

Article

# Identification of a Threshold Minimum Area for Reflectance Retrieval from Thermokarst Lakes and Ponds Using Full-Pixel Data from Sentinel-2

Pedro Freitas <sup>1,\*</sup> , Gonçalo Vieira <sup>1,2,\*</sup> , João Canário <sup>2,3</sup>, Diogo Folhas <sup>3</sup> and Warwick F. Vincent <sup>2</sup>

<sup>1</sup> Centro de Estudos Geográficos, Instituto de Geografia e Ordenamento do Território (CEG/IGOT), Universidade de Lisboa, 1600-276 Lisboa, Portugal

<sup>2</sup> Centre d'Études Nordiques (CEN), Département de Biologie, Takuvik Joint International Laboratory, Université Laval, Québec, QC G1V 0A6, Canada; joao.canario@tecnico.ulisboa.pt (J.C.); warwick.vincent@bio.ulaval.ca (W.F.V.)

<sup>3</sup> Centro de Química Estrutural, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal; diogo.folhas@tecnico.ulisboa.pt

\* Correspondence: pedro-freitas@campus.ul.pt (P.F.); vieira@campus.ul.pt (G.V.); Tel.: +351-962168892 (P.F.); +351-930600381 (G.V.)

Received: 3 February 2019; Accepted: 15 March 2019; Published: 18 March 2019



**Abstract:** Thermokarst waterbodies caused by permafrost thawing and degradation are ubiquitous in many subarctic and Arctic regions. They are globally important components of the biogeochemical carbon cycle and have potential feedback effects on climate. These northern waters are mostly small lakes and ponds, and although they may be mapped using very high-resolution satellites or aerial photography, these approaches are generally not suitable for monitoring purposes, due to the cost and limited availability of such images. In this study we evaluated the potential use of widely available high-resolution imagery from Sentinel-2 (S2) for the characterization of the spectral reflectance of thermokarst lakes and ponds. Specifically, we aimed to define the minimum lake area that could be reliably imaged, and to identify challenges and solutions for remote sensing of such waters in the future. The study was conducted in subarctic Canada, in the vicinity of Whapmagoostui-Kuujuarapik (Nunavik, Québec), an area in the sporadic permafrost zone with numerous thermokarst waterbodies that vary greatly in size. Ground truthing lake reflectance data were collected using an Unmanned Aerial System (UAS) fitted with a multispectral camera that collected images at 13 cm resolution. The results were compared with reflectance from Sentinel-2 images, and the effect of lake area on the reflectance response was assessed. Our results show that Sentinel-2 imagery was suitable for waterbodies larger than 350 m<sup>2</sup> once their boundaries were defined, which in the two test sites would allow monitoring from 11% to 30% of the waterbodies and 73% to 85% of the total lake area. Challenges for remote sensing of small lakes include the confounding effects of water reflection (both direct radiation and diffuse), wind and shadow. Given the small threshold area and frequent revisit time, Sentinel-2 provides a valuable approach towards the continuous monitoring of waterbodies, including ponds and small lakes such as those found in thermokarst landscapes. UASs provide a complementary approach for ground truthing and boundary definition.

**Keywords:** lakes; ponds; permafrost; reflectance; satellite; Sentinel-2; subarctic; Unmanned Aerial Systems

## 1. Introduction

Climate warming and associated changes in the circumpolar north are causing accelerated permafrost thaw and mobilization of organic carbon pools that have accumulated over thousands of years [1]. In many areas, the degradation of ice-rich permafrost results in subsidence of local terrain and the formation of thermokarst lakes and ponds [2]. These contain high concentrations of permafrost-derived organic carbon [3] that can be converted to carbon dioxide and methane via microbial processes [4,5]. The resultant greenhouse gas (GHG) emissions from thermokarst waters have the potential to induce positive feedbacks in the climate system, leading to increased global temperatures and accelerated thermokarst erosion [6,7]. Recent calculations show that the rate of carbon release to the atmosphere from deep thermokarst lakes could more than double by the end of this century [8], and there is a pressing need to integrate thermokarst lake and pond dynamics into Global Climate Models.

Global warming may cause the eventual disappearance of thermokarst waterbodies in the subarctic and the formation of new ones in the Arctic as the permafrost boundary continues to migrate northwards. These changes will cause modifications in GHG fluxes, and there is therefore growing interest in the analysis of lake and pond dynamics in thermokarst landscapes [5,7,9]. The waterbodies show large variations in their limnological, biogeochemical and geomorphological characteristics, even across small areas, as indicated by large differences in their optical properties [4,10,11]. GHG production in these waters appears to correlate with two optically active components: dissolved organic matter (DOM) and total suspended solids (TSS) [4,6,12]. The thawing landscapes are also experiencing an increase in shrub and tree communities, together with peat accumulation [4,6,13,14], but this increasing terrestrial and wetland productivity may be offset by the GHG emissions from thermokarst lakes [15].

Remote sensing imagery from medium, high and very high-resolution sensors could potentially be used to detect and monitor thermokarst waters [16,17], and this information would be valuable for the modelling of GHG fluxes [10]. Currently, there are enormous gaps concerning the mapping and monitoring at the global scale of small thermokarst lakes and ponds [6,16,18–20], and there is a need to find solutions using existing technologies and data [9].

The European Space Agency has identified waterbodies as a key component for Earth observation due to their linkages with variables such as surface temperature, soil moisture and methane emission rates [21]. Methane fluxes in thermokarst lakes appear to increase with decreasing lake area, which affects the density stratification of the water column and the tendency towards anoxia and methanogenesis [5]. On the other hand, the varying dimensions of these lakes result in a high degree of complexity that is difficult to capture and integrate, especially for waterbodies with areas less than 10,000 m<sup>2</sup> [16].

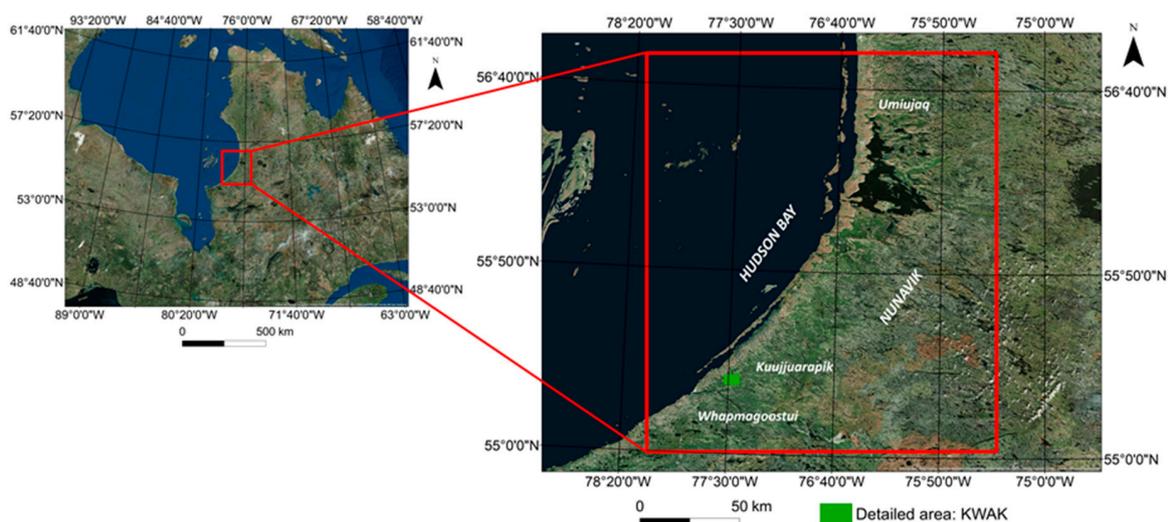
Several studies have applied satellite remote sensing to the mapping and monitoring of small waterbodies in permafrost landscapes. In an analysis of tundra lakes at three Arctic sites, imagery from Landsat-5 and MODIS was combined with higher resolution data from other satellites (TerraSAR-X and KOMPSTAT-2) and aerial photography; this showed that extrapolation of the usual Pareto distribution for lake frequency versus area would overestimate the number of small waterbodies [16]. QuickBird imagery has been used to model and forecast methane emissions in permafrost regions, and lake formation and disappearance monitored using Landsat imagery [7]. GaoFen-2 imagery has been used to monitor the response of thermokarst lakes to rapid warming and to develop algorithms for detecting lake coastlines and their dynamics [17]. Sentinel-2 (S2) has been used to improve water indexes with the application of innovative pansharpener methods [22], and to study limnological variables such as chlorophyll *a* (Chl *a*), water color, colored dissolved organic matter (CDOM) and dissolved organic carbon (DOC) [23]. In the Eastern Hudson Bay area of Canada, several studies have examined the evolution of thermokarst lakes and ponds by measuring changes in their limnological and optical properties with different types of high-resolution remote sensing imagery [4,6,10–12].

The increased availability and decreased cost of high-resolution satellite imagery are allowing for the development of new techniques and applications in remote thermokarst lake areas [9]. However, studies are still lacking on the evaluation of the use of S2 imagery for the detection and monitoring of small waterbodies with areas from tens to thousands of square meters. S2 satellites, first launched in 2015, open a new window of opportunity for the study of thermokarst lakes and ponds, especially due to improved spatial, spectral and temporal resolution. The development of Unmanned Aerial Systems (UAS) as well as the miniaturization of multispectral and hyperspectral cameras, also provide new opportunities for field data collection that can be used for ground truthing, thereby widening the possibilities for the satellite applications. However, the combination of ground-based measurements has spatiotemporal constraints inherent to each respective collection system [24].

Here we aimed to identify the minimum threshold of water body area that would allow the use of S2 imagery to characterize and monitor thermokarst lakes and ponds. We surmised that a comparison of S2 imagery with UAS multispectral imagery as ground truthing data would allow a size threshold for adequate characterization of the water body spectral characteristics from satellite imagery. Specifically, we compared very high-resolution (13 cm) data from 109 thermokarst lakes and ponds, obtained with a multispectral camera in an UAS with S2 data for optical and near-infrared wavebands.

## 2. Study Area

The study was conducted in subarctic Canada, in the valley of the Kwakwatanikapistikw River, which lies near Whapmagoostui-Kuujuarapik on the eastern coast of Hudson Bay, Canada (Figure 1). The regional climate is subarctic with average daily air temperatures of  $-20^{\circ}\text{C}$  in the coldest month (January) and  $+11^{\circ}\text{C}$  in the warmest month (August) [25]. The mean annual precipitation is c. 650 mm, 40% of it as snow. The air freezing index is twice the thaw freezing index, and the period with daily minimum temperature above  $0^{\circ}\text{C}$  is 126 days, with the growing season from mid-May to late September [6,12]. Winds prevail from the east and southeast in autumn and winter, and from the northeast in spring and summer [12]. Changes in the snow and ice regimes in the Hudson Bay have modified substantially the regional climate [25]. Mean annual air temperatures have increased by about  $2^{\circ}\text{C}$  and precipitation by 100 mm since 1950, with a marked increase since the 1990s [13].

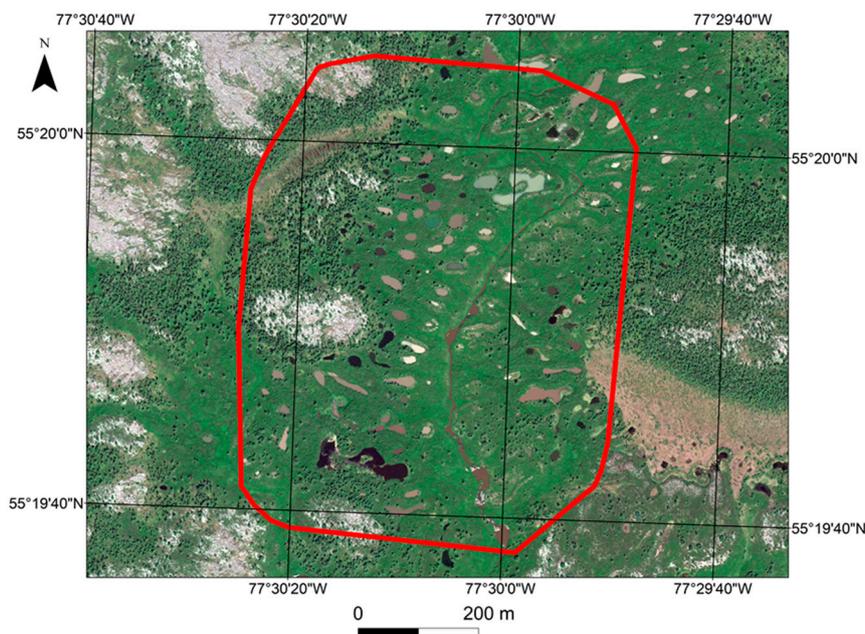


**Figure 1.** Location of the Kwakwatanikapistikw River (KWAK) study area (Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community).

The Eastern Hudson Bay region shows one of the largest glacio-isostatic rebound rates on Earth (1.1 m per century) due to the Laurentide Ice Sheet, which was responsible for the sculpting of the

terrain into low altitude and low gradient glacial landforms [12,25]. The deglaciation occurred at about 8000 yr BP, followed by the Tyrrel Sea transgression (up to 250 to 300 m asl) at c. 7900 yr BP and subsequent regression, which led to the deposition of marine clays and silt, as well as coastal sands [6,25]. Many of the terrestrial features occurring below 250 m in the coastal region are the result of landscape processes over the last thousand years. These include periglacial processes, permafrost formation, peat deposition, soil formation, aeolian processes, isolation of lacustrine basins and vegetation colonization [25]. At the end of the Little Ice Age, about 200–300 years ago, warming led to the onset of the permafrost degradation phase in the region [14,25–28], favoring the erosion of palsas and lithalsas, and generating numerous thermokarst lakes and ponds [6,12].

The Kwakwatanikapistikw (KWAK) Valley (latitude 55°19'50" N; longitude 77°30'05" W) is a long-term monitoring site of the Centre of Northern Studies (CEN) of Laval University (Figure 2) and is in the sporadic permafrost zone. It is one of the most well-studied lake areas in subarctic Canada and hence was selected for the present study. The area for ground truthing by UAS survey was 0.5 km<sup>2</sup> and included numerous thermokarst lakes and ponds that range in size from meters to tens of meters in length. The waterbodies close to rock outcrops show a darker color, are older and have a higher organic matter content (higher DOC and lower TSS), while the lakes further away from rock outcrops are more recent, more turbid and show more TSS and lower DOC [6,12].



**Figure 2.** Detailed KWAK study area with the boundary of the Unmanned Aerial System survey shown in red (true color composite from WorldView-4 – © 2018 DigitalGlobe, Inc., Longmont CO USA 80503).

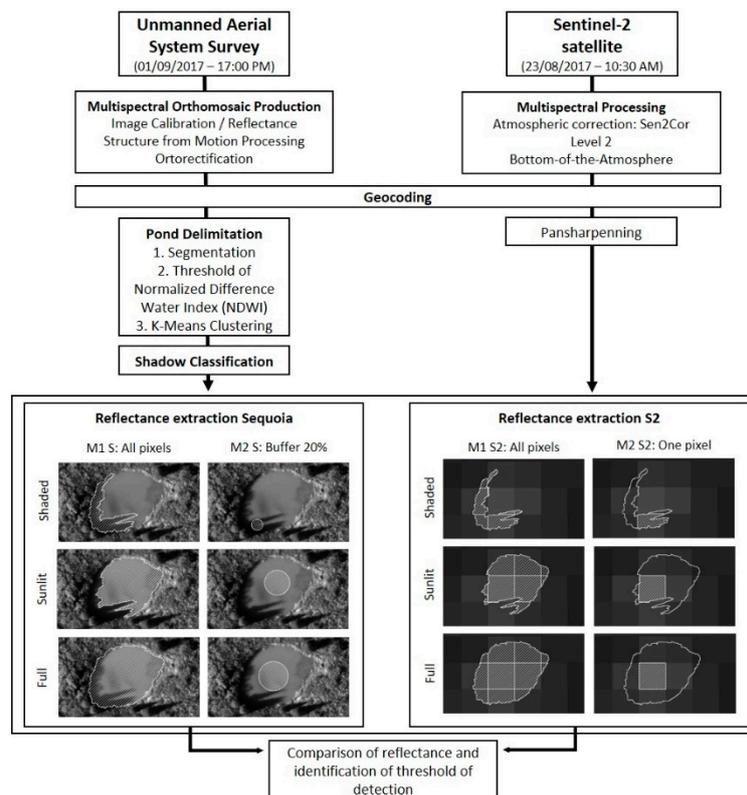
The area has experienced large landscape changes over the last 50 years, with degradation of palsas and lithalsas, expansion of thermokarst lakes and ponds, and pronounced shrubification [14,26,28] and terrestrialization, with the expansion of shrub and tree populations, along with peat accumulation [6,13]. The lakes and ponds are generally 1 to 3 m deep [4] and are surrounded by dense shrub communities (*Betula glandulosa*, *Salix sp.*, *Alnus sp.* and *Myrica gale*), sparse conifers (*Picea mