAT PROJECT OFFICE. 2012. *Goals of the GOSAT Project* [Online]. Onogawa, Japan: Japan National Institute of Environmental Studies. Available: <http://www.gosat.nies.go.jp/eng/gosat/info.htm> [Accessed 30 September 2012].

References

- GROSSE, G., ROMANOVSKY, V. E., WALTER, M. A., H., L. & ZIMOV, S. 2008. Distribution of Thermokarst Lakes and Ponds at Three Yedoma Sites in Siberia. *Ninth International Conference on Permafrost*. Fairbanks: Institute of Northern Engineering UAF.
- 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.
- HALL, D. K., FAGRE, D. B., KLASNER, F., LINEBAUGH, G. & LISTON, G. E. 1994. Analysis of ERS 1 synthetic aperture radar data of frozen lakes in northern Montana and implications for climate studies. *Journal of Geophysical Research*, 99, 22,473-22,482.
- HAMILTON, J. D., KELLY, C. A., RUDD, J. W. M., HESSLEIN, R. H. & ROULET, N. T. 1994. Flux to the atmosphere of CH₄ and CO₂ from wetland ponds on the Hudson Bay lowlands (HBLs). *Journal of Geophysical Research: Atmospheres*, 99, 1495-1510.
- HEIN, R., CRUTZEN, P. J. & HEIMANN, M. 1997. An inverse modeling approach to investigate the global atmospheric methane cycle. *Global Biogeochemical Cycles*, 11, 43-76.
- HILTON, J. 1984. Airborne remote sensing for freshwater and estuarine monitoring. *Water Research*, 18, 1195-1223.
- HINKEL, K. M., EISNER, W. R., BOCKHEIM, J. G., NELSON, F. E., PETERSON, K. M. & DAI, X. 2003. Spatial Extent, Age, and Carbon Stocks in Drained Thaw Lake Basins on the Barrow Peninsula, Alaska. *Arctic, Antarctic, and Alpine Research*, 35, 291-300.
- HINKEL, K. M., JONES, B. M., EISNER, W. R., CUOMO, C. J., BECK, R. A. & FROHN, R. 2007. Methods to assess natural and anthropogenic thaw lake drainage on the western Arctic coastal plain of northern Alaska. *Journal of Geophysical Research: Earth Surface*, 112, F02S16.
- HOUWELING, S., KAMINSKI, T., DENTENER, F., LELIEVELD, J. & HEIMANN, M. 1999. Inverse modeling of methane sources and sinks using the adjoint of a global transport model. *Journal of Geophysical Research: Atmospheres*, 104, 26137-26160.
- HUTTUNEN, J. T., ALM, J., LIIKANEN, A., JUUTINEN, S., LARMOLA, T., HAMMAR, T., SILVOLA, J. & MARTIKAINEN, P. J. 2003. Fluxes of methane, carbon dioxide and nitrous oxide in boreal lakes and potential anthropogenic effects on the aquatic greenhouse gas emissions. *Chemosphere*, 52, 609-621.
- IPCC 2001. Climate Change 2001: The Scientific Basis. *Cambridge University Press*. Cambridge, United Kingdom and New York, NY, USA.

- IPCC 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. *In*: SOLOMON, S., D., MANNING, M., CHEN, Z., MARQUIS, M., AVERYT, K. B., TIGNOR, M. & MILLER, H. L. (eds.) *Cambridge University Press*. Cambridge, United Kingdom and New York, NY, USA.
- JÄHNE, B. 1987. On the parameters influencing air-water gas exchange. *Journal of Geophysical Research: Atmospheres*, 92, 1937-1950.
- JAPAN AEROSPACE EXPLORATION AGENCY 2011. *GOSAT / IBUKI Data Users Handbook*, Ministry of Environment (Japan).
- JEFFRIES, M. O., MORRIS, K., WEEKS, W. F. & WAKABAYASHI, H. 1994. Structural and stratigraphic features and ERS 1 synthetic aperture radar backscatter characteristics of ice growing on shallow lakes in NW Alaska, winter 1991–1992. *Journal of Geophysical Research: Oceans*, 99, 22459-22471.
- JORGENSEN, M. T. & SHUR, Y. 2007. Evolution of lakes and basins in northern Alaska and discussion of the thaw lake cycle. *Journal of Geophysical Research: Earth Surface*, 112, F02S17.
- JUUTINEN, S., ALM, J., LARMOLA, T., SAARNIO, S., MARTIKAINEN, P. J. & SILVOLA, J. 2004. Stand-specific diurnal dynamics of CH₄ fluxes in boreal lakes: Patterns and controls. *Journal of Geophysical Research: Atmospheres*, 109, D19313.
- KALIO, K. 2012. Water quality estimation by optical remote sensing in boreal lakes. *Monographs of the Boreal Environment Research*, No. 39, 54.
- KAPLAN, J. O., BIGELOW, N. H., PRENTICE, I. C., HARRISON, S. P., BARTLEIN, P. J., CHRISTENSEN, T. R., CRAMER, W., MATVEYEVA, N. V., MCGUIRE, A. D., MURRAY, D. F., RAZZHIVIN, V. Y., SMITH, B., WALKER, D. A., ANDERSON, P. M., ANDREEV, A. A., BRUBAKER, L. B., EDWARDS, M. E. & LOZHKIN, A. V. 2003. Climate change and Arctic ecosystems: 2. Modeling, paleodata-model comparisons, and future projections. *J. Geophys. Res.*, 108, 8171.
- KHALIL, M. A. K., RASMUSSEN, R. A., SHEARER, M. J., DALLUGE, R. W., REN, L. X. & DUAN, C. L. 1998. Measurements of methane emissions from rice fields in China. *Journal of Geophysical Research: Atmospheres*, 103, 25181-25210.
- KLING, G., KIPPHUT, G. & MILLER, M. 1992. The flux of CO₂ and CH₄ from lakes and rivers in arctic Alaska. *Hydrobiologia*, 240, 23-36.
- LEHNER, B. & DÖLL, P. 2004. Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, 296, 1-22.

References

- LEMKE, P., REN, J., ALLEY, R. B., ALLISON, I., CARRASCO, J., FLATO, G., FUJII, Y., KASER, G., MOTE, P., THOMAS, R. H. & T. ZHANG, T. 2007. Observations: Changes in Snow, Ice and Frozen Ground. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom: Cambridge University.
- LUNARDINI, V. J. 1996. Climatic warming and the degradation of warm permafrost. *Permafrost and Periglacial Processes*, 7, 311-320.
- MAXWELL, B. 1992. Arctic Climate: Potential for Change under Global Warming. *In: F. STUART CHAPIN, I., ROBERT, L. J., JAMES, F. R., GAIUS, R. S., JOSEF, S., ELLEN W. CHUA - F. STUART CHAPIN III, R. L. J. J. F. R. G. R. S. J. S. & ELLEN, W. C. (eds.) Arctic Ecosystems in a Changing Climate*. San Diego: Academic Press.
- MCKEOWN, D. M. & DENLINGER, J. L. 1984. Map-guided feature extraction from aerial imagery. *In: DEPARTMENT, C. S. (ed.)*. Carnegie Mellon University.
- MEIRINK, J. F., BERGAMASCHI, P., FRANKENBERG, C., D'AMELIO, M. T. S., DLUGOKENCKY, E. J., GATTI, L. V., HOUWELING, S., MILLER, J. B., RÖCKMANN, T., VILLANI, M. G. & KROL, M. C. 2008. Four-dimensional variational data assimilation for inverse modeling of atmospheric methane emissions: Analysis of SCIAMACHY observations. *Journal of Geophysical Research: Atmospheres*, 113, D17301.
- MIKALOFF FLETCHER, S. E., TANS, P. P., BRUHWILER, L. M., MILLER, J. B. & HEIMANN, M. 2004. CH₄ sources estimated from atmospheric observations of CH₄ and its ¹³C/¹²C isotopic ratios: 2. Inverse modeling of CH₄ fluxes from geographical regions. *Global Biogeochemical Cycles*, 18, GB4005.
- MOORE, G. K. 1980. Satellite remote sensing of water turbidity / Sonde de télémessure par satellite de la turbidité de l'eau. *Hydrological Sciences Bulletin*, 25, 407-421.
- MORGENSTERN, A., GROSSE, G., GUNTHER, F., FEDOROVA, I. & SCHIRRMEISTER, L. 2011. Spatial analyses of thermokarst lakes and basins in Yedoma landscapes of the Lena Delta. *The Cryosphere*, 5.
- MORRIS, K., JEFFRIES, M. O. & WEEKS, W. F. 1995. Ice processes and growth history on Arctic and sub-Arctic lakes using ERS-1 SAR data. *Polar Record*, 31, 115-128.
- MORRISSEY, L. A. & LIVINGSTON, G. P. 1992. Methane Emissions From Alaska Arctic Tundra: An Assessment of Local Spatial Variability. *J. Geophys. Res.*, 97, 16661-16670.

- OECHEL, W. C. & VOURLITIS, G. L. 1994. The effects of climatic change on land–atmosphere feedbacks in arctic tundra regions. *Trends in Ecology & Evolution*, 9, 324-329.
- OLEFELDT, D., TURETSKY, M. R., CRILL, P. M. & MCGUIRE, A. D. 2013. Environmental and physical controls on northern terrestrial methane emissions across permafrost zones. *Global Change Biology*, 19, 589-603.
- PALMER, P. I. 2008. Quantifying sources and sinks of trace gases using space-borne measurements: current and future science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366, 4509-4528.
- PING, C. L., MICHAELSON, G. J. & KIMBLE, J. M. 1997. Carbon storage along a latitudinal transect in Alaska. *Nutrient Cycling in Agroecosystems*, 49, 235-242.
- RHODE, H. 1990. A Comparison of the Contribution of Various Gases to the Greenhouse Effect. *Science*, 248, 1217-19.
- ROACH, J. K., GRIFFITH, B. & VERBYLA, D. 2012. Comparison of three methods for long-term monitoring of boreal lake area using Landsat TM and ETM+ imagery. *Canadian Journal of Remote Sensing*, 38, 427-440.
- ROBINSON, G. M. 2009. Statistics, Overview. In: EDITORS-IN-CHIEF: ROB, K. & NIGEL, T. (eds.) *International Encyclopedia of Human Geography*. Oxford: Elsevier.
- ROULET, N. T., JANO, A., KELLY, C. A., KLINGER, L. F., MOORE, T. R., PROTZ, R., RITTER, J. A. & ROUSE, W. R. 1994. Role of the Hudson Bay lowland as a source of atmospheric methane. *Journal of Geophysical Research: Atmospheres*, 99, 1439-1454.
- SHAKHOVA, N., SERGIENKO, V. & SEMILETOV, I. 2009. The contribution of the East Siberian shelf to the modern methane cycle. *Herald of the Russian Academy of Sciences*, 79, 237-246.
- SMITH, L. C., SHENG, Y. & MACDONALD, G. M. 2007. A first pan-Arctic assessment of the influence of glaciation, permafrost, topography and peatlands on northern hemisphere lake distribution. *Permafrost and Periglacial Processes*, 18, 201-208.
- STOW, D., HOPE, A., BOYNTON, W., PHINN, S., WALKER, D. & AUERBACH, N. 1998. Satellite-derived vegetation index and cover type maps for estimating carbon dioxide flux for arctic tundra regions. *Geomorphology*, 21, 313-327.
- TAKEUCHI, W., NAKANO, T. & YASUOKA, Y. 2011. Estimation of Methane Emission from West Siberian Lowland with Subpixel Land Cover Characterization Between MODIS and ASTER

References

- Land Remote Sensing and Global Environmental Change. *In:* RAMACHANDRAN, B., JUSTICE, C. O. & ABRAMS, M. J. (eds.). Springer New York.
- WALTER, K. M., EDWARDS, M. E., GROSSE, G., ZIMOV, S. A. & CHAPIN, F. S. 2007a. Thermokarst Lakes as a Source of Atmospheric CH₄ During the Last Deglaciation. *Science*, 318, 633-636.
- WALTER, K. M., ENGRAM, M., DUGUAY, C. R., JEFFRIES, M. O. & CHAPIN, F. S. I. 2008. Potential Use of Synthetic Aperture Radar for Estimating Methane Ebullition from Arctic Lakes. *JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION*, 44, 305-315.
- WALTER, K. M., SMITH, L. & CHAPIN, S. 2007b. Methane bubbling from northern lakes: present and future contributions to the global methane budget. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365, 1657-1676.
- WALTER, K. M., VAS, D. A., BROSIUS, L., CHAPIN III, F. S., ZIMOV, S. A. & Q., Z. 2010. Estimating methane emissions from northern lakes using ice-bubble surveys. *Limnology and Oceanography: Methods*, 8, 592-609.
- WALTER, K. M., ZIMOV, S. A., CHANTON, J. P., VERBYLA, D. & CHAPIN, F. S. 2006. Methane bubbling from Siberian thaw lakes as a positive feedback to climate warming. *Nature*, 443, 71-75.
- WANNINKHOF, R. 1992. Relationship between wind speed and gas exchange over the ocean. *Journal of Geophysical Research: Oceans*, 97, 7373-7382.
- WASHENFELDER, R. A., TOON, G. C., BLAVIER, J. F., YANG, Z., ALLEN, N. T., WENBERG, P. O., VAY, S. A., MATROSS, D. M. & DAUBE, B. C. 2006. Carbon dioxide column abundances at the Wisconsin Tall Tower site. *Journal of Geophysical Research: Atmospheres*, 111, D22305.
- WASHENFELDER, R. A., WENBERG, P. O. & TOON, G. C. 2003. Tropospheric methane retrieved from ground-based near-IR solar absorption spectra. *Geophysical Research Letters*, 30, 2226.
- WETTERICH, S., RUDAYA, N., TUMSKOY, V., ANDREEV, A. A., OPEL, T., SCHIRRMEISTER, L. & MEYER, H. 2011. Last Glacial Maximum records in permafrost of the East Siberian Arctic. *Quaternary Science Reviews*, 30, 3139-3151.
- WORTHY, D. E. J., LEVIN, I., HOPPER, F., ERNST, M. K. & TRIVETT, N. B. A. 2000. Evidence for a link between climate and northern wetland methane emissions. *Journal of Geophysical Research: Atmospheres*, 105, 4031-4038.

- YAMANOUCHI, T. 2011. Early 20th century warming in the Arctic: A review. *Polar Science*, 5, 53-71.
- YAN, X., CAI, Z., OHARA, T. & AKIMOTO, H. 2003. Methane emission from rice fields in mainland China: Amount and seasonal and spatial distribution. *Journal of Geophysical Research: Atmospheres*, 108, 4505.
- ZHU, R., LIU, Y., XU, H., HUANG, T., SUN, J., MA, E. & SUN, L. 2010. Carbon dioxide and methane fluxes in the littoral zones of two lakes, east Antarctica. *Atmospheric Environment*, 44, 304-311.
- ZHU, X., QIANLAI, Z., CHEN, M., SIRIN, A., MELILLO, J., KICKLIGHTER, D., SOKOLOV, A. & SONG, L. 2013. Methane emissions from wetlands: biogeochemical, microbial, and modeling perspectives from local to global scales. *Global Change Biology*, n/a.
- ZHUANG, Q., MELILLO, J. M., KICKLIGHTER, D. W., PRINN, R. G., MCGUIRE, A. D., STEUDLER, P. A., FELZER, B. S. & HU, S. 2004. Methane fluxes between terrestrial ecosystems and the atmosphere at northern high latitudes during the past century: A retrospective analysis with a process-based biogeochemistry model. *Global Biogeochemical Cycles*, 18, GB3010.
- ZIMOV, S., DAVYDOV, S., ZIMOVA, G., DAVYDOVA, A., SCHUUR, E., DUTTA, K. & CHAPIN III, F. 2006. Permafrost carbon: Stock and decomposability of a globally significant carbon pool. *Geophysical Research Letters*, 33.
- ZIMOV, S. A., VOROPAEV, Y. V., DAVYDOV, S. P., ZIMOVA, G. M., DAVYDOVA, A. I., CHAPIN, F. S. & CHAPIN, M. C. 2001. Flux of methane from north siberian aquatic systems: Influence on atmospheric methane. In: PAEPE, R., MELNIKOV, V., OVERLOOP, E. & GOROKHOV, V. (eds.) *Permafrost Response on Economic Development, Environmental Security and Natural Resources*. Springer Netherlands.
- ZIMOV, S. A., VOROPAEV, Y. V., SEMILETOV, I. P., DAVIDOV, S. P., PROSIANNIKOV, S. F., CHAPLIN, F. S., CHAPIN, M. C., TRUMBORE, S. & TYLER, S. 1997. North Siberian lakes: a methane source fueled by Pleistocene carbon. *Science*, 277, 800-802.