GIScience & Remote Sensing

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tgrs20

Diversity of lakes and ponds in the forest-tundra ecozone: from limnicity to limnodiversity

Pedro Freitas, Gonçalo Vieira, Diana Martins, João Canário, Pedro Pina, Birgit Heim, Bennet Juhls, Raoul-Marie Couture & Warwick F. Vincent

To cite this article: Pedro Freitas, Gonçalo Vieira, Diana Martins, João Canário, Pedro Pina, Birgit Heim, Bennet Juhls, Raoul-Marie Couture & Warwick F. Vincent (2024) Diversity of lakes and ponds in the forest-tundra ecozone: from limnicity to limnodiversity, GIScience & Remote Sensing, 61:1, 2427309, DOI: 10.1080/15481603.2024.2427309

To link to this article: https://doi.org/10.1080/15481603.2024.2427309

Published online: 13 Nov 2024.

GISCIENCE & REMOTE SENSING
2024, VOL. 61, NO. 1, 2427309
https://doi.org/10.1080/15481603.2024.2427309

Taylor & Francis
Taylor & Francis Group

Diversity of lakes and ponds in the forest-tundra ecozone: from limnicity to limnodiversity

Pedro Freitas [a,b], Gonçalo Vieira [a,b], Diana Martins [a], João Canário [b,c], Pedro Pina [d,e], Birgit Heim [f], Bennet Juhls [f], Raoul-Marie Couture [b,g] and Warwick F. Vincent [b,h]

[a]Centro de Estudos Geográficos, Laboratório Associado TERRA, Instituto de Geografia e Ordenamento do Território, Universidade de Lisboa, Lisbon, Portugal; [b]Centre d'études nordiques (CEN), Université Laval, Quebec, QC, Canada; [c]Centro de Química Estrutural, Institute of Molecular Sciences and Department of Chemical Engineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal; [d]Departamento de Ciências da Terra, Universidade de Coimbra, Coimbra, Portugal; [e]Instituto Dom Luiz, Universidade de Coimbra, Coimbra, Portugal; [f]Polar Terrestrial Environmental Systems, Alfred Wegener Institute (AWI), Helmholtz Centre for Polar and Marine Research, Potsdam, Germany; [g]Département de Chimie, Université Laval, Quebec, Canada; [h]Takuvik International Laboratory & Département de Chimie, Université Laval, Quebec, Canada

ABSTRACT

Arctic and subarctic landscapes have unique hydrological and limnological features and are now experiencing rapid change due to climate warming and permafrost thaw. The highly abundant lakes, ponds, and rivers across these landscapes play an increasingly important role in global biogeochemical cycles and are sentinels of environmental changes. However, studying these remote waters poses challenges for both in situ sampling and remote-sensing analysis. Here we developed a synergistic remote-sensing strategy that combined PlanetScope and Sentinel-2 satellite data to estimate limnicity (water fraction per land surface), limnodensity (density of water bodies), and limnodiversity (optical diversity of water bodies) along a boreal forest-tundra transect, from the non-permafrost to the continuous permafrost zones of western Nunavik (Subarctic Canada). Our analyses show that this region hosts 335,281 water bodies, around 90% in the 0.0001 to 0.01 km^2 size range. In bedrock outcrops, large water bodies were mostly associated with glacially carved depressions (higher limnicity). In contrast, small water bodies were predominately found in sedimentary infills along valleys (higher limnodensity). The discontinuous permafrost zone had the highest limnodensity and limnodiversity. This was likely due to permafrost thaw (thermokarst), particularly the collapse, subsidence, and erosion of palsas (organic permafrost mounds), resulting in ponds with black- and brown-colored waters, and lithalsas (mineral permafrost mounds), resulting in ponds with brown, light-brown, and sometimes white-colored waters. Some of these limnodense and limnodiverse landscapes, although covering only 2 to 7% of the total area of the study region, contained over one-third (34%) of the total number of water bodies, 97% of which were <0.01 km^2; they accounted for a small proportion of the total black-colored water bodies (23%), but a high proportion of the total brown- (60%) and light brown-colored water bodies (92%) throughout the region. This research underscores the utility of optical satellite remote sensing for assessing water body types and for evaluating their individual and distinct aquatic responses to climate change. The dataset may be used to improve the modeling of carbon fluxes by better categorizing small water bodies affected by organic or mineral soil type settings. This is an important factor dictating biogeochemical responses, with effects on albedo, climate feedbacks, and ecosystem dynamics in the boreal forest-tundra region. The framework developed here may be applied to landscapes elsewhere in the world that have high densities of water bodies of variable size and optical properties.

ARTICLE HISTORY
Received 22 June 2024
Accepted 2 November 2024

KEYWORDS
Forest-tundra ecozone; permafrost; lakes; ponds; color; limnicity

1. Introduction

The Arctic and subarctic have distinctive hydrological and limnological settings due to the ice-sheet dynamics of the Pleistocene, which gave rise to a postglacial landscape of wetlands, widespread water bodies, and extensive permafrost (Lehner and Döll 2004; Smith, Sheng, and Macdonald 2007; Verpoorter et al. 2014; Webb et al. 2022). These landscapes are undergoing marked changes due to climate warming, at a faster and stronger pace than any other region, with local and global consequences (Biskaborn et al. 2019; Rantanen et al. 2022; Schuur et al. 2015).

High-latitude water bodies are sentinels of landscape change and environmental state shifts linked to climate warming, permafrost thaw, and changes in precipitation and snow regimes (Magnuson et al. 1997; Saros et al. 2023; Webb et al. 2022). These climate impacts affect aquatic ecosystems and their

CONTACT Pedro Freitas pedro-freitas@edu.ulisboa.pt

catchments, resulting in autochthonous modifications such as shifts in primary production, nutrient cycling, and microbial activity, as well as allochthonous modifications such as inputs of organic matter and contaminants from the surrounding landscape. These water quality modifications further affect ecosystem services, biodiversity, and the climate system (Vincent et al. 2012). In the northern circumpolar region, rocky highlands typically support large water bodies (>0.01 km^2) associated with glacially carved depressions and tectonic processes, while lowlands often sustain wetlands able to support an enormous variety of ponds (<0.01 km^2), which are known to be biogeochemically more active (Holgerson and Raymond 2016; Muster et al. 2017; Vincent and Laybourn-Parry 2008).

Permafrost covers 15% of the exposed land surface in the Northern Hemisphere and hosts a large portion of the world's wetlands (5–25%) (Gorham 1991; Olefeldt et al. 2016; Treat et al. 2024). Within these, peatlands are biogeochemically distinct ecosystems (Gorham 1991). These water-saturated environments can act as carbon sinks (Harris et al. 2021; Zhang and Väliranta 2024) but also include shallow and morphologically and optically diverse lakes and ponds that are intense carbon sources (Arsenault et al. 2022; Hassan et al. 2023; Kuhn et al. 2021; Taillardat et al. 2024). The formation and development of wetlands depend on various factors, including climate (moisture, temperature, and precipitation), soils (organic or mineral), topography, and vegetation structure (fen, bog, shrublands, and forest) (Gorham 1991). These factors also influence the genesis, role, evolution, and limnological characteristics of the water bodies (Arsenault et al. 2022; Bouchard et al. 2014). However, studying landscapes replete with biogeochemically diverse water bodies poses specific challenges not only for traditional methods focused on in situ sampling but also for remote-sensing analysis.

Disturbances in permafrost wetlands are often due to abrupt permafrost thaw processes, such as the formation of thermokarst water bodies (permafrost thaw lakes and ponds). These can lead to increased soil carbon emissions, more than doubling those from gradual thaw (≈125 to 190%) (Walter Anthony et al. 2018). However, they are not accurately depicted in Earth-System models (Kuhn et al. 2018; Turetsky et al. 2020; Vonk, Tank, and Walvoord 2019; Walter Anthony et al. 2018). Diverse types of thermokarst water bodies occur in different permafrost zones, primarily linked to distinct geomorphological processes, namely polygonal tundra degradation in the continuous permafrost zone and degradation of cryogenic mounds (organic-rich palsa and mineral-rich lithalsa) in the discontinuous and sporadic permafrost zones (Bouchard et al. 2017; Grosse, Jones, and Arp 2013).

As wetlands and peatlands expand, permafrost thermokarst water bodies are predicted to increase in number and size, covering 545,000 km^2 by 2100 and 1,048,000 km^2 by 2300, potentially doubling carbon emissions to 72 ± 60 Pg (Heslop et al. 2020; Turetsky et al. 2020). Initially small (<0.01 km^2) and shallow (<5 m), these water bodies play disproportionate biogeochemical roles in the landscape by processing large amounts of carbon and nutrients, caused by frequent water column mixing, shoreline erosion processes, high terrestrial inputs per water volume, and other interactions with the surrounding landscape (Arsenault et al. 2022; Bégin and Vincent 2017; Heslop et al. 2020; Holgerson and Raymond 2016; Kuhn et al. 2018; Zandt, Liebner, and Welte 2020). They are especially strong emitters of the potent greenhouse gas methane, as well as carbon dioxide (Abnizova et al. 2012; Negandhi et al. 2013; Wik et al. 2016; Zandt, Liebner, and Welte 2020). These conditions lead to distinct aquatic ecosystem responses, challenging the effectiveness of current models based on existing lentic ecosystem data (Arsenault et al. 2022; Taillardat et al. 2024).

Recent studies have focused on the biogeochemical distinctiveness of various lentic systems, ranging from large lakes to small water bodies, including peatland and thermokarst ponds (Arsenault et al. 2022; Hassan et al. 2023; Taillardat et al. 2024). In particular, Arsenault et al. (2022) and Taillardat et al. (2024) found that while small water bodies, such as non-peatland and peatland ponds, generally represented a specific and well-defined category within lentic systems due to the more significant role of the organic matter inputs compared to large water bodies, modeling the behavior of small thermokarst water bodies proved challenging, with bimodal pH distributions and distinct patterns in nutrients and dissolved organic carbon. This strong thermokarst biogeochemical variability can be attributed in part to their diverse origins, from organic-rich (peat) or mineral-rich (marine silt, clay, and sand deposits) soils (Arsenault et al.

2022; Heslop et al. 2020). These preconditions affect aquatic ecosystem structure and functioning, including dissolved oxygen availability, turbidity, underwater light, redox potential, primary production, and microbial community structure (Arsenault et al. 2022; Bouchard et al. 2011; Deshpande et al. 2017).

Optical remote sensing offers a promising approach for assessing and monitoring small water bodies, including thermokarst, at the regional or even global scale, and also allows the upscaling of in situ observations of water properties. However, the success of this approach is often constrained by persistent cloud cover and resultant limited optical data, challenges on precise automatic delineation of small water bodies (<0.01 km^2), as well as on retrieving their morphometrical and optical characteristics at native sensor resolution (Freitas et al. 2024; Mullen et al. 2023; Muster et al. 2017; Olefeldt et al. 2021; Pekel et al. 2016). Remote sensing workflows for overcoming these problems are currently lacking.

Advances and continuous acquisitions of satellite imagery and its distribution are not only allowing consistent and improved water body delineations (Cooley et al. 2017; Freitas et al. 2024; Mullen et al. 2023) but also spatial assessments of water transparency and color, as well as the monitoring of optically active constituents such as algal populations, colored dissolved organic matter, and suspended particulate matter (Juhls et al. 2022; Pahlevan et al. 2022; Shang et al. 2021; Toming et al. 2016). These technologies provide insights into the role and evolution of high-latitude water bodies, supporting regional and global monitoring and modeling efforts (Arsenault et al. 2022; Hassan et al. 2023; Kokelj and Jorgenson 2013).

Here, we developed a synergistic optical remote-sensing strategy combining PlanetScope-Dove (PS-D) and Sentinel-2 (S2) imagery for assessing the distribution and diversity of water bodies along a latitudinal gradient across the tundra forest zone of western Nunavik (Subarctic Canada) from non-permafrost to continuous permafrost. This region provided the ideal pond-rich and lake-rich landscapes (limnoscapes) to develop and apply this approach that will likely be of application in similar regions of the circumpolar North. In particular, we hypothesized that thermokarst landscapes, resulting from contemporary permafrost degradation along the forest-tundra ecozone, may contain not only high densities of small water bodies (<0.01 km^2) due to patterned ground, ice segregation, and cryoturbation processes occurring under suitable climatic cold conditions but also high optical diversity of waters depending on the geomorphological setting, namely following mineral-rich lithalsa or organic-rich palsa degradation.

Our approach involved deriving a Very-High Resolution (VHR) water body delineation database using a trained Mask R-CNN deep learning model (HLWATER) over PS-D data (Freitas et al. 2024) and then retrieving their optical properties using S2 data (HLWATER-Optical). This combined data analysis allowed the estimation of the water fraction per land surface (limnicity), the density of water bodies (limno-density), and the optical diversity of water bodies (limnodiversity). The analyses covered a broad range of water body sizes, from small (10^{-4}–10^{-2} km^2) to large (>0.01 km^2), thereby allowing an overall evaluation of waters across subarctic forest-tundra landscapes, including permafrost thaw (thermokarst) environments.

2. Study area

The study region was located in western Nunavik, on the eastern side of Hudson Bay, Canada, and covered a total area of 41,832 km^2 within the latitudes 54° to 58° N and the longitudes 74° to 78° W. This region lies in the subarctic forest-tundra transition region, from non-permafrost to continuous permafrost zones (Figure 1). The region is part of the Precambrian Canadian Shield, characterized by granitic-gneissic formations (Bhiry et al. 2011). It has undergone multiple glacial cycles, with the most recent being the Wisconsin glaciation, which ended with the northward and eastward retreat of the Laurentide Ice Sheet at 8 ka (Harden et al. 1992). Consequently, this region exhibits one of the fastest glacial-isostatic uplifts in the world of about 1.3–1.5 cm per year (Andrews 1968; Bhiry et al. 2011).

Following the glacial retreat, the transgression of the Tyrrell Sea at ≈7.9 ka filled large glacial valleys and structural depressions with marine clay, silt, and sand deposits (Bhiry et al. 2011). These deposits are currently found along coastal valleys at elevations between 150 and 300 m (Allard and Seguin 1987; Veilleux, Bhiry, and Decaulne 2020). The marine regression favored vegetation colonization, peatland accumulation (≈6–4.7 ka), periglacial

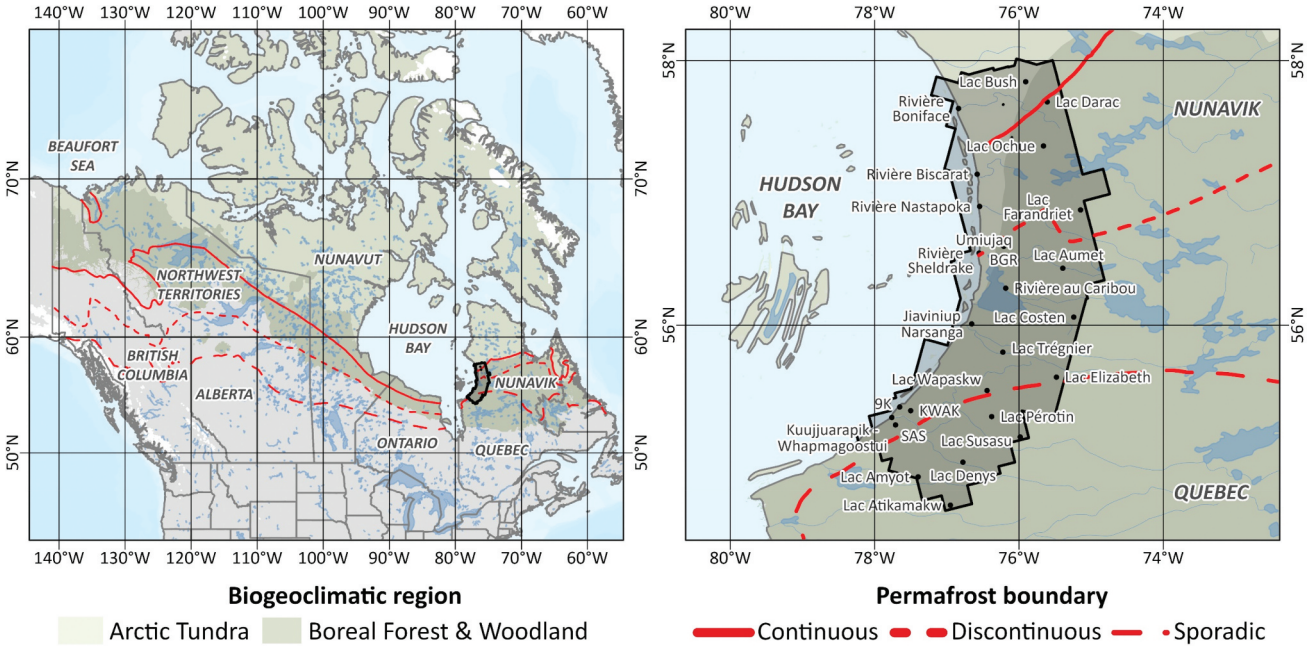

Figure 1. Regional study site location in Canada and western Nunavik (eastern Hudson Bay), extending across the boreal forest-tundra transition zone, and from the non-permafrost to continuous permafrost zones. The permafrost boundaries (displayed in red) represent the southern limit of each zone. The base information is from the Government of Canada (CanVec, permafrost atlas and biogeoclimatic regions).

processes with permafrost aggradation (≈1.9–1.2 ka), and the formation of cryogenic mounds such as lithalsas and palsas (Allard and Seguin 1987; Bhiry et al. 2011). As a result, this region contains some of the most pristine wetlands and peatlands in the world (Darnajoux et al. 2015).

Since the end of the Little Ice Age (≈0.3 ka) and especially in recent decades, the region has experienced accelerated permafrost degradation due to increasing air temperatures (0.4°C per decade) and precipitation, particularly during the summer (Beck et al. 2015; Bhiry et al. 2011; Fortier et al. 2023; Owczarek et al. 2020; Payette et al. 2004; Vallée and Payette 2007). This degradation has led to the formation and proliferation of widespread thermokarst water bodies, with diverse limnological, optical, and morphometrical characteristics, as well as greenhouse gas emission potential (Bouchard et al. 2014; Folhas et al. 2020; Freitas et al. 2019, 2022; Laurion et al. 2010; Matveev, Laurion, and Vincent 2019; Wang et al. 2018; Watanabe et al. 2011). These thermokarst water bodies, until recently unaccounted for geographically, vary in color from black/brown to light-brown/whitish, depending on the properties of the soil, genesis, and age, specifically following the degradation of organic-rich palsas or mineral-rich lithalsas, respectively (Bégin and Vincent 2017; Bouchard et al. 2011, 2014; Watanabe et al. 2011) (Figure 2).

3. Material and methods

3.1. General workflow

Beginning with a PS-D mosaic for the regional study sector, we used the HLWATER trained deep learning Mask R-CNN model from Freitas et al. (2024) to derive a Very-High Resolution (VHR) database with 335,281 automatically delineated water bodies, setting the basis for their optical characterization. This dataset was used as a reference for zonal (for water bodies >0.01 km^2) and pixel-targeted (for water bodies <0.01 km^2) water body reflectance retrievals using S2 satellite data, generating the HLWATER-Optical database. The PS-D and S2 imagery were combined through local co-registration using the AROSICS Python package developed by Scheffler et al. (2017). To mitigate the impacts of outliers on the subsequent clustering results, we first implemented a multivariate unsupervised k-Nearest Neighbors (kNN) algorithm using the PyOD Python library of Zhao, Nasrullah, and Li (2019). Finally, for automatically establishing the main water body

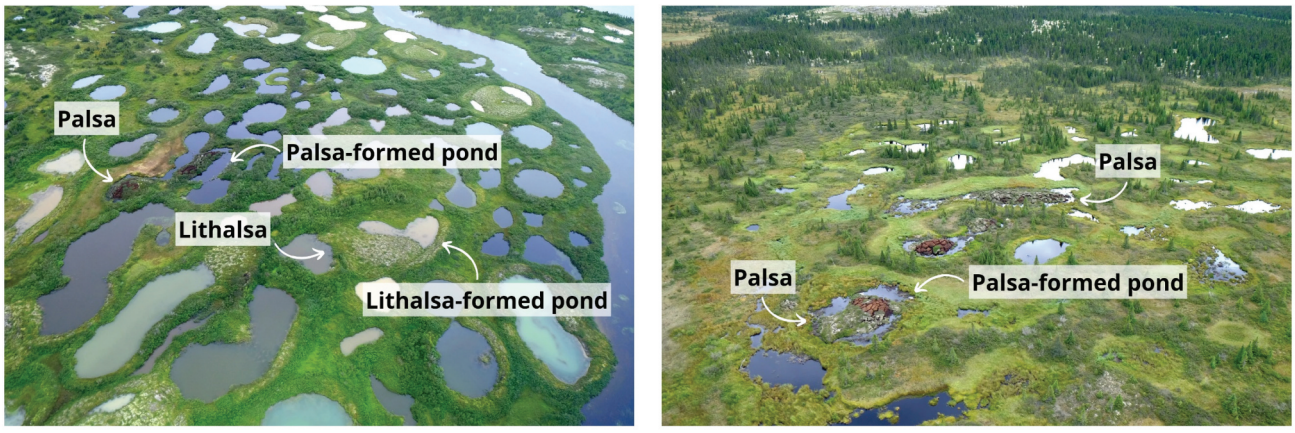

Figure 2. Oblique aerial photographs of thermokarst landscapes with palsa-formed (organic soil) and lithalsa-formed (mineral soil) ponds in the wetlands of BGR (56°36′34.84″N; 76°12′58.55″W) in the Sheldrake River basin and in the bog peatland of SAS (55° 13′6.84″N; 77°42′26.95″W) close to Sasapimakwananisikw River. The former have whitish/light brown colored waters (light scattering by mineral particulates), and the latter have dark brown or black colored waters (absorption due to high concentrations of particulate and dissolved organic matter from peat).

optical clusters, we used K-Means as unsupervised clustering method (Ahmed, Seraj, and Mohammed 2020) (Figure 3).

The HLWATER model resultant database was analyzed at the regional scale to evaluate limnicity (water fraction per land surface) and limnodensity (number

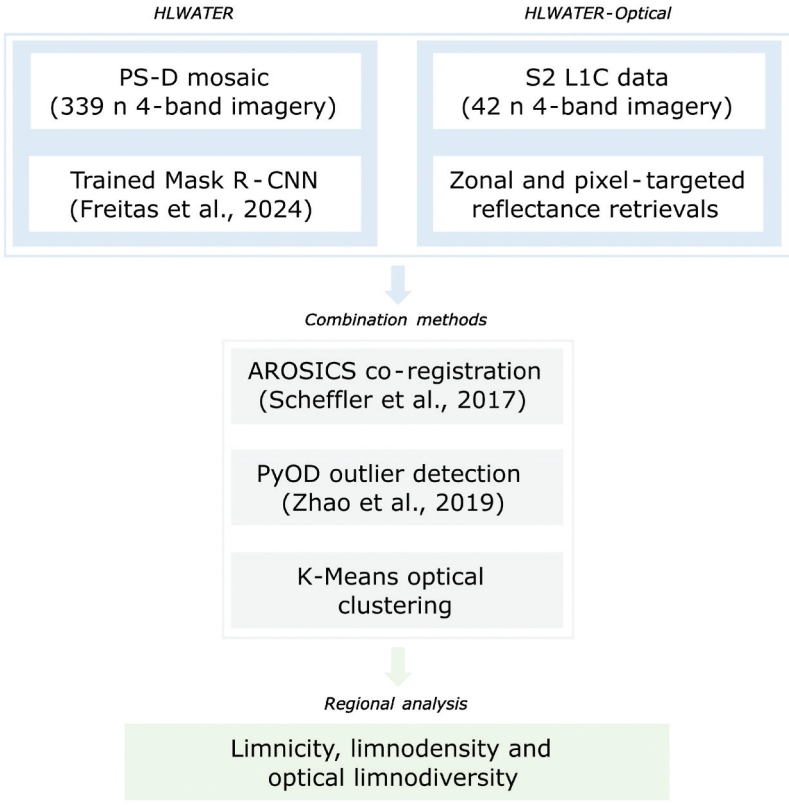

Figure 3. General methodological workflow showing the combination of materials, methods, and outputs for achieving Very-High Resolution (VHR) water delineation (HLWATER) and optical assessment (HLWATER-Optical) products at the regional scale.

of individual centroids per land surface). In addition, the HLWATER-Optical database enabled the analysis of the optical properties (notably color) of the water bodies, resulting in the evaluation of their optical limnodiversity (diversity of optical groups per land surface), highlighting its significance as a key hydrological and biogeochemical component of the rapidly changing northern-landscapes.

3.2. HLWATER

We employed the HLWATER model to classify a PS-D mosaic with approximately 3-m spatial resolution and 4-band spectral resolution (blue, green, red, and NIR). This mosaic comprised 339 scenes acquired from 2017 to 2019 during the ice and snow-free period between July and September. The implementation of HLWATER allowed the automatic delineation of 335,281 water bodies in the regional sector, covering lakes (>0.01 km^2), ponds (<0.01 km^2), rivers, streams, and creeks.

As noted in Freitas et al. (2024), the fully autonomous performance of HLWATER varied across different landscape units, achieving mean Intersection over Union (IoU) 0.5 F1 scores between 0.53 and 0.71 and mean F1 scores between 0.62 and 0.95. The model demonstrated an optimal minimum detection water body size threshold of 166 m^2 for western Nunavik (eastern Hudson Bay). Consequently, delineations below this threshold were automatically removed.

For feature tile classification, the confidence of the HLWATER was set to 50%, the best alternative to maximize true positive and minimize false positive and false negative classifications (e.g. highest F1 score) (Freitas et al. 2024). To ensure data quality, all false positives were reviewed and removed as necessary. This procedure was conducted by a single operator through visual inspection of the PS-D mosaic as a near-infrared false color composite (R – near-infrared, G – green, B – blue), aided by available VHR satellite imagery (e.g. ESRI World Imagery, BingMaps, Google Earth) following Qayyum et al. (2020), Nitze et al. (2021) and Mullen et al. (2023).

3.3. HLWATER-Optical

Using the PS-D mosaic for a homogeneous optical assessment of the water bodies at the regional scale was impractical due to the numerous scenes acquired on different dates. Consequently, using this mosaic would lead to the properties of the water bodies being determined by temporal variations instead of spatial distributions. In addition, PS-D data is known to suffer from geometric, radiometric, and calibration issues (e.g. low signal-to-noise ratio), which particularly affects its sensitivity for water applications (Dash and Ogutu 2016; Frazier and Hemingway 2021; Maciel et al. 2020; Mullen et al. 2023). In contrast, S2 has demonstrated proven capabilities for inland and coastal water quality monitoring over wide regional sectors and timeframes (Zeng et al. 2023), including small water bodies (Freitas et al. 2019). As a result, this assessment was conducted using S2 10-m data (visible and NIR range). S2 provided high-quality spectral information, enabling the retrieval of optical properties for the maximum possible number of water bodies from the HLWATER database, thus forming the HLWATER-Optical dataset.

Based on the analysis of the S2 granules extent and metadata downloaded from Google Earth Engine, a total of 14 granules were required to cover the entire regional sector, each comprehending an approximate area of 110 × 110 km (Gorelick et al. 2017). To minimize the impact of cloud cover across the region, we selected data from 23 August 2019 under predominantly cloud-free conditions. Specifically, 13 out of 14 granules exhibited less than 10% of cloud cover, with one granule showing less than 20%. In addition, this date ensured mid warm season conditions, which were important for assessing the optical properties of the water bodies.

The S2 data were screened for no cloud coverage and downloaded as Level-1C (L1C) from the Copernicus Data Space Ecosystem. The water bodies that clouds and shadows could still impact were automatically removed from this assessment by considering the S2 L2A Sen2Cor Scene Classification Map (SCL) corresponding classes. The SCL product is automatically generated during the execution of the Sen2Cor standard atmospheric correction model and in this case was downloaded from Google Earth Engine for the matching imagery sets (Gorelick et al. 2017). We based our analysis on Top-of-the-Atmosphere (TOA) reflectance to retrieve the optical properties of the water bodies since the Bottom-of-the-Atmosphere (BOA) S2 Level-2A atmospherically corrected data results from Sen2Cor, which is designed for the land surface (Warren et al. 2019). Due to the strong atmospheric scattering affecting the TOA blue spectral

reflectance, which is known to be particularly pronounced over water bodies, we excluded the S2 blue band (490 nm) from our assessment (Kutser et al. 2004; Soriano-González et al. 2022). As a result, we supported our optical assessment focusing on the S2 green (560 nm) and red (665 nm) bands. Toming et al. (2016) demonstrated that S2 green and red TOA reflectance can be successfully used for retrieving inland water body optical properties (color) and colored dissolved organic matter. Additionally, we utilized the S2 NIR band (842 nm) to identify outliers.

3.4. Satellite imagery combination

To ensure consistency between the water body boundaries derived from the deep learning model over the PS-D mosaic (HLWATER) and the S2 acquisitions (HLWATER-Optical), we developed a framework to mitigate co-registration incompatibilities, as well as avoid sub-pixel errors on reflectance due to water body size and morphometric complexity and also considering potential adjacency effects, such as surrounding vegetation scattering and cast shadows impacts (Freitas et al. 2022; Paulino et al. 2022). Given the extensive area of the study region and the large number of water bodies in the HLWATER database, these were critical steps for obtaining accurate reflectance measurements and ensuring the robustness of the HLWATER-Optical database.

3.4.1. Local co-registration

The PS-D and S2 satellites exhibit differences in data acquisition conditions, including viewing geometry, sensor characteristics, and in-product consistency. Consequently, all S2 scenes were co-registered to the PS-D mosaic using the COREG_LOCAL function from AROSICS (Scheffler et al. 2017). We used cubic resampling, a grid resolution of 50 pixels and a window size of 64×64 pixels. The tie point filter level was set to 2 out of 3 (e.g. Reliability, MSSIM, but RANSAC), which was found to enhance the algorithm's ability to perform higher absolute shift adjustments using more tie points in challenging and significantly displaced sectors.

Analyzing the AROSICS outputs per S2 granule, the mean number of valid True Positive (TP) tie points ranged from 399 to 31,602, depending on the extent of the S2 scene over our PS-D mosaic, with a global mean and standard deviation of 11,950 and 13,480 tie points, respectively (Table 1). Per granule, the RMSE absolute shift vector length ranged from 5 to 8 m, as the mean absolute computed shifts, showing a global mean of 4.6 m and standard deviation of 0.8 m. Typically, the necessary resolved absolute mean shifts were less than 1/2 S2 pixel (Table 1).

Locally, the computed and corrected absolute mean shifts (m) were dependent on the different PS-D individual scenes used to produce our mosaic dataset (Figure 4). Scenes acquired on the same day usually showed similar and uniform adjustment patterns, including mean shift directions. General adjustments of less than 1/2 S2 pixels were necessary, with the greatest adjustments reaching up to 5 pixels (e.g. 56 m) in specific sectors, underscoring the importance of performing this

Table 1. S2 to PS-D co-registration statistics. Valid True Positive (TP) tie points, Root Mean Squared Error (RMSE) of absolute shift vector length, Mean Squared Error (MSE) of absolute shift vector length, Mean Absolute Error (MAE) of absolute shift vector length, absolute shift, and angle per S2 granule adjustment to the PS-D mosaic and overall statistics (Mean, Median and Standard Deviation). Each granule's geographic extension is shown in Figure 4.

Granule	Valid TP	RMSE (m)	MSE (m)	MAE (m)	Absolute Shift (m)	Angle (°)
UPA	463	4.9	23.7	4.7	4.7	129.5
UPB	2917	5.5	30.8	5.4	5.4	143.7
UUF	12658	4.3	18.8	4.1	4.1	92.2
UUG	29917	4.6	21.3	4.4	4.4	118.5
UVF	2404	5.2	27.3	5.0	5.0	96.4
UVG	25695	4.5	20.3	4.2	4.2	116.8
VUH	4696	4.9	24.2	4.7	4.7	133.2
VUJ	4637	3.6	12.7	3.5	3.5	111.8
VUK	2449	3.6	12.8	3.5	3.5	112.9
VVH	36297	4.5	20.1	4.1	4.1	162.2
VVJ	34393	5.9	34.2	5.0	5.0	126.1
VVK	8945	5.5	30.8	4.6	4.6	112.4
VWH	831	3.5	12.3	3.4	3.4	145.6
VWJ	997	3.4	11.3	3.3	3.3	154.4
$\bar{X}$	11,950	4.6	21.5	4.3	4.3	125.4
M_d	6,821	4.5	20.8	4.2	4.2	122.0
σ	13,480	0.8	7.5	0.7	0.7	20.7

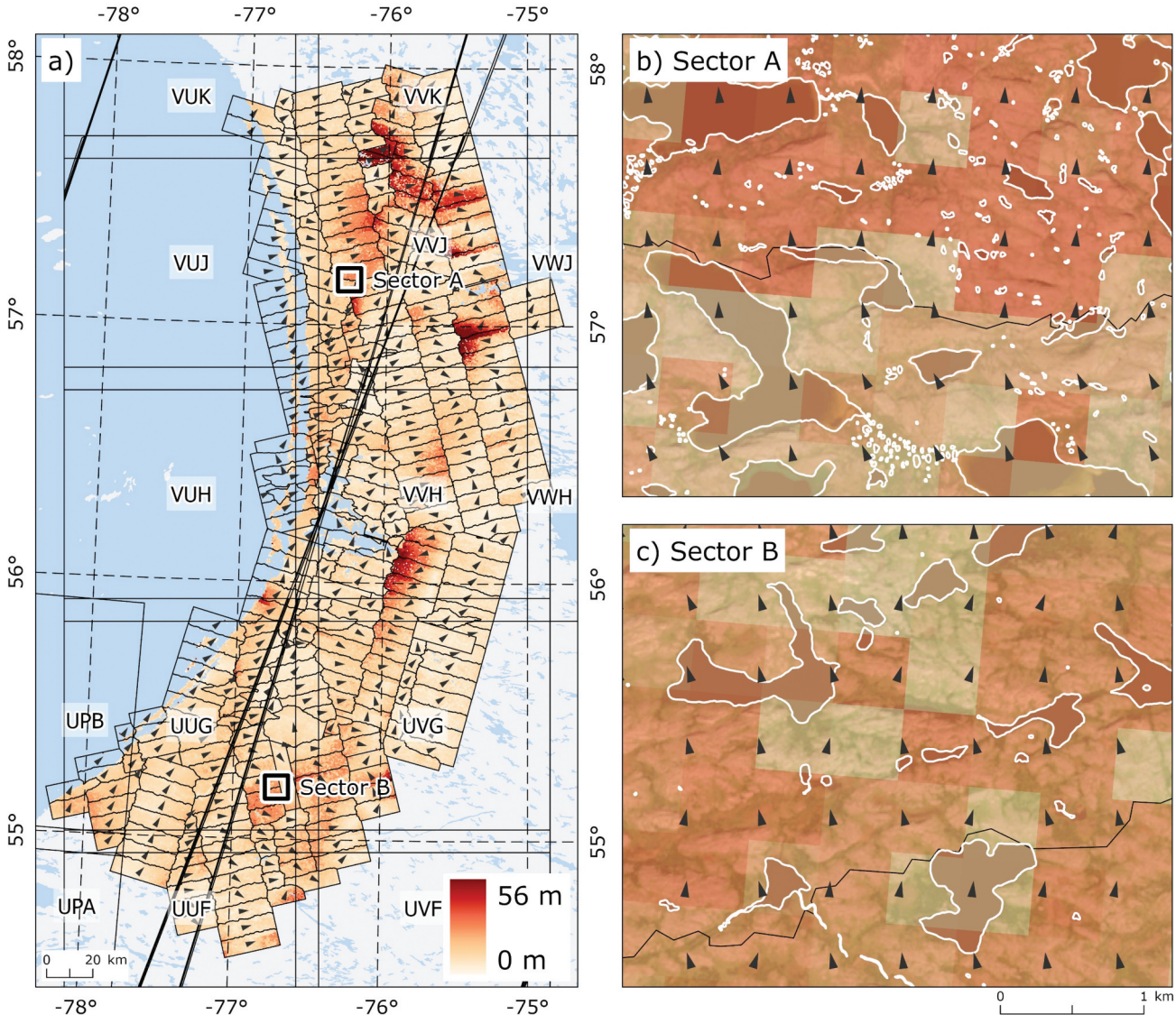

Figure 4. S2 to PS-D local co-registration: a) computed and corrected absolute mean shifts (m) between S2 (outlined black granules) and PS-D scenes (small black seamlines) and mean shift direction (black arrow); b) detail of sector A and c) sector B, showing the water surface extent output from HLWATER (white boundaries) and absolute mean shifts (m) over the S2 true color composite. The base data information is from the Government of Canada (CanVec).

co-registration. Small deviations could result in significant errors in assessing water signals, particularly for small (<0.01 km^2) and elongated water bodies.

In some cases, AROSICS was unable to fully resolve the necessary adjustments. Visual inspections revealed that some problems persisted in areas requiring higher mean absolute shifts, showing that polynomial adjustments were sometimes insufficient. AROSICS assumed that large deformations between the image sets did not exist, but such deformations can indeed occur between some sensors (Scheffler et al. 2017). Consequently, we excluded all water bodies showing inconsistent co-registration in those sectors from subsequent steps. In addition, for water bodies smaller than 0.01 km^2, only the farthest pixel from the shorelines was targeted as representative of its optical properties, mitigating other localized problems.

3.4.2. Zonal and pixel-targeted reflectance retrievals

Reflectance retrieval is a challenge in irregular and especially in small water bodies. To avoid adjacency and water level change effects, we employed two strategies for retrieving reflectance from the S2

imagery using the HLWATER product as reference. For larger water bodies, we collected zonal statistics considering several pixels, but for smaller water bodies, we selected single pixels (Figure 5).

For water bodies larger than 0.01 km^2, we first calculated the Euclidian distance from the shoreline to the center of each water body to identify the farthest point possible from the shore (e.g. maximum Euclidian distance within water body margins). These points served as the basis for calculating a circular buffer with 20% of the area of the respective water body (Freitas et al. 2019). Whenever the circular buffer extended beyond the water body boundaries, it was clipped to its extent. These buffers were then used to compute zonal statistics, namely, to obtain the median value of all pixels, based on the S2 imagery for each water body (Figure 5).

For water bodies smaller than 0.01 km^2, we implemented a more robust strategy due to the challenges posed by their size for reflectance retrievals at the S2 spatial resolution, particularly for those smaller than 350 m^2 (Freitas et al. 2019). This strategy considered the morphometric complexity of the water bodies, as well as the positioning of the S2 pixels in relation to the shorelines, as follows:

(1) HLWATER water fraction calculation at S2 10-m grid resolution (sub-pixel classification).
(2) Assessment of the S2 pixels fully within HLWATER water body boundaries.
(3) Calculation of the Euclidian distance of each pixel center to the shorelines.
(4) Selection of the farthest pixel from the shoreline to represent the optical properties of the water body (Figure 5).

Steps 3 and 4 were implemented to accommodate and avoid adjacency impacts. For the small water bodies, the distance of the farthest pixel from the shoreline ranged from a minimum of 2 m to a maximum of 46 m, with a mean, median, and standard deviation of 11, 8, and 7 m, respectively.

Due to the high abundance of water bodies smaller than 0.01 km^2 and the Pareto-like frequency distribution observed for western Nunavik (Freitas et al. 2024), some

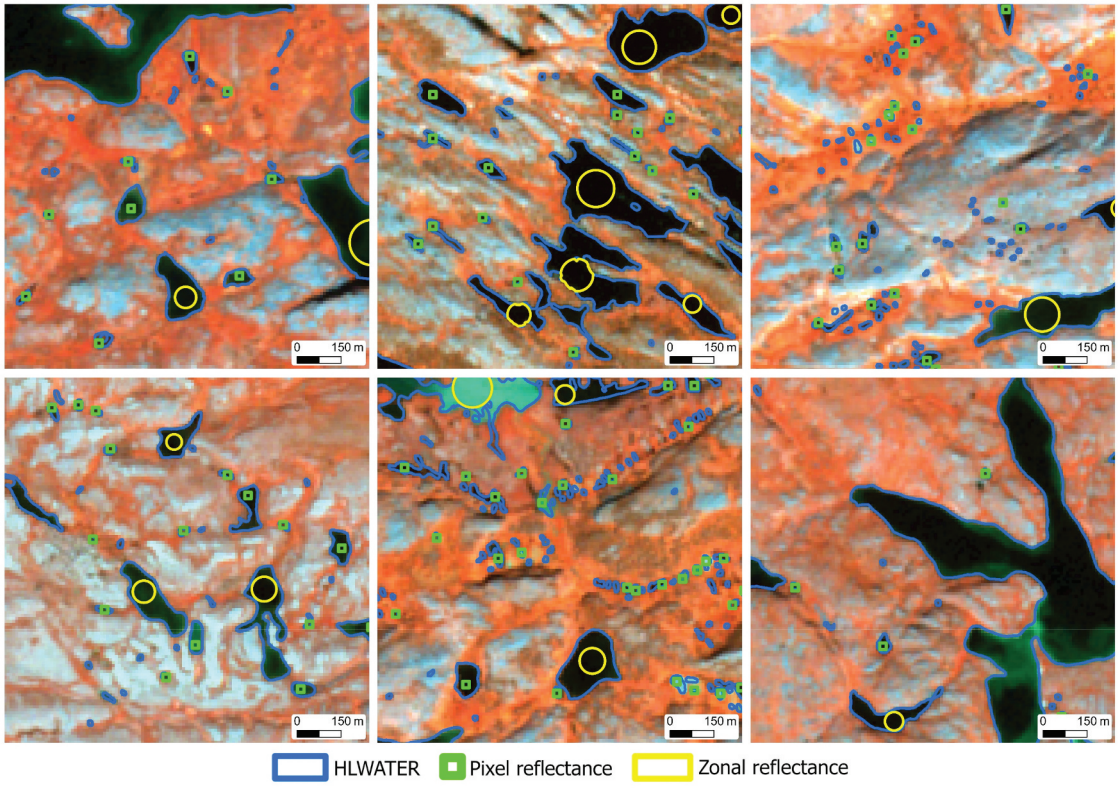

Figure 5. Automated selection of HLWATER water body areas for reflectance retrieval over the S2 scene (false color composite: *R* – NIR; G – Green; B – blue). In water bodies larger than 0.01 km^2, reflectance was calculated (median values) from within a circle encompassing 20% of the lake area (yellow polygons). In water bodies smaller than 0.01 km^2, reflectance was obtained from a single pixel (green squares). The small blue polygons without green pixels inside were excluded from the analysis (inability to accommodate full S2 pixels).

water bodies did not meet the size and shape criteria to accommodate at least one full S2 pixel within their boundaries and were excluded from further analysis (Figure 5). Additionally, impacts from cloud cover and punctual co-registration incompatibilities prevented the integration of all 335,281 water bodies in our HLWATER database. Ultimately, a total of 167,755 (50%) were selected. The S2 10-m pixel size, within the presented workflow, allowed us to acquire information on water bodies ranging from 0.000166 km^2 (166 m^2) to 160,244 km^2. The mean, median, and standard deviation of water body sizes were 0.035 km^2, 0.00147 km^2, and 0.700 km^2, respectively.

3.4.3. Outlier detection algorithm implementation

The use of the PS-D mosaic comprehending summer scenes between 2017 and 2019 for water body delineation (HLWATER) and the S2 from 23 August 2019 for reflectance retrieval (HLWATER-Optical) could locally create inconsistencies, for example, associated with modifications in the vegetation close to the shorelines or even drainage or drying events. To mitigate such inconsistencies, which would impact reflectance without relation to water properties, we implemented a data-driven, unsupervised kNN multivariate outlier detection algorithm, using the NIR band and the red/green ratio as references in the PyOD Python library developed by Zhao, Nasrullah, and Li (2019). The red/green ratio is commonly used for colored dissolved organic matter and water color analysis in organic-rich water bodies (Kutser et al. 2004; Toming et al. 2016), while the NIR band is used for removing or correcting sun-glint (Kay et al., 2009) and land-adjacency impacts (Paulino et al. 2022). The removal of outliers also allowed to further analyze the whole dataset focusing on the prevailing optical groups.

Although simple, kNN demonstrates state-of-the-art performance for detecting both global and dependency anomalies in data (Han et al. 2022). During the implementation, we used an outlier fraction of 0.01 (Figure 6). Increasing this threshold did not significantly alter the general band reflectance statistics and led to the removal of very turbid water bodies (e.g. whitish), limiting the accurate assessment of their optical representativity at the regional scale.

Notably, the kNN algorithm caused minor changes in mean, median, standard deviations, Interquartile Range (IQR), first and third quartiles, as well as coefficient of variations TOA reflectance across all bands, indicating that outliers were isolated cases and did not affect the main data distribution (Table 2). At the end, the kNN implementation identified 1,314 outliers, which were excluded from further analysis. This resulted in a total of 166,441 water bodies being used for the optical assessment and included in the HLWATER-Optical product.

3.5. Regional analysis

The regional analysis was conducted using 1 km^2 hexagonal grids as geographical units over the HLWATER and the HLWATER-Optical datasets. HLWATER covering 335,281 water bodies was used for the calculation of the limnicity and limnodensity. The analysis of the limnicity or the distribution of the water fraction of land surface per km^2 provided insight into the spatial extent of the main water bodies within the landscape.

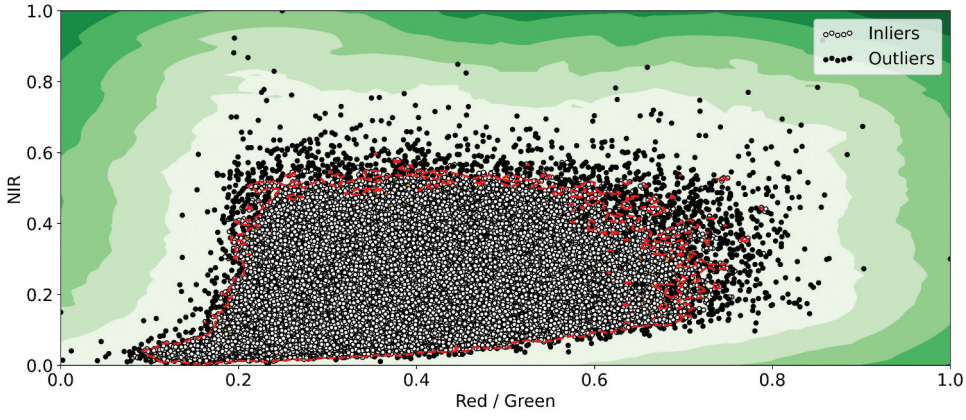

Figure 6. Identification of the outliers and inliers according to the multivariate combination of the S2 red/green ratio and NIR using the kNN outlier detection algorithm.

Table 2. S2 TOA reflectance (unitless, 0–1) band statistics considering all water bodies before and after implementing the kNN outlier detection algorithm.

Date	S2 band	Min	Max	$\bar{X}$	σ	M_d	IQR	1st Q	3rd Q	CV
Before kNN	Blue	0.071	0.199	0.080	0.009	0.078	0.006	0.076	0.081	0.110
	Green	0.044	0.219	0.056	0.013	0.053	0.010	0.049	0.059	0.237
	Red	0.023	0.224	0.040	0.015	0.036	0.013	0.031	0.044	0.381
	NIR	0.015	0.342	0.062	0.036	0.048	0.048	0.034	0.082	0.582
	Red/Green	0.528	1.026	0.709	1.137	0.683	1.267	0.631	0.740	1.604
After kNN	Blue	0.071	0.197	0.080	0.008	0.078	0.006	0.076	0.081	0.102
	Green	0.044	0.219	0.056	0.013	0.053	0.010	0.049	0.059	0.223
	Red	0.026	0.224	0.040	0.014	0.036	0.013	0.031	0.044	0.354
	NIR	0.017	0.213	0.061	0.035	0.048	0.048	0.034	0.082	0.569
	Red/Green	0.595	1.026	0.706	1.121	0.683	1.271	0.631	0.739	1.587

Table 3. HLWATER database summary statistics on the total water bodies of the regional sector, split by 0.01 km^2 size threshold and considering a minimum detection size threshold of 166 m^2.

Features	n	$\bar{X}$	σ	M_d	Total area
Water bodies > 0.01 km^2	32,232	0.17 km^2	1.59 km^2	0.03 km^2	5,582 km^2
Water bodies < 0.01 km^2	303,049	0.001 km^2	0.002 km^2	0.0004 km^2	313 km^2

The limnodensity, here defined as the density of inland water bodies per km^2, allowed us to determine the concentration of small lentic water systems across the region. For this calculation, each water body was considered as an individual entity, and their centroids were computed and used as geographic references.

The HLWATER-Optical dataset allowed the color analysis of 166,441 water bodies and estimates of "limnodiversity" as the total number of optical clusters computed by the centroid of each water body per km^2. The optical clusters were defined using K-Means clustering with optimized seed locations in ArcGIS Pro 3.2, based on the green and red bands from S2 (Khan 2012). The number of optical water body groups was set to 11, automatically determined by the optimal K based on Pseudo-F Statistics scores (e.g. similar to the Calinski and Harabasz score – ratio between-cluster dispersion and within-cluster dispersion). Thus, a 1 km^2 hexagon showing 11 different cluster members would indicate the maximum possible scale of limnodiversity of our regional sector. Finally, we graphically combined limnodensity with limnodiversity through a bivariate classification. This analysis allowed us to better analyze how small water bodies were responsible for sustaining contrasting limnodense and limnodiverse landscapes in the regional sector.

4. Results

4.1. Limnicity

The implementation of the HLWATER model allowed the classification of a total of 335,281 water bodies over the PS-D mosaic. Of these, 90% (303,049) were smaller than 0.01 km^2 and covered a total area of 313 km^2. The remaining 10% (32,232) were larger than 0.01 km^2 and covered a total area of 5,583 km^2 (Table 3).

Analyzing the limnicity at the regional scale, the majority of the 1 km^2 hexagonal grids revealed small water fractions. About 86% of the grids showed water fractions smaller than 32.1% (<6.8 = 36%; 6.8–17.4 = 30%; 17.4–32.1 = 20%). The remaining 14% had water fractions greater than this threshold (32.1–56.9 = 10%; >56.9 = 4%). Higher water fractions were observed toward the northern latitudes and farther from the coast, especially above the 56.5° N parallel and from the boundary of the sporadic permafrost zone toward the continuous permafrost zone (Figure 7).

High water fractions were usually indicative of large rivers and large lakes. An example was the Great Whale River, the largest river in eastern Hudson Bay, which has its mouth near the village of the Inuit and Cree community of Kuujjuarapik-Whapmagoostui (Bhiry et al. 2011; Owczarek et al. 2020). Cases of large hydrologically interconnected lakes included Lake Tikirartuuq and Lake Kakiattualuk, located in glacially carved granitic depressions that are typical of the vast highlands of the Canadian Shield (Leboeuf and Fournier 2015; Magnuson et al. 1997; Nitze et al. 2018).

Conversely, smaller water fractions were observed along the coast and in the lowlands of more or less open U-shaped valleys, typically filled with Quaternary

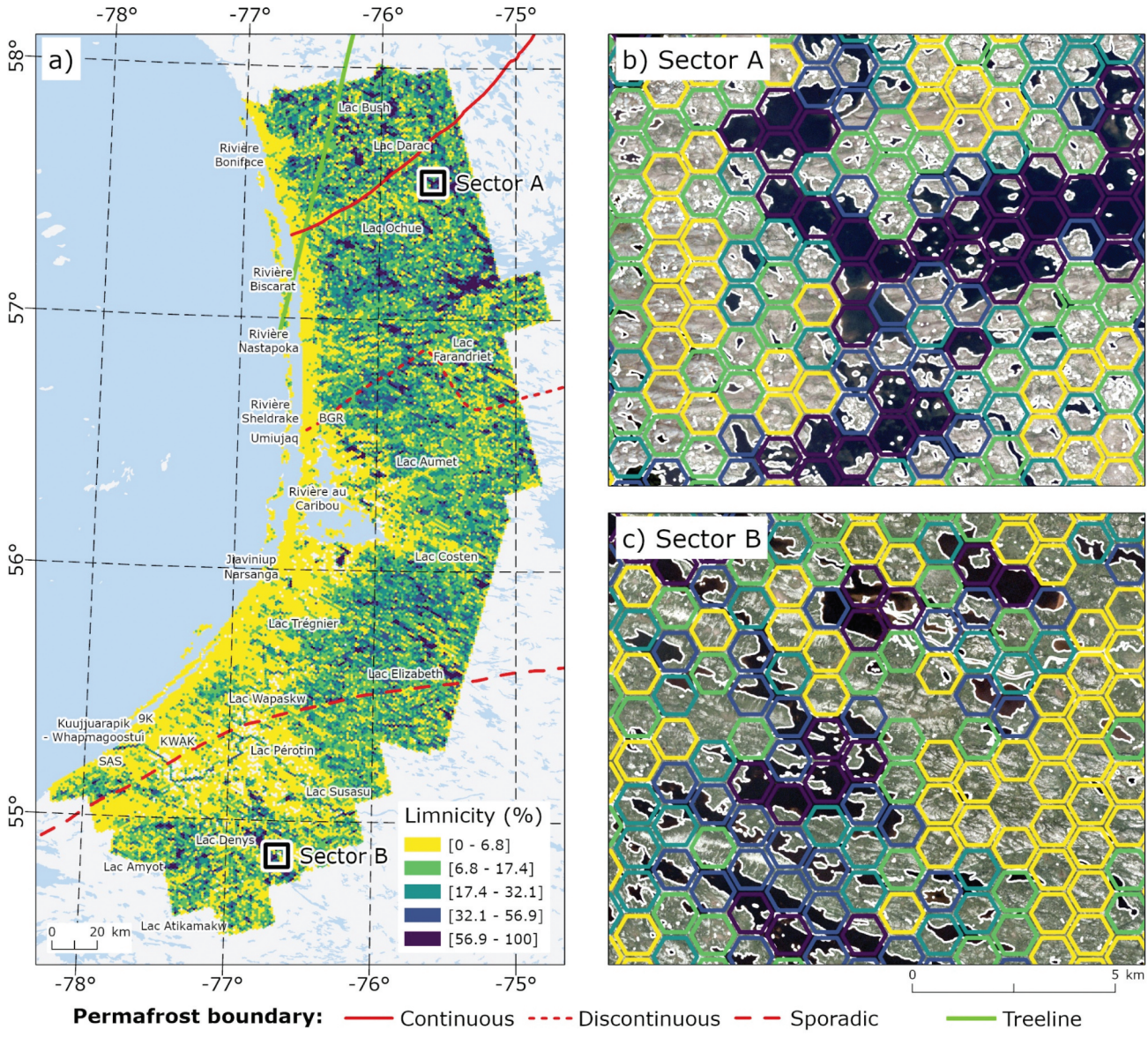

Figure 7. Limnicity (water fraction of land surface per 1 km^2) for western Nunavik: a) regional map of limnicity with permafrost zones and treeline limits, b) detail of sector A and c) sector B, showing the water surface extent output from HLWATER (white boundaries) over the PS-D mosaic (true color composite) and the hexagons with the limnicity classification. Classes were defined using the natural breaks method. The base information is from the Government of Canada (CanVec and permafrost atlas).

marine surficial deposits left by the transgression of the Tyrrell Sea at ≈7.9 ka (Bhiry et al. 2011; Leboeuf and Fournier 2015). These low water fraction grids were especially prevalent below the 56.5° N parallel, in non-permafrost sectors as well as between the discontinuous and sporadic permafrost zones.

4.2. Limnodensity

While the limnicity analysis illustrated the geographical distribution of large water bodies in the regional sector, the calculation of the limnodensity (number of water body centroids per 1 km^2) provided a different perspective by highlighting the representativity of small water bodies in certain valleys (Figure 8). Although 63% of the area revealed the existence of up to 8 water bodies per km^2, some valleys exhibited exceptionally high limnodensities, up to 229 water bodies per km^2 (Figure 8).

Along the coast and on islands, small water bodies were mainly associated with outcrops, specifically rock concavities filled with water. Further inland, often along valley lineaments perpendicular to the coast, and

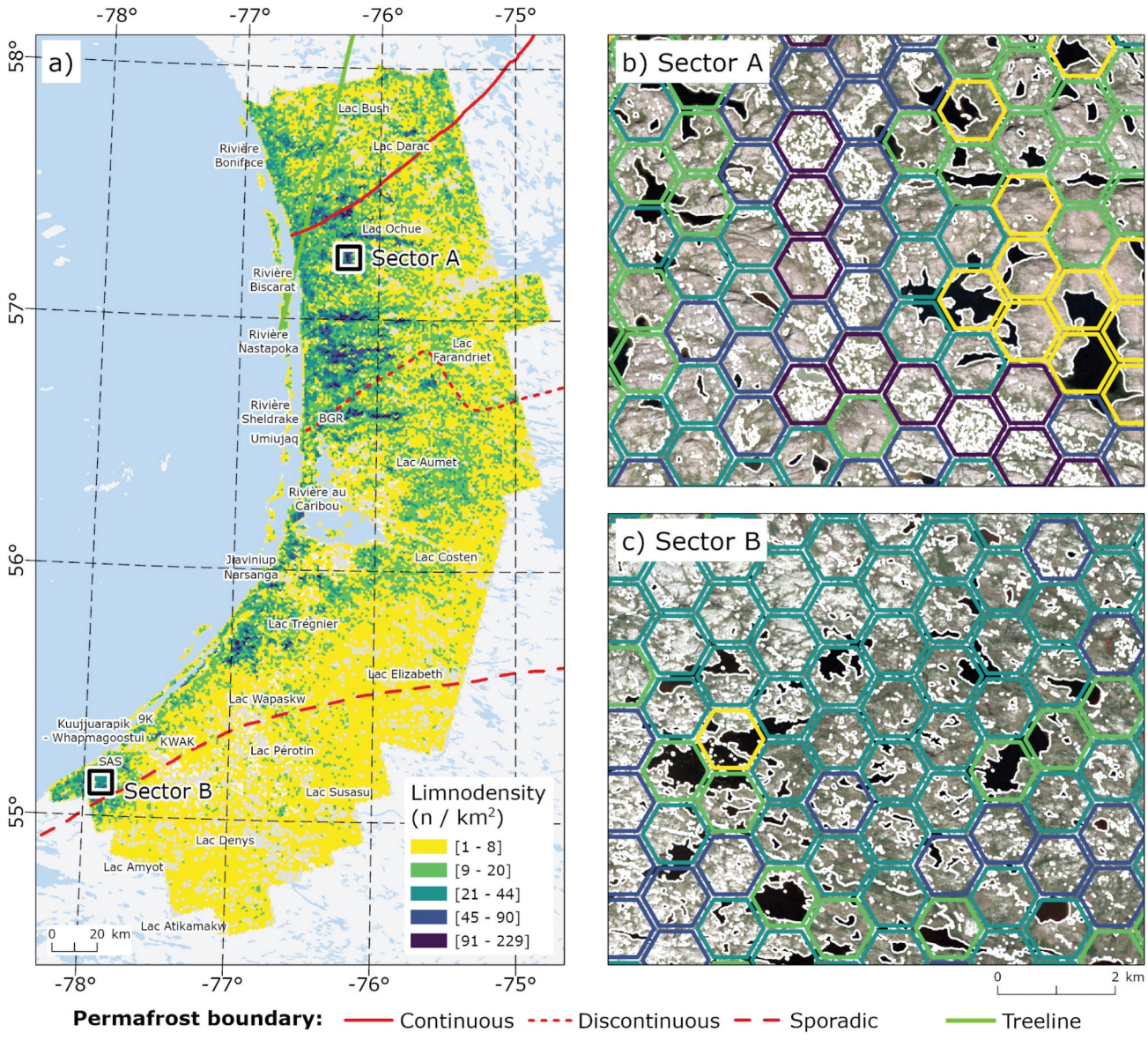

Figure 8. Limnodensity (number of water body centroids per 1 km^2) for western Nunavik: a) regional map of limnodensity with permafrost zones and treeline limits b) detail of sector A and c) sector B, showing the water surface extent output from HLWATER (white boundaries) over the PS-D mosaic (true color composite) and the hexagons with the limnodensity classification. Classes were defined using the natural breaks method. The base information is from the Government of Canada (CanVec and permafrost atlas).

especially in the discontinuous permafrost zone, small water bodies extended into glacial valleys filled with Quaternary deposits (Leboeuf and Fournier 2015). In some valleys, such as those of the Sheldrake River basin, these water bodies extended up to 45 km far from the coast. These limnodense landscapes in the discontinuous permafrost zone, or near this zone, were linked to advanced stages of permafrost degradation, characterized by the collapse, erosion, and subsidence of permafrost mounds such as palsas and lithalsas (e.g. BGR) (Allard and Seguin 1987; Coulombe, Bouchard, and Pienitz 2016; Matveev et al. 2016). In this zone, high limnodensities were observed not only perpendicular to the coast but also parallel, depending on the direction of the axis of the main valleys (Figure 8b). In the sporadic permafrost zone, high-density small water body hotspots became isolated, such as near KWAK and the Lac Trégnier.

4.3. Limnodiversity

For evaluating the limnodiversity of the regional sector, namely the existence, representativity and

geographical distribution of the different optical (color) groups, we used K-Means clustering, which automatically identified a total of 11 clusters, based on the green and red TOA reflectance of the HLWATER-Optical database (Figure 9). Generally, clusters 1 through 5 showed TOA reflectance (0–1 unitless) ranging from minimum median values of 0.048 in the green and 0.029 in the red to maximum median values of 0.063 in the green and 0.049 in the red. The IQR ranged from a minimum of 0.0015 in the green and 0.0018 in the red in cluster 1 to a maximum of 0.0033 in the green and 0.0038 in the red in cluster 5. Visual inspection indicated that these clusters predominantly referred to black (e.g. apparent dark conditions) colored water bodies (Figure 10).

From cluster 5 to cluster 6, the IQR in the green changed from 0.0033 to 0.0046 and in the red from 0.0038 to 0.0048, potentially indicating an important shift in the optical properties of the water bodies integrated in subsequent clusters. In cluster 7, the IQR continued to increase to 0.0072 in the green and 0.0074 in the red, and in cluster 8, from 0.0093 in the green to 0.0103 in the red. Visual inspections of these clusters revealed brownish colors (Figure 10). Finally, clusters 9 to 11 exhibited IQR values over 0.01 in both the green and red bands, with cluster 11 reaching a maximum IQR value of 0.02 in the red band. Visual inspection of these clusters indicated light brown (clusters 9 and 10) to white (cluster 11) colored waters (Figure 10).

In terms of regional representativeness, there was a clear predominance of black-colored water clusters (Figure 10). Specifically, cluster 1 (52,021 water bodies = 31%) and cluster 2 (34,035 water bodies = 20%) represented over half of the total water body occurrences. Combined, the black color clusters (1, 2, 4 and 5) accounted for 91% of the water bodies, followed by the brownish clusters (6, 7 and 8) with 8%, and the light brownish/whitish water bodies (9, 10 and 11) representing 1.3%. This highlighted that the latter were rare and special cases within the predominantly black-colored water body landscape.

The distribution of cluster types showed clear geographical controls and regional gradual changes. Specifically, there was a noticeable transition from widespread regional distribution dominated by black colored water bodies to clusters of brown and light-brown colored water bodies concentrated in small areas, which typically coincided with high limnodensity hotspots (Figure 11). In black-colored clusters, while cluster 1 (lowest TOA reflectance) corresponded to large water bodies located in bedrock sectors, clusters 2 through 5 corresponded to smaller water bodies moving from widespread distributions to progressive confinements to the main limnodensity hotspots. The brown (clusters 6 to 8) and light brown (clusters 9 to 11) clusters were concentrated in the high limnodensity hotspots within the sporadic to discontinuous, but mainly in the discontinuous permafrost zone. In particular, the light-brownish water bodies were primarily observed in the high limnodensity hotspots of the northern valleys within the discontinuous permafrost zone, with some isolated cases occurring in the sporadic and continuous permafrost zones.

In general, the coexistence of small water bodies with black, brown, and light-brown colors in specific glacial valleys within the discontinuous

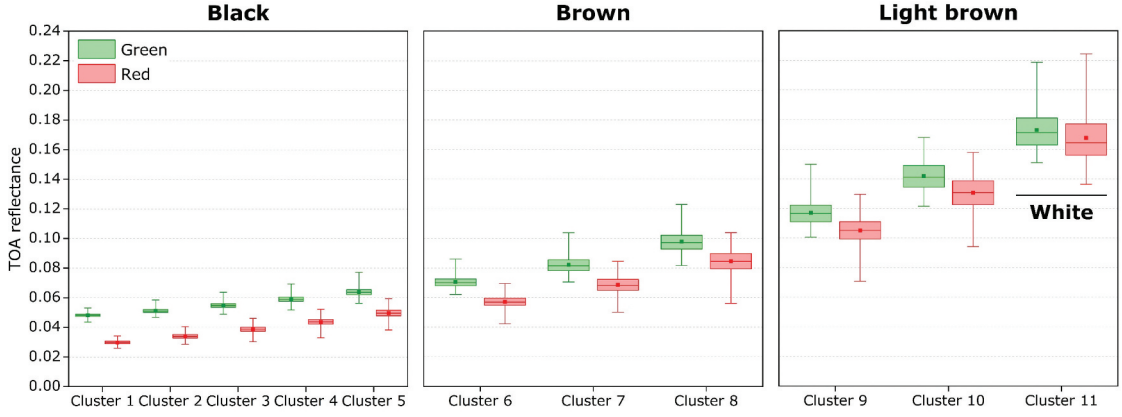

Figure 9. Green and red TOA reflectance (0–1 unitless) statistics per clusters of water bodies grouped by color regimes based on shifts in the Interquartile-Range (IQR).

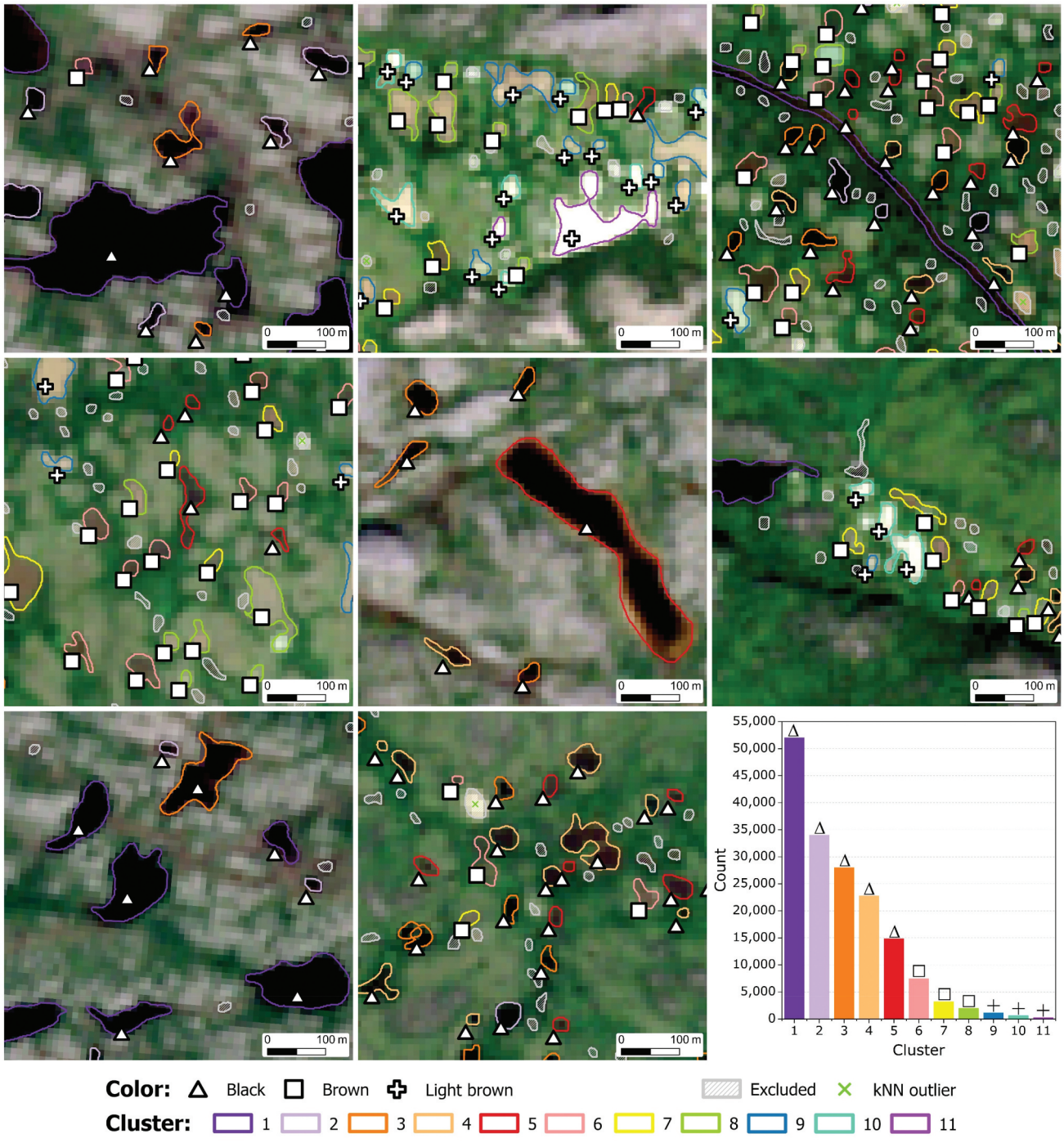

Figure 10. Cluster color frequencies in western Nunavik (bottom right) with examples for the dominant color of the water body comparison (black, brown and light brown). The S2 scene is a true color composite (RGB).

permafrost zone, characterized by thermokarst processes due to their unique climatological, geological, and hydrological settings, resulted in the most limnodiverse landscapes. In this permafrost zone, some thermokarst landscapes revealed 5 to 11 color clusters, further showcasing their limnological singularity in the regional sector (Figure 12).

4.4. Limnodensity, limnodiversity, and the case study of the Sheldrake River basin

We graphically combined limnodiversity (total number of optical clusters computed by the centroid of each water body per km^2) with limnodensity (total number of water body centroids per km^2) to evaluate

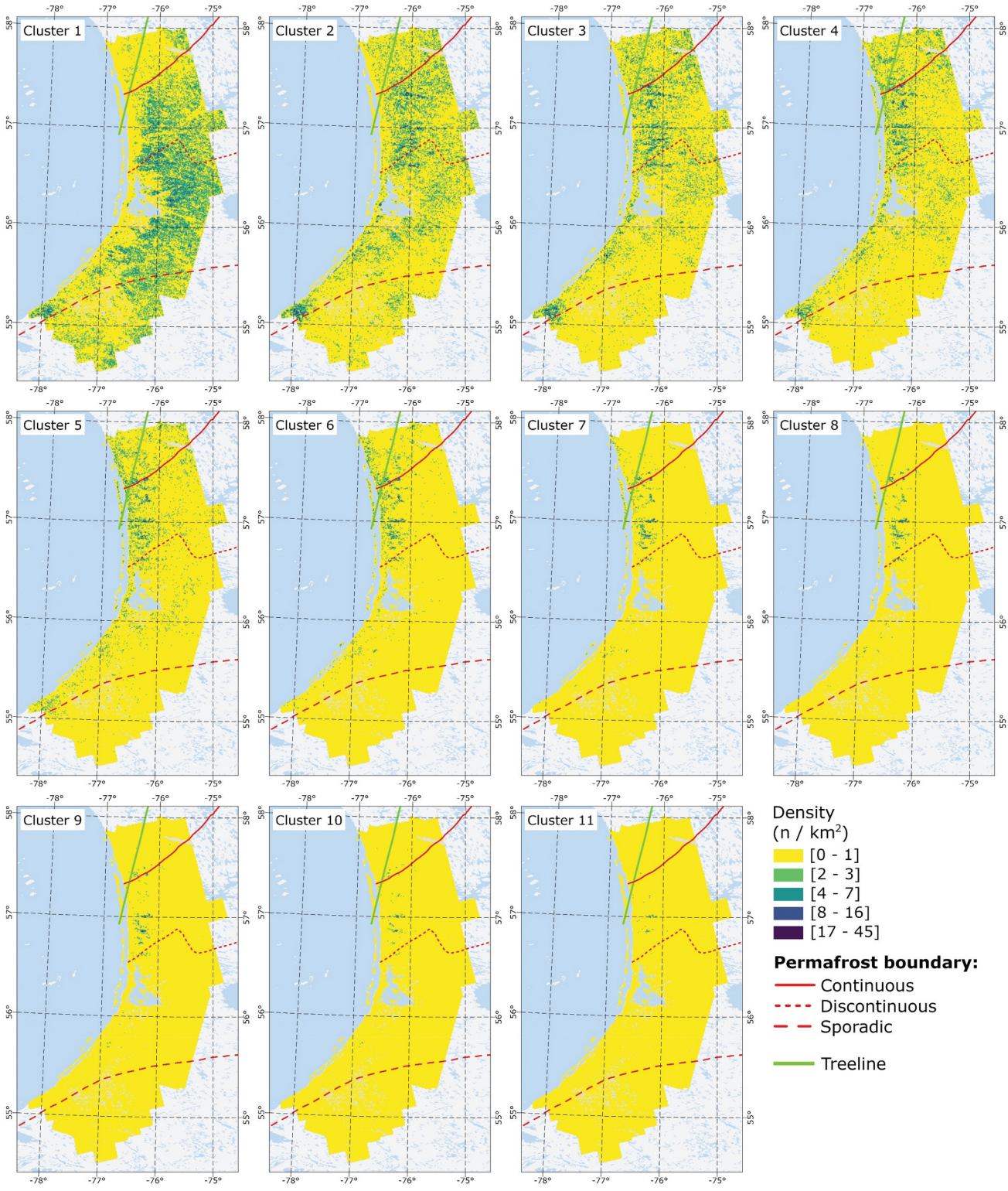

Figure 11. Geographical distribution of the water body color clusters in western Nunavik, expressed by density (total number of water bodies centroids per 1 km^2). Clusters 1, 2 3, 4, and 5 – black water bodies; clusters 6, 7, and 9 – brown water bodies; clusters 9, 10, and 11 – light brown/whitish water bodies. The base information is from the Government of Canada (CanVec and permafrost atlas).

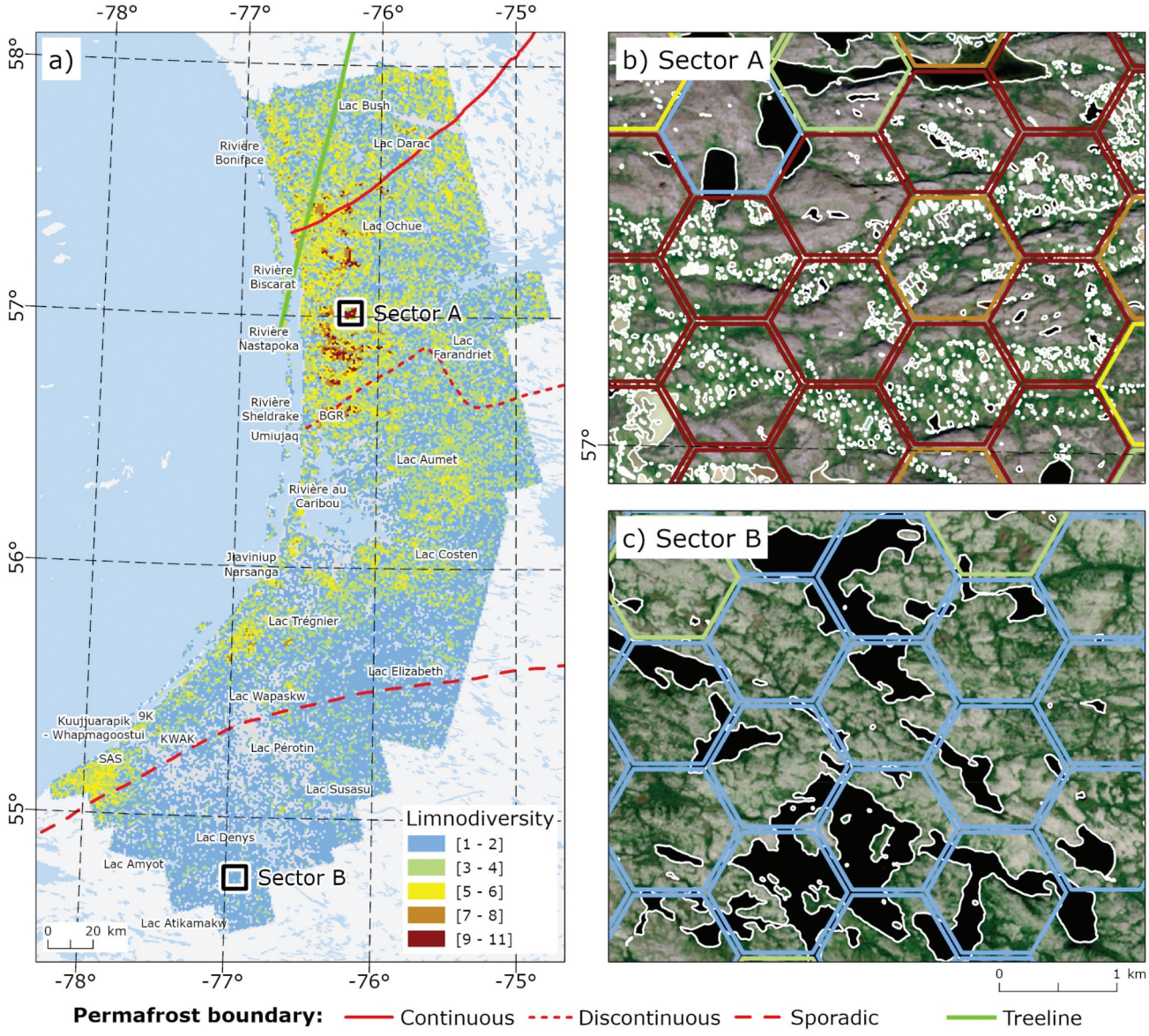

Figure 12. Limnodiversity (total number of optical clusters computed by the centroid of each water body per km^2) for western Nunavik: a) regional map of limnodiversity with permafrost zones and treeline limits; b) detail of sector A and c) sector B, showing the water surface extent output from HLWATER (white boundaries) over the S2 scene (true color RGB composite). The base information is from the Government of Canada (CanVec and permafrost atlas).

how small water bodies were responsible for sustaining high limnodense and limnodiverse landscapes (Figure 13).

Most of western Nunavik showed low limnodiversity and low limnodensity, representing 63% of the classified grids. These landscapes also showed higher limnicity. The largest water bodies, such as glacial basin lakes, had similar optical properties, with low green and red TOA reflectance and were classified within this group (Figure 13c). The second most common classification was moderate limnodiversity and moderate limnodensity, representing 11% of the classified grids. This class was typically linked to peatlands (e.g. close to SAS, Figure 13). On the opposite side of the spectrum, 2% of the grids were classified as having high limnodiversity and high limnodensity (Figure 13b). These areas were densely packed with small water bodies of varied colors. Some of these hotspots were well-defined in the landscape. Typically, they were surrounded by grids with moderate limnodiversity and high limnodensity, representing 7%.

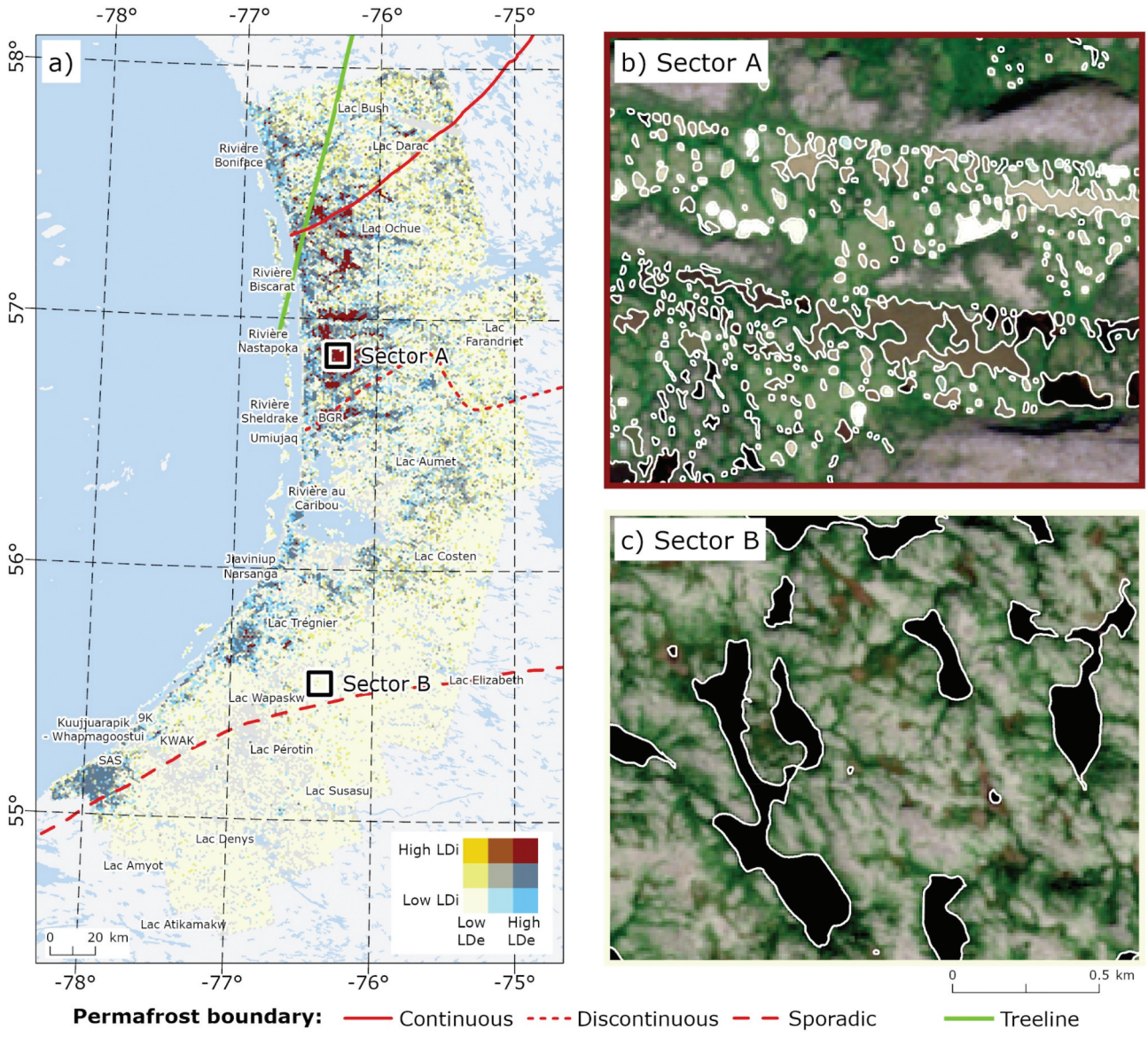

Figure 13. Bivariate classification of limnodiversity (LDi – total number of optical clusters computed by the centroid of each water body per km^2; classification intervals: 1–3; 3–6; 6–11) and limnodensity (LDe – total number of water body centroids per km^2; classification intervals: 1–10; 10–20; 20–229) for western Nunavik: a) regional map with the former bivariate classification, permafrost zones and treeline limits, b) detail of sector A and c) sector B, showing the water surface extent output from HLWATER (white boundaries) over the S2 scene (true color RGB composite). The base information is from the Government of Canada (CanVec and permafrost atlas).

This effect was attributed to the varying optical properties of the water bodies along the valleys, as observed in the Sheldrake River basin (Figures 13 and 14).

The Sheldrake River basin exemplified how the optical properties of thermokarst ponds varied along some valleys according to the characteristics and environmental conditions of the landscapes, such as the type of peatland and underlying surface deposits, impacting water optical quality properties (Figure 14). The valley showed a strong variability in green and red reflectance, related to different scattering and absorption properties of individual lentic systems. A gradient of decreasing reflectance was observed from the west, near the mouth of the Sheldrake River to the east, with an abrupt decrease at 76° 08′ W. This reduction in reflectance, along with the change from high limnodiversity and high limnodensity to moderate limnodiversity and high limnodensity, coincided

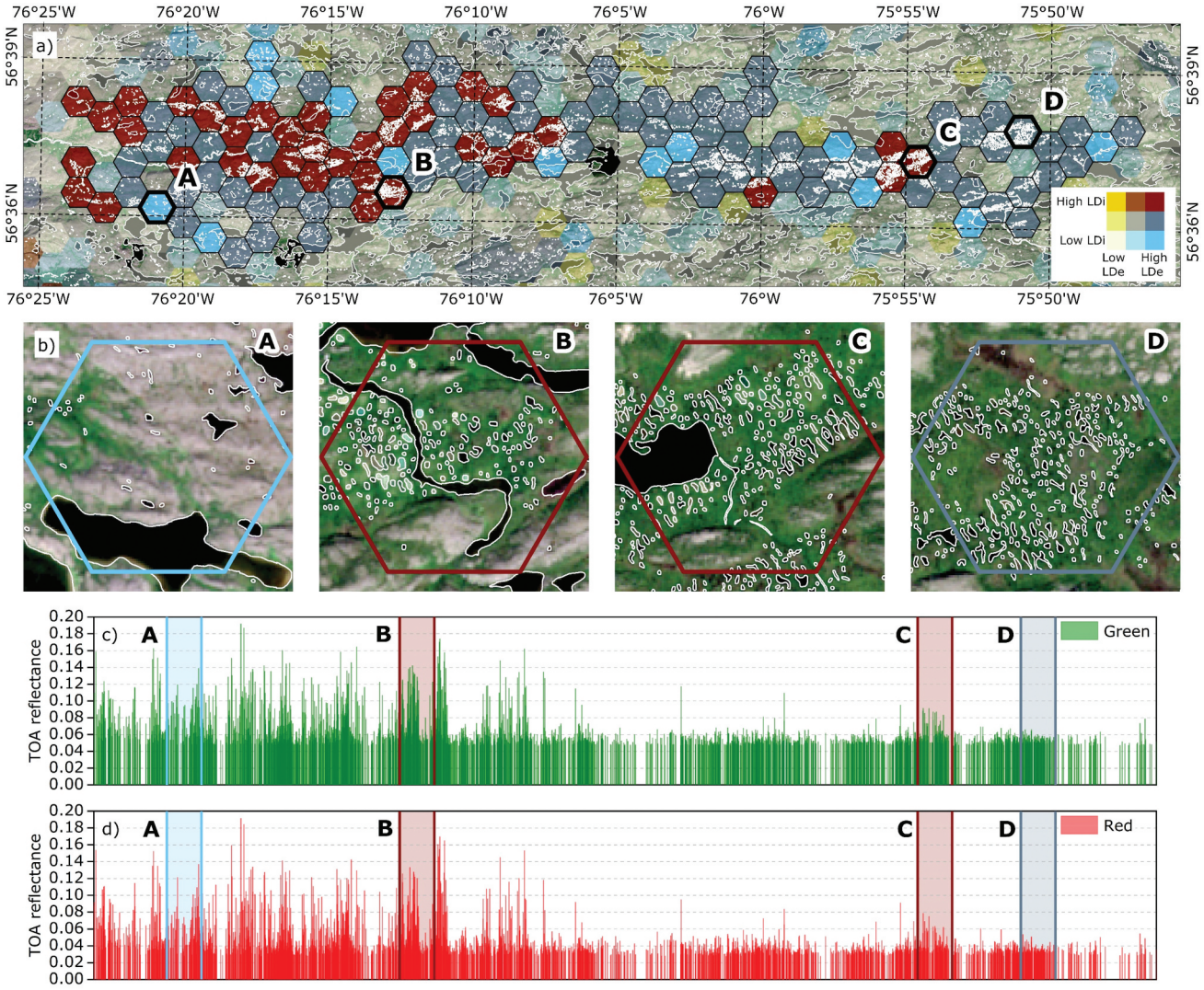

Figure 14. Bivariate classification between limnodiversity (LDi) and limnodensity (LDe) of Figure 11 for part of the Sheldrake River basin (black grids): a) general map and selected cases of 1 km^2 grids (A, B, C and B); b) detail of grid A,B, C and D; c) and d) green and red TOA reflectance for the water bodies along the valley. Grid B is the BGR study site of Figure 2. The white boundaries are the surface water extent outputs from HLWATER over the S2 scene (true color RGB composite).

with the inland limit of Tyrrell Sea paleo-marine sediments mapped by Allard and Seguin (1987) in this and surrounding basins. Accordingly, the gradient of decreasing reflectance appeared primarily controlled by the prevalence of scattering by suspended sediments in lakes with beds in fluviomarine deposits in the downstream sector of the basin (from 76° 25′ to 76° 08′ W), which eastwards had surficial organic deposits, adding absorption (Allard and Seguin 1987). In high limnodensity hotspots, small water bodies closer to the coast and along the Sheldrake River were predominantly lithalsa-formed ponds (Figure 14b) and further inland gave rise to palsa-formed ponds, marking the significant increase in reflectance and color that we found, as also supported by Allard and Seguin (1987) (Figure 14d).

5. Discussion

5.1. Synergistic remote sensing assessments of small water bodies: potential and uncertainties

We have developed a methodological framework that significantly enhanced the retrieval of spatial metrics and optical properties, specifically from small water

bodies, by integrating data and products from PS-D and S2 satellites. This framework was designed to be replicable in other regions and may be especially useful for studying small water bodies around the world. In addition, the new PlanetScope SuperDove constellation, with its improved radiometric quality (e.g. signal-to-noise ratio), offers more spectral information (coastal blue – 443 nm; blue – 490 nm; green I – 531 nm; green – 565 nm; red – 665 nm; red edge – 705 nm; NIR I – 740 nm; NIR II – 865 nm) at the same spatial resolution (≈3 m) and is now harmonized to S2 acquisitions. This enhancement leverages the volume of data (e.g. increased revisit time) and maximizes the probability of acquiring clear-skies imagery.

The development of our processing and the quality of our results were only possible by using the HLWATER model over a 3-m PS-D mosaic and further ensuring optimal co-registration between that mosaic and 10-m S2 tiles through the implementation of AROSICS (Scheffler et al. 2017). Although, in general, adjustments of less than 1/2 S2 pixels were done, in some sectors these reached up to 5 S2 pixels (e.g. 56 m). In those cases, AROSICS was not always able to fully resolve the co-registration, which implied the need to perform visual inspections along with manual removal of some small water bodies from the analyses. In addition, our reflectance extraction method for small water bodies, considering the farthest pixel from the shorelines, complemented this innovative approach.

We selected S2 imagery acquired under clear sky conditions and for the mid-warm season; therefore, the scenes were mostly free from impacts related to clouds, cloud shadows, and snow/ice covers. For removing specific water bodies impacted by these factors, the use of the S2 L2A SCL correspondent classes proved efficient. However, further careful analysis and consideration of the quality of this product is advisable when S2 scenes are acquired under varying cloud cover conditions due to frequent omission and commission errors (Qiu, Zhu, and He 2019).

The use of K-Means based on the highest Pseudo-F Statistics score allowed the definition of the number of color clusters, setting the limnodiversity scale according to the overall water body optics of the regional sector. Since K-Means is known to be particularly sensitive to outliers, implementing the kNN outlier detection algorithm beforehand was important (Ahmed, Seraj, and Mohammed 2020; Khan 2012; Zhao, Nasrullah, and Li 2019).

According to assessments using historical data on optical and geochemical lake characteristics in this region (Bouchard et al. 2011, 2014; Breton et al. 2009; Laurion et al. 2010; Watanabe et al. 2011), along with our own field work surveys (Freitas et al. 2019, 2022) and visual inspections of the color clustering, the results were not only uniform but also consistent with the optical in situ characteristics of the water bodies. The spatial validation of the HLWATER model presented by Freitas et al. (2024) showed that the water body delineation was robust even for small lakes and ponds. However, to extend the HLWATER-Optical approach to time-series analysis, further exploration of atmospheric correction models and their applicability to small and optically diverse water bodies is needed (Warren et al. 2019).

Water bodies experience water table oscillations through time, making their boundaries shrink or expand (Pekel et al. 2016; Pickens et al. 2020). Xu et al. (2022) showed that global lakes and reservoirs larger than 1 km^2 experienced rising water levels from 2003 to 2021, while large uncertainties remain in the case of small water bodies. Our selection procedures for spectral signatures avoided the problems associated with lake expansion. However, especially for small lakes and ponds, applications of our method to other periods will result in improved results if running again the HLWATER model over near-synchronous PS-D and S2 mosaics. This will be an important step for time-series analysis. In the current approach, the outlier removal technique that was applied guaranteed that water bodies drained in the short interval between the water body delineation (HLWATER) and the S2 retrievals (HLWATER-Optical) would not be included in the analysis.

5.2. Thermokarst pond genesis, representativity, and evolution: the importance of optical remote sensing retrievals

In eastern Hudson Bay, thermokarst water bodies are no older than 300 years (Bouchard et al. 2011). Accordingly, their formation and evolution can be attributed not only to the end of the Little Ice Age but also to recent climate change, where increasing air temperatures and precipitation have accelerated permafrost degradation, a trend observed in most

boreal regions (Bhiry et al. 2011; Bouchard et al. 2014; Mccrystall et al. 2021; Payette et al. 2004; Vallée and Payette 2007). As a result, monitoring these small lentic systems is essential for assessing how climate change has been shaping the hydrology, energy budgets, and biogeochemistry of the Arctic and Subarctic lowlands (Arsenault et al. 2022; Hassan et al. 2023; Holgerson and Raymond 2016; Muster et al. 2017; Taillardat et al. 2024).

Our research identified specific optical groups associated with distinct permafrost conditions. Brown and light brown water bodies were concentrated in limnodensity hotspots in the discontinuous and sporadic permafrost zones and appeared to be associated with thermokarst landscapes resulting from degrading organic-rich palsas (e.g. black to brown colored waters) or mineral-rich lithalsas (e.g. light brown to whitish colored waters), which are typical geomorphological features of these permafrost zones (Bouchard et al. 2017).

Comparison of our clustering results with the permafrost probability and the Mean Annual Ground Temperature (MAGMT – 2000–2016) map products developed by Obu et al. (2019) gave consistent results. While black-colored clusters that represent the clear water lake-type without scattering particles were observed in a wide range of conditions, an increasing concentration of brown- and light brown-colored clusters that represent water bodies dominated by scattering suspended material was found in higher permafrost probability sectors, as well as lower MAGTM sectors, supporting their thermokarst genesis (Figure 15). Along some valleys, lithalsa-formed pond landscapes near the coast gave way to palsa-formed pond ones further inland (that was more represented by hotspots with high limnodensity, but moderate limnodiversity, characterized by lower reflectance clusters). Although light brown and white ponds were rare (1.3%), they are known to be generally younger and may serve as markers and sentinels of change in thermokarst landscapes (Bouchard et al. 2014).

Shifts of decreasing reflectance along the valleys toward the east appear to be linked to the eastern boundary of Tyrrell Sea sedimentary deposits (Allard and Seguin 1987). Thus, changes in pond colors can primarily be attributed to sedimentological controls of the lake bed and lake shores along the valleys, namely the clayey silts of the Tyrrell Sea deposition near the coast in the west, followed by alluvial plain silty and sandy sediments with beds of organic debris and then fibrous peat rich in wood materials toward the east further inland (Allard and Seguin 1987). These diverse ponds in thermokarst landscapes have been subject of numerous studies (Bouchard et al. 2011, 2014; Breton et al. 2009; Folhas et al. 2020; Freitas et al. 2019; Laberge-Carignan et al. 2024; Laurion et al. 2010; Matveev et al. 2016; Wang et al. 2018; Watanabe et al. 2011), but these studies were mainly restricted to localized sites or local scales. With the spatial upscaling that the HLWATER-Optical remote sensing approach provides the data of these studies

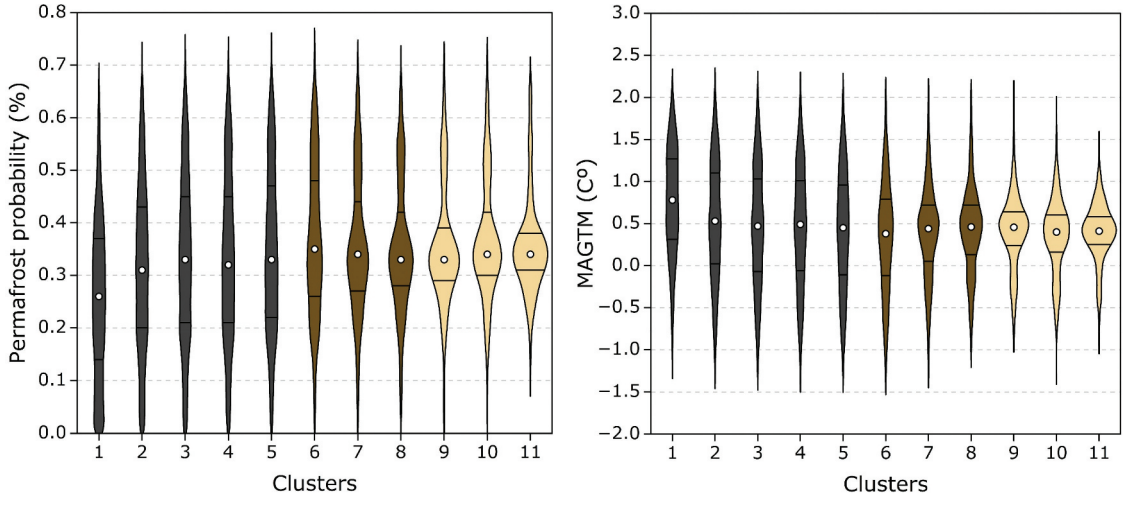

Figure 15. HLWATER-Optical color clusters, permafrost probability (%) and Mean Annual Ground Temperature (MAGMT – C°) from 2000 to 2016 by Obu et al. (2019). Clusters 1, 2 3, 4, and 5 – black water bodies; clusters 6, 7, and 9 – brown water bodies; clusters 9, 10, and 11 – light brown/whitish water bodies.

can be set in much larger spatial context (Olefeldt et al. 2021).

Considering the bivariate classification between limnodiversity and limnodensity, some classes were likely indicative of thermokarst landscapes (high limnodiversity and high limnodensity, as well as medium limnodiversity and high limnodensity). Although these represented a relatively small fraction of the total area (2% and 7%, respectively), they accounted for over a third of the total number of water bodies (113,759 $n = 34\%$), 97% of which showed sizes below 0.01 km^2 (110,926 n) (Bégin and Vincent 2017; Laurion et al. 2010). In comparison, a smaller fraction of black water bodies was set within these landscapes, representing 23%, while brown and light brown water bodies represented 60% and 92%, respectively. Certain optical clusters in the study sector included only small ponds, showing high optical diversity, while glacial erosion water bodies were included in the lowest green and red reflectance clusters as black-colored water bodies (1–5).

Black-colored water bodies showed a much higher green relative to red reflectance (Figure 7), which indicates oligotrophic conditions with only minor particle scattering in the water column. Conversely, the significant proportion of the optical signal from brown water bodies is indicative of higher particle content. For the high-scattering water bodies that all do not belong to the black-colored clusters, the reduced green relative to red reflectance could be indicative of higher colored dissolved organic matter absorption due to allochthonous terrestrial input in the green wavelength region that was also described by Kutser et al. (2004). This supports the idea that these limnoscapes that contain water bodies with a reduced green relative to red reflectance sustain important sources of carbon for greenhouse fluxes, with possible regional and global implications (Arsenault et al. 2022; Laurion et al. 2010). Light brown and white ponds should be recent and constrained by high inputs of inorganic particulate matter from thawing lithalsas (e.g. silt and clay), leading to high green and red reflectance characterized by less absorption and by this, revealing a distinct organic biogeochemistry (Bouchard et al. 2011; Coulombe, Bouchard, and Pienitz 2016).

While large water bodies are influenced by large hydrographic basin processes, small thermokarst water bodies drain small basins (Arsenault et al. 2022). These basins or catchments, although small, may be dynamic concerning terrestrial inputs into the water bodies, due to freeze-thaw cycles and permafrost disturbances, affecting their biogeochemistry at different rates in space and time (Coulombe, Bouchard, and Pienitz 2016; Grosse, Jones, and Arp 2013). These variations are not feasible to assess by costly and field-demanding in situ observations but can be efficiently tracked by an optical remote sensing approach such as HLWATER-Optical.

Wauthy et al. (2018) analyzed in situ datasets of 253 ponds across 14 regions in the circumpolar north, documenting a general "browning" in optical properties due to an increasing dominance of land-derived organic matter from permafrost thaw. In complement to that study, our approach can track optical changes in small water bodies over time, addressing how various lentic systems react to terrestrial ecosystem changes, such as permafrost thaw, terrestrialization, shrubification, and shifts in plant community structure over large areas. This will allow for an improved understanding of the small-scale but biogeochemically intense processes occurring across the circumpolar North.

6. Conclusions

We developed a remote-sensing synergistic strategy combining PS-D and S2 acquisitions for assessing the limnicity, limnodensity, and limnodiversity of a boreal forest tundra transect located in Nunavik, Subarctic Canada. The proposed indices of limnodensity and limnodiversity, together with limnicity, provided an unprecedented overview of lake and pond variability of this transition zone. Deglaciation history and geomorphological controls, together with permafrost conditions, marked the regional patterns in limnicity, limnodensity, limnodiversity, and lake color dominance that were found.

Our detailed limnological assessment was made possible by implementing the HLWATER model, which consisted of a trained Mask-RCNN deep learning model, over a 3-m spatially resolved PS-D mosaic, for generating a very high-resolution water body delineation database, and then using this product as a reference for water reflectance retrievals from the HLWATER-Optical dataset, itself derived from S2 imagery. The satellite scenes were combined through local co-registration and outlier filtering, and the

HLWATER and HLWATER-Optical datasets are now made freely available for other applications.

Our framework allowed the successful delineation and investigation of 335,281 water bodies across a latitudinal gradient from non-permafrost to the continuous permafrost zone. Of these, 90% (303,049) were smaller than 0.01 km^2, highlighting the vast number of small water bodies in the region. While limnicity largely reflected the distribution of large water bodies, limnodensity allowed us to identify the significant presence and distribution of small lakes and ponds in several glacial valleys. In the discontinuous permafrost zone, certain valleys, such as the Sheldrake River basin, exhibited a remarkably high limnodensity, with up to 229 water bodies per km^2 extending up to 45 km inland. Although small in size, these high-density water bodies are likely to play a critical role in key biogeochemical processes, such as carbon and nutrient cycling, supporting also a wide variety of benthic communities, plankton populations, and microbial activities (Abnizova et al. 2012; Bégin and Vincent 2017; Kuhn et al. 2018; Laurion et al. 2010; Matveev et al. 2016). Such a detailed geographical analysis of these water bodies for this region had not been conducted prior to this study, underscoring the novelty of our approach and the resultant datasets.

According to the HLWATER-Optical dataset of 166,441 water bodies, large water bodies in glacially carved basins generally exhibited the lowest green and red reflectance due to high water transparency by low content of scattering particles. In contrast, some valleys in the discontinuous permafrost zone were characterized by a combination of high limnodensity and high limnodiversity due to different contributions of scattering by particles and absorption by organic matter. The analyses revealed active thermokarst landscapes that are currently experiencing permafrost degradation. These are undergoing marked changes and show different stages of mineral-rich lithalsa and organic-rich palsa degradation, resulting in the proliferation of small thermokarst ponds that have diverse optical properties even across small geographic areas.

The optical lake clusters with the highest green and red reflectance indicative of white-, light brown-, and brown-colored waters due to high scattering on particles were primarily found near the coast, associated with lake beds in lithalsas and the clayey silts of the Tyrrell Sea sediment deposition. Further inland, these clusters gave way to lower reflectance waters due to lower scattering particles, associated with lake beds in palsas, alluvial plain silty, and sandy sediments with beds of organic debris and peat accumulation with fibrous wood pieces (Allard and Seguin 1987). The variability in thermokarst water bodies underscores the need to differentiate between those over organic or mineral deposits to better predict their functioning and biogeochemical roles (Arsenault et al. 2022; Kokelj and Jorgenson 2013; Taillardat et al. 2024).

The research framework and the remote sensing processing tools presented here can be extended to other regions and through time, providing insights into the genesis, role, and evolution of small water bodies, including peatland, thermokarst, and river delta water bodies. By integrating in situ observations, our approach could further enhance the spatial and temporal understanding of aquatic biogeochemical processes. These results also draw attention to the importance of continuous monitoring and advanced remote sensing techniques for an improved understanding of climate change impacts on northern landscapes and ecosystems.

Acknowledgments

This research is a contribution to the College on Polar and Extreme Environments (POLAR2E) of the University of Lisbon. We thank the Planet's Education and Research Program, through which we were able to obtain the Planet CubeSat Images for this research. The Indigenous communities of Kuujjuarapik/Whapmagoostui and Umiujaq are gratefully thanked for allowing us to conduct research on their traditional lands. In addition, we thank Dr. David Scheffler for the development of the AROSICS Python package and for providing help for its successful implementation, and members of the CEN Aquatic Geochemistry Group at Université Laval led by Raoul-Marie Couture, for offering a broader perspective on the role of thaw lakes and ponds in the Arctic and Subarctic.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the Portuguese Foundation for Science and Technology (FCT, Portuguese Polar Program – PROPOLAR) under the projects THAWPOND, by the Centre of Geographical Studies (IGOT, Universidade de Lisboa – FCT I.P. UIDB/00295/2020 and UIDP/00295/2020) and is also part of the project PERMAMERC – Mercury Biogeochemistry, Fate, and

Impact in Permafrost Thaw Ecosystems (PTDC/CTA-AMB/4744/2020). Additional support was provided from ArcticNet (NCE), Sentinel North (CFREF), NSERC and the Center for Northern Studies (CEN), Université Laval. Pedro Freitas was funded by FCT under PhD grant SFRH/BD/145278/2019.

Data availability statement

The PlanetScope images can be downloaded from https://www.planet.com/products/platform/. The HLWATER V1.0 model used to produce the surface water extent product presented here is published at https://doi.org/10.5281/zenodo.10203553. The HLWATER curated water surface extent product for western Nunavik (eastern Hudson Bay) is published at https://doi.org/10.5281/zenodo.12196313. The HLWATER-Optical dataset for the same region is published at https://doi.org/10.5281/zenodo.12197263.

ORCID

Pedro Freitas http://orcid.org/0000-0003-0752-0490
Gonçalo Vieira http://orcid.org/0000-0001-7611-3464
Diana Martins http://orcid.org/0000-0001-8940-4762
João Canário http://orcid.org/0000-0002-5190-446X
Pedro Pina http://orcid.org/0000-0002-3199-7961
Birgit Heim http://orcid.org/0000-0003-2614-9391
Bennet Juhls http://orcid.org/0000-0002-5844-6318
Raoul-Marie Couture http://orcid.org/0000-0003-4940-3372
Warwick F. Vincent http://orcid.org/0000-0001-9055-1938

References

Abnizova, A., J. Siemens, M. Langer, and J. Boike. 2012. "Small Ponds with Major Impact: The Relevance of Ponds and Lakes in Permafrost Landscapes to Carbon Dioxide Emissions." *Global Biogeochemical Cycles* 26 (2), GB2041. https://doi.org/10.1029/2011GB004237.

Ahmed, M., R. Seraj, and S. Mohammed. 2020. "The K-Means Algorithm: A Comprehensive Survey and Performance Evaluation." *Electronics* 9 (8): 1295. https://doi.org/10.3390/electronics9081295.

Allard, M., and K. Seguin. 1987. "The Holocene Evolution of Permafrost Near the Tree Line, on the Eastern Coast of Hudson Bay (Northern Quebec)." *Canadian Journal of Earth Sciences* 24 (11): 2206–2222. https://doi.org/10.1139/e87-209.

Andrews, J. T. 1968. "Postglacial Rebound in Arctic Canada: Similarity and Prediction of Uplift Curves." *Canadian Journal of Earth Sciences* 5 (1): 39–47. https://doi.org/10.1139/e68-004.

Arsenault, J., J. Talbot, L. E. Brown, J. Holden, K. Martinez-Cruz, A. Sepulveda-Jauregui, G. T. Swindles, M. Wauthy, and J.-F. Lapierre. 2022. "Biogeochemical Distinctiveness of Peatland Ponds, Thermokarst Waterbodies, and Lakes." *Geophysical Research Letters* 49 (11): e2021GL097492. https://doi.org/10.1029/2021GL097492.

Beck, I., R. Ludwig, M. Bernier, E. Lévesque, and J. Boike. 2015. "Assessing Permafrost Degradation and Land Cover Changes (1986–2009) Using Remote Sensing Data Over Umiujaq, Sub-Arctic Québec." *Permafrost and Periglacial Processes* 26 (2): 129–141. https://doi.org/10.1002/ppp.1839.

Bégin, P. N., and W. F. Vincent. 2017. "Permafrost Thaw Lakes and Ponds as Habitats for Abundant Rotifer Populations." *Arctic Science* 3 (2): 354–377. https://doi.org/10.1139/as-2016-0017.

Bhiry, N., A. Delwaide, M. Allard, Y. Bégin, L. Filion, M. Lavoie, C. Nozais, et al. 2011. "Environmental Change in the Great Whale River Region, Hudson Bay: Five Decades of Multidisciplinary Research by Centre d'études Nordiques (CEN)." *Écoscience* 18 (3): 182–203. https://doi.org/10.2980/18-3-3469.

Biskaborn, B. K., S. L. Smith, J. Noetzli, H. Matthes, G. Vieira, D. A. Streletskiy, P. Schoeneich, et al. 2019. "Permafrost is Warming at a Global Scale." *Nature Communications* 10 (1). https://doi.org/10.1038/s41467-018-08240-4.

Bouchard, F., P. Francus, R. Pienitz, and I. Laurion. 2011. "Sedimentology and Geochemistry of Thermokarst Ponds in Discontinuous Permafrost, Subarctic Quebec, Canada." *Journal of Geophysical Research* 116 (G2): G00M04. https://doi.org/10.1029/2011JG001675.

Bouchard, F., P. Francus, R. Pienitz, I. Laurion, and S. Feyte. 2014. "Subarctic Thermokarst Ponds: Investigating Recent Landscape Evolution and Sediment Dynamics in Thawed Permafrost of Northern Québec (Canada)." *Arctic, Antarct Alp Res* 46 (1): 251–271. https://doi.org/10.1657/1938-4246-46.1.251.

Bouchard, F., L. A. MacDonald, K. W. Turner, J. R. Thienpont, A. S. Medeiros, B. K. Biskaborn, J. Korosi, R. I. Hall, R. Pienitz, and B. B. Wolfe. 2017. "Paleolimnology of Thermokarst Lakes: A Window into Permafrost Landscape Evolution." *Arctic Science* 3 (2): 91–117. https://doi.org/10.1139/as-2016-0022.

Breton, J., C. Vallières, I. Laurion, and Y. Prairie. 2009. "Limnological Properties of Permafrost Thaw Ponds in Northeastern Canada." *Canadian Journal of Fisheries and Aquatic Sciences* 66 (10): 1635–1648. https://doi.org/10.1139/F09-108.

Cooley, S. W., L. C. Smith, L. Stepan, and J. Mascaro. 2017. "Tracking Dynamic Northern Surface Water Changes with High-Frequency Planet CubeSat Imagery." *Remote Sensing* 9 (12): 1306. https://doi.org/10.3390/rs9121306.

Coulombe, O., F. Bouchard, and R. Pienitz. 2016. "Coupling of Sedimentological and Limnological Dynamics in Subarctic Thermokarst Ponds in Northern Québec (Canada) on an Interannual Basis." *Sedimentary Geology* 340:15–24. https://doi.org/10.1016/j.sedgeo.2016.01.012.

Darnajoux, R., F. Lutzoni, J. Miadlikowska, and J. Bellenger. 2015. "Determination of Elemental Baseline Using Peltigeralean Lichens from Northeastern Canada (Québec): Initial Data Collection for Long Term Monitoring of the Impact of Global Climate Change on Boreal and Subarctic

Area in Canada." *Science of the Total Environment* 533:1–7. https://doi.org/10.1016/j.scitotenv.2015.06.030.

Dash, J., and B. O. Ogutu. 2016. "Recent Advances in Space-Borne Optical Remote Sensing Systems for Monitoring Global Terrestrial Ecosystems." *Progress in Physical Geography* 40 (2): 322–351. https://doi.org/10.1177/0309133316639403.

Deshpande, B. N., F. Maps, A. Matveev, and W. F. Vincent. 2017. "Oxygen Depletion in Subarctic Peatland Thaw Lakes." *Arctic Science* 3 (2): 406–428. https://doi.org/10.1139/as-2016-0048.

Folhas, D., A. C. Duarte, M. Pilote, W. F. Vincent, P. Freitas, G. Vieira, A. M. S. Silva, R. M. B. O. Duarte, and J. Canário. 2020. "Structural Characterization of Dissolved Organic Matter in Permafrost Peatland Lakes." *Water* 12 (11): 3059. https://doi.org/10.3390/w12113059.

Fortier, P., N. L. Young, J. Lemieux, M. Walvoord, and R. Fortier. 2023. "Long-Term, High-Resolution Permafrost Monitoring Reveals Coupled Energy Balance and Hydrogeologic Controls on Talik Dynamics Near Umiujaq (Nunavik, Québec, Canada)." *Water Resources Research* 59 (1): e2022WR032456. https://doi.org/10.1029/2022WR032456.

Frazier, A. E., and B. L. Hemingway. 2021. "A Technical Review of Planet Smallsat Data: Practical Considerations for Processing and Using PlanetScope Imagery." *Remote Sensing* 13 (19): 3930. https://doi.org/10.3390/rs13193930.

Freitas, P., G. Vieira, J. Canário, D. Folhas, and W. F. Vincent. 2019. "Identification of a Threshold Minimum Area for Reflectance Retrieval from Thermokarst Lakes and Ponds Using Full-Pixel Data from Sentinel-2." *Remote Sensing* 11 (6): 657. https://doi.org/10.3390/rs11060657.

Freitas, P., G. Vieira, J. Canário, W. F. Vincent, P. Pina, and C. Mora. 2024. "A Trained Mask R-CNN Model Over PlanetScope Imagery for Very-High Resolution Surface Water Mapping in Boreal Forest-Tundra." *Remote Sensing of Environment* 304:114047. https://doi.org/10.1016/j.rse.2024.114047.

Freitas, P., G. Vieira, C. Mora, J. Canário, and W. F. Vincent. 2022. "Vegetation Shadow Casts Impact Remotely Sensed Reflectance from Permafrost Thaw Ponds in the Subarctic Forest-Tundra Zone." *Environmental Earth Sciences* 81 (22). https://doi.org/10.1007/s12665-022-10640-1.

Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore. 2017. "Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone." *Remote Sensing of Environment* 202:18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Gorham, E. 1991. "Northern Peatlands: Role in the Carbon Cycle and Probable Responses to Climatic Warming." *Ecological Applications* 1 (2): 182–195. https://doi.org/10.2307/1941811.

Grosse, G., B. Jones, and C. Arp. 2013. "Thermokarst Lakes, Drainage, and Drained Basins." In *Glacial and Periglacial Geomorphology*, edited by J. Shroder, R. Giardino, and J. Harbor, 325–353. Vol. 8. Academic Press. https://doi.org/10.1016/B978-0-12-374739-6.00216-5.

Han, S., X. Hu, H. Huang, M. Jiang, and Y. Zhao. 2022. "ADBench: Anomaly Detection Benchmark." *SSRN Electronic Journal* 35:32142–32159. https://doi.org/10.2139/ssrn.4266498.

Harden, J. W., E. T. Sundquist, R. F. Stallard, and R. K. Mark. 1992. "Dynamics of Soil Carbon During Deglaciation of the Laurentide Ice Sheet." *Science (80-)* 258 (5090): 1921–1924. https://doi.org/10.1126/science.258.5090.1921.

Harris, L. I., K. Richardson, K. A. Bona, S. J. Davidson, S. A. Finkelstein, M. Garneau, J. Mclaughlin, et al. 2021. "The Essential Carbon Service Provided by Northern Peatlands." *Frontiers in Ecology and the Environment* 20 (4): 222–230. https://doi.org/10.1002/fee.2437.

Hassan, M., J. Talbot, J. Arsenault, K. Martinez-Cruz, A. Sepulveda-Jauregui, J. Hoyos-Santillan, and J. Lapierre. 2023. "Linking Dissolved Organic Matter to CO2 and CH4 Concentrations in Canadian and Chilean Peatland Pools." *Global Biogeochemical Cycles* 37 (4): e2023GB007715. https://doi.org/10.1029/2023GB007715.

Heslop, J. K., K. M. Walter Anthony, M. Winkel, A. Sepulveda-Jauregui, K. Martinez-Cruz, A. Bondurant, G. Grosse, and S. Liebner. 2020. "A Synthesis of Methane Dynamics in Thermokarst Lake Environments." *Earth-Science Reviews* 210:103365. https://doi.org/10.1016/j.earscirev.2020.103365.

Holgerson, M. A., and P. A. Raymond. 2016. "Large Contribution to Inland Water CO2 and CH4 Emissions from Very Small Ponds." *Nature Geoscience* 9 (3): 9(222–226. https://doi.org/10.1038/ngeo2654.

Juhls, B., A. Matsuoka, M. Lizotte, G. Bécu, P. P. Overduin, J. El Kassar, E. Devred, et al. 2022. "Seasonal Dynamics of Dissolved Organic Matter in the Mackenzie Delta, Canadian Arctic Waters: Implications for Ocean Colour Remote Sensing." *Remote Sensing of Environment* 283:113327. https://doi.org/10.1016/j.rse.2022.113327.

Kay, S., J. D. Hedley, and S. Lavender 2009. "Sun Glint Correction of High and Low Spatial Resolution Images of Aquatic Scenes: A Review of Methods for Visible and Near-Infrared Wavelengths." *Remote Sensing* 1 (4): 697–730. https://doi.org/10.3390/rs1040697.

Khan, F. 2012. "An Initial Seed Selection Algorithm for K-Means Clustering of Georeferenced Data to Improve Replicability of Cluster Assignments for Mapping Application." *Applied Soft Computing* 12 (11): 3698–3700. https://doi.org/10.1016/j.asoc.2012.07.021.

Kokelj, S. V., and M. T. Jorgenson. 2013. "Advances in Thermokarst Research." *Permafrost and Periglacial Processes* 24 (2): 108–119. https://doi.org/10.1002/ppp.1779.

Kuhn, M. A., L. M. Thompson, J. C. Winder, L. P. P. Braga, A. J. Tanentzap, D. Bastviken, and D. Olefeldt. 2021. "Opposing Effects of Climate and Permafrost Thaw on CH_4 and CO_2 Emissions from Northern Lakes." *AGU Advances* 2 (4): e2021AV000515. https://doi.org/10.1029/2021AV000515.

Kuhn, M., E. J. Lundin, R. Giesler, M. Johansson, and J. Karlsson. 2018. "Emissions from Thaw Ponds Largely Offset the Carbon Sink of Northern Permafrost Wetlands." *Scientific Reports* 8 (1): 1–7. https://doi.org/10.1038/s41598-018-27770-x.

Kutser, T., D. C. Pierson, K. Y. Kallio, A. Reinart, and S. Sobek. 2004. "Mapping Lake CDOM by Satellite Remote Sensing."

Remote Sensing of Environment 94 (4): 535–540. https://doi.org/10.1016/j.rse.2004.11.009.

Laberge-Carignan, A., M. Pilote, D. Mercier, L. Florence, D. Folhas, and R.-M. Couture. 2024. "Seasonal Contrasts in Dissolved Selenium Dynamics in Subarctic Thaw Lakes." *ACS Earth and Space Chemistry* 8 (7): 1359–1369. https://doi.org/10.1021/acsearthspacechem.4c00041.

Laurion, I., W. F. Vincent, S. MacIntyre, L. Retamal, C. Dupont, P. Francus, and R. Pienitz. 2010. "Variability in Greenhouse Gas Emissions from Permafrost Thaw Ponds." *Limnology & Oceanography* 55 (1): 115–133. https://doi.org/10.4319/lo.2010.55.1.0115.

Leboeuf, A., and R. A. Fournier. 2015. "A Multisensor Multiresolution Method for Mapping Vegetation Status, Surficial Deposits, and Historical Fires Over Very Large Areas in Northern Boreal Forests of Quebec, Canada." *Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8 (11): 5199–5211. https://doi.org/10.1109/JSTARS.2015.2477780.

Lehner, B., and P. Döll. 2004. "Development and Validation of a Global Database of Lakes, Reservoirs and Wetlands." *Journal of Hydrology* 296 (1–4): 1–22. https://doi.org/10.1016/j.jhydrol.2004.03.028.

Maciel, D. A., E. Márcia, L. De Moraes, C. F. Barbosa, V. S. Martins, R. F. Júnior, H. Oliveira, L. Augusto, S. De Carvalho, and F. D. L. Lobo. 2020. "Evaluating the Potential of CubeSats for Remote Sensing Reflectance Retrieval Over Inland Waters." *International Journal of Remote Sensing* 41 (7): 2807–2817. https://doi.org/10.1080/2150704X.2019.1697003.

Magnuson, J. J., K. E. Webster, R. A. Assel, C. J. Bowser, P. J. Dillon, J. G. Eaton, H. E. Evans, et al. 1997. "Potential Effects of Climate Changes on Aquatic Systems: Laurentian Great Lakes and Precambrian Shield Region." *Hydrological Process* 11 (8): 825–871. https://doi.org/10.1002/(SICI)1099-1085(19970630)11:8<825:AID-HYP509>3.0.CO;2-G.

Matveev, A., I. Laurion, B. N. Deshpande, N. Bhiry, and W. F. Vincent. 2016. "High Methane Emissions from Thermokarst Lakes in Subarctic Peatlands." *Limnology & Oceanography* 61 (S1): S150–S164. https://doi.org/10.1002/lno.10311.

Matveev, A., I. Laurion, and W. F. Vincent. 2019. "Winter Accumulation of Methane and Its Variable Timing of Release from Thermokarst Lakes in Subarctic Peatlands." *Journal of Geophysical Research Biogeosciences* 124 (11): 3521–3535. https://doi.org/10.1029/2019JG005078.

Mccrystall, M. R., J. Stroeve, M. Serreze, B. C. Forbes, and J. A. Screen. 2021. "New Climate Models Reveal Faster and Larger Increases in Arctic Precipitation Than Previously Projected." *Nature Communications* 12 (1): 6765. https://doi.org/10.1038/s41467-021-27031-y.

Mullen, A. L., J. D. Watts, B. M. Rogers, M. L. Carroll, C. D. Elder, J. Noomah, Z. Williams, et al. 2023. "Using High-Resolution Satellite Imagery and Deep Learning to Track Dynamic Seasonality in Small Water Bodies." *Geophysical Research Letters* 50 (7): e2022GL102327. https://doi.org/10.1029/2022GL102327.

Muster, S., K. Roth, M. Langer, S. Lange, F. Cresto Aleina, A. Bartsch, A. Morgenstern, et al. 2017. "PeRL: A Circum-Arctic Permafrost Region Pond and Lake Database." *Earth System Science Data* 9 (1): 317–348. https://doi.org/10.5194/essd-9-317-2017.

Negandhi, K., I. Laurion, M. J. Whiticar, P. E. Galand, X. Xu, C. Lovejoy, and V. Shah. 2013. "Small Thaw Ponds: An Unaccounted Source of Methane in the Canadian High Arctic." *PLoS One* 8 (11): e78204. https://doi.org/10.1371/journal.pone.0078204.

Nitze, I., G. Grosse, B. M. Jones, V. E. Romanovsky, and J. Boike. 2018. "Remote Sensing Quantifies Widespread Abundance of Permafrost Region Disturbances Across the Arctic and Subarctic." *Nature Communications* 9 (1): 5423. https://doi.org/10.1038/s41467-018-07663-3.

Nitze, I., K. Heidler, S. Barth, and G. Grosse. 2021. "Developing and Testing a Deep Learning Approach for Mapping Retrogressive Thaw Slumps." *Remote Sensing* 13 (21): 4294. https://doi.org/10.3390/rs13214294.

Obu, J., S. Westermann, A. Bartsch, N. Berdnikov, H. H. Christiansen, A. Dashtseren, R. Delaloye, et al. 2019. "Northern Hemisphere Permafrost Map Based on TTOP Modelling for 2000–2016 at 1 Km2 Scale." *Earth-Science Reviews* 193:299–316. https://doi.org/10.1016/j.earscirev.2019.04.023.

Olefeldt, D., S. Goswami, G. Grosse, D. Hayes, G. Hugelius, P. Kuhry, A. D. Mcguire, et al. 2016. "Circumpolar Distribution and Carbon Storage of Thermokarst Landscapes." *Nature Communications* 7 (1): 13043. https://doi.org/10.1038/ncomms13043.

Olefeldt, D., M. Hovemyr, M. A. Kuhn, D. Bastviken, T. J. Bohn, J. Connolly, P. Crill, et al. 2021. "The Boreal–Arctic Wetland and Lake Dataset (BAWLD)." *Earth System Science Data* 13 (11): 5127–5149. https://doi.org/10.5194/essd-13-5127-2021.

Owczarek, P., M. Opa, S. Boudreau, P. Lajeunesse, and Ł. Stachnik. 2020. "Re-Activation of Landslide in Sub-Arctic Areas Due to Extreme Rainfall and Discharge Events (The Mouth of the Great Whale River, Nunavik, Canada)." *Science of the Total Environment* 744:140991. https://doi.org/10.1016/j.scitotenv.2020.140991.

Pahlevan, N., B. Smith, K. Alikas, J. Anstee, C. Barbosa, C. Binding, M. Bresciani, et al. 2022. "Simultaneous Retrieval of Selected Optical Water Quality Indicators from Landsat-8, Sentinel-2, and Sentinel-3." *Remote Sensing of Environment* 270:112860. https://doi.org/10.1016/j.rse.2021.112860.

Paulino, R. S., V. S. Martins, E. M. L. M. Novo, C. C. F. Barbosa, L. A. S. De Carvalho, and F. N. Begliomini. 2022. "Assessment of Adjacency Correction Over Inland Waters Using Sentinel-2 MSI Images." *Remote Sensing* 14 (8): 1829. https://doi.org/10.3390/rs14081829.

Payette, S., A. Delwaide, M. Caccianiga, and M. Beauchemin. 2004. "Accelerated Thawing of Subarctic Peatland Permafrost Over the Last 50 Years." *Geophysical Research Letters* 31 (18). https://doi.org/10.1029/2004GL020358.

Pekel, J. F., A. Cottam, N. Gorelick, and A. S. Belward. 2016. "High-Resolution Mapping of Global Surface Water and Its Long-Term Changes." *Nature* 540 (7633): 418–422. https://doi.org/10.1038/nature20584.

Pickens, A. H., M. C. Hansen, M. Hancher, S. V. Stehman, A. Tyukavina, P. Potapov, B. Marroquin, and Z. Sherani. 2020. "Mapping and Sampling to Characterize Global Inland Water Dynamics from 1999 to 2018 with Full Landsat Time-Series." *Remote Sensing of Environment* 243:111792. https://doi.org/10.1016/j.rse.2020.111792.

Qayyum, N., S. Ghuffar, H. M. Ahmad, A. Yousaf, and I. Shahid. 2020. "Glacial Lakes Mapping Using Multi Satellite PlanetScope Imagery and Deep Learning." *ISPRS International Journal of Geo-Information* 9 (10): 560. https://doi.org/10.3390/ijgi9100560.

Qiu, S., Z. Zhu, and B. He. 2019. "Fmask 4.0: Improved Cloud and Cloud Shadow Detection in Landsats 4–8 and Sentinel-2 Imagery." *Remote Sensing of Environment* 231:111205. https://doi.org/10.1016/j.rse.2019.05.024.

Rantanen, M., A. Y. Karpechko, A. Lipponen, K. Nordling, O. Hyvärinen, K. Ruosteenoja, T. Vihma, and A. Laaksonen. 2022. "The Arctic Has Warmed Nearly Four Times Faster Than the Globe Since 1979." *Communications Earth & Environment* 3 (1): 168. https://doi.org/10.1038/s43247-022-00498-3.

Saros, J. E., C. D. Arp, R. Couture, F. Dean, S. Macintyre, S. Mcgowan, M. Ayala-Borda, et al. 2023. "Sentinel Responses of Arctic Freshwater Systems to Climate: Linkages, Evidence, and a Roadmap for Future." *Arctic Science* 9 (2): 356–392. https://doi.org/10.1139/as-2022-0021.

Scheffler, D., A. Hollstein, H. Diedrich, K. Segl, and P. Hostert. 2017. "AROSICS: An Automated and Robust Open-Source Image Co-Registration Software for Multi-Sensor Satellite Data." *Remote Sensing* 9 (7): 676. https://doi.org/10.3390/rs9070676.

Schuur, E. A. G., A. D. McGuire, C. Schädel, G. Grosse, J. W. Harden, D. J. Hayes, G. Hugelius, et al. 2015. "Climate Change and the Permafrost Carbon Feedback." *Nature* 520 (7546): 171–179. https://doi.org/10.1038/nature14338.

Shang, Y., G. Liu, Z. Wen, P. Jacinthe, K. Song, B. Zhang, L. Lyu, S. Li, X. Wang, and X. Yu. 2021. "Remote Estimates of CDOM Using Sentinel-2 Remote Sensing Data in Reservoirs with Different Trophic States Across China." *Journal of Environmental Management* 286:112275. https://doi.org/10.1016/j.jenvman.2021.112275.

Smith, L. C., Y. Sheng, and G. M. Macdonald. 2007. "A First Pan-Arctic Assessment of the Influence of Glaciation, Permafrost, Topography and Peatlands on Northern Hemisphere Lake Distribution." *Permafrost and Periglacial Processes* 18 (2): 201–208. https://doi.org/10.1002/ppp.581.

Soriano-González, J., E. P. Urrego, S. Xavier, E. Angelats, C. Alcaraz, A. Ru, C. Vicente, and E. Tenjo. 2022. "Towards the Combination of C2RCC Processors for Improving Water Quality Retrieval in Inland and Coastal Areas." *Remote Sensing* 14 (5): 1124. https://doi.org/10.3390/rs14051124.

Taillardat, P., A. Linkhorst, C. P. Deblois, A. Prijac, L. Gandois, A. Tremblay, and M. Garneau. 2024. "A Carbon Source in a Carbon Sink: Carbon Dioxide and Methane Dynamics in Open - Water Peatland Pools." *Global Biogeochemical Cycles* 38 (4): e2023GB007909. https://doi.org/10.1029/2023GB007909.

Toming, K., T. Kutser, A. Laas, M. Sepp, B. Paavel, and T. Nõges. 2016. "First Experiences in Mapping Lake Water Quality Parameters with Sentinel-2 MSI Imagery." *Remote Sensing* 8 (8): 640. https://doi.org/10.3390/rs8080640.

Treat, C. C., A. M. Virkkala, E. Burke, L. Bruhwiler, A. Chatterjee, J. B. Fisher, J. Hashemi, et al. 2024. "Permafrost Carbon: Progress on Understanding Stocks and Fluxes Across Northern Terrestrial Ecosystems." *Journal of Geophysical Research Biogeosciences* 129 (3): e2023JG007638. https://doi.org/10.1029/2023JG007638.

Turetsky, M. R., B. W. Abbott, M. C. Jones, K. W. Anthony, D. Olefeldt, E. A. G. Schuur, G. Grosse, et al. 2020. "Carbon Release Through Abrupt Permafrost Thaw." *Nature Geoscience* 13 (2): 138–143. https://doi.org/10.1038/s41561-019-0526-0.

Vallée, S., and S. Payette. 2007. "Collapse of Permafrost Mounds Along a Subarctic River Over the Last 100 Years (Northern Québec)." *Geomorphology* 90 (1–2): 162–170. https://doi.org/10.1016/j.geomorph.2007.01.019.

Veilleux, S., N. Bhiry, and A. Decaulne. 2020. "Talus Slope Characterization in Tasiapik Valley (Subarctic Québec): Evidence of Past and Present Slope Processes." *Geomorphology* 349:106911. https://doi.org/10.1016/j.geomorph.2019.106911.

Verpoorter, C., T. Kutser, D. A. Seekell, and L. J. Tranvik. 2014. "A Global Inventory of Lakes Based on High-Resolution Satellite Imagery." *Geophysical Research Letters* 41 (18): 6396–6402. https://doi.org/10.1002/2014GL060641.

Vincent, W. F., I. Laurion, R. Pienitz, and K. M. W. Anthony. 2012. "Climate Impacts on Arctic Lakes." In *Climatic Change and Global Warming of Inland Waters: Impacts and Mitigation for Ecosystems and Societies*, edited by C. R. Goldman, M. Kumagai, and R. D. Robarts, 27–42. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118470596.

Vincent, W. F., and J. Laybourn-Parry. 2008. *Polar Lakes and Rivers*. Oxford, UK: Oxford University Press.

Vonk, J. E., S. E. Tank, and M. A. Walvoord. 2019. "Integrating Hydrology and Biogeochemistry Across Frozen Landscapes." *Nature Communications* 10 (1): 3–6. https://doi.org/10.1038/s41467-019-13361-5.

Walter Anthony, K., T. Schneider von Deimling, I. Nitze, S. Frolking, A. Emond, R. Daanen, P. Anthony, P. Lindgren, B. Jones, and G. Grosse. 2018. "21st-Century Modeled Permafrost Carbon Emissions Accelerated by Abrupt Thaw Beneath Lakes." *Nature Communications* 9 (1): 3262. https://doi.org/10.1038/s41467-018-05738-9.

Wang, L., M. Jolivel, P. Marzahn, M. Bernier, and R. Ludwig. 2018. "Thermokarst Pond Dynamics in Subarctic Environment Monitoring with Radar Remote Sensing." *Permafrost and Periglac Process* 29 (4): 231–245. https://doi.org/10.1002/ppp.1986.

Warren, M. A., S. G. H. Simis, V. Martinez-Vicente, K. Poser, M. Bresciani, K. Alikas, E. Spyrakos, C. Giardino, and A. Ansper. 2019. "Assessment of Atmospheric Correction Algorithms for the Sentinel-2A MultiSpectral Imager Over Coastal and Inland Waters." *Remote Sensing of Environment* 225:267–289. https://doi.org/10.1016/j.rse.2019.03.018.

Watanabe, S., I. Laurion, K. Chokmani, R. Pienitz, and W. F. Vincent. 2011. "Optical Diversity of Thaw Ponds in Discontinuous Permafrost: A Model System for Water Color Analysis." *Journal of Geophysical Research* 116 (G2). G02003 https://doi.org/10.1029/2010JG001380.

Wauthy, M., M. Rautio, K. S. Christoffersen, L. Forsström, I. Laurion, H. L. Mariash, S. Peura, and W. F. Vincent. 2018. "Increasing Dominance of Terrigenous Organic Matter in Circumpolar Freshwaters Due to Permafrost Thaw." *Limnology and Oceanography Letters* 3 (3): 186–198. https://doi.org/10.1002/lol2.10063.

Webb, E. E., A. K. Liljedahl, J. A. Cordeiro, M. M. Loranty, C. Witharana, and J. W. Lichstein. 2022. "Permafrost Thaw Drives Surface Water Decline Across Lake-Rich Regions of the Arctic." *Nature Climate Change* 12 (9): 841–846. https://doi.org/10.1038/s41558-022-01455-w.

Wik, M., R. K. Varner, K. W. Anthony, S. Macintyre, and D. Bastviken. 2016. "Climate-Sensitive Northern Lakes and Ponds are Critical Components of Methane Release." *Nature Geoscience* 9 (2): 99–105. https://doi.org/10.1038/NGEO2578.

Xu, N., Y. Ma, Z. Wei, C. Huang, G. Li, H. Zheng, and X. H. Wang. 2022. "Satellite Observed Recent Rising Water Levels of Global Lakes and Reservoirs." *Environmental Research Letters* 17 (7): 074013. https://doi.org/10.1088/1748-9326/ac78f8.

Zandt, M. H., S. Liebner, and C. U. Welte. 2020. "Roles of Thermokarst Lakes in a Warming World." *Trends Microbiol* 28 (9): 769–779. https://doi.org/10.1016/j.tim.2020.04.002.

Zeng, F., C. Song, Z. Cao, K. Xue, S. Lu, T. Chen, and K. Liu. 2023. "Monitoring Inland Water via Sentinel Satellite Constellation: A Review and Perspective." *ISPRS Journal of Photogrammetry & Remote Sensing* 204:340–361. https://doi.org/10.1016/j.isprsjprs.2023.09.011.

Zhang, H., and M. Väliranta. 2024. "To Better Detect Drivers of Peatland Carbon Accumulation Rates and Patterns." *Environmental Research Letters* 19 (4): 041004. https://doi.org/10.1088/1748-9326/ad33d6.

Zhao, Y., Z. Nasrullah, and Z. Li. 2019. "PyOD: A Python Toolbox for Scalable Outlier Detection." *Journal of Machine Learning Research* 20 (96): 1–7. http://jmlr.org/papers/v20/19-011.html.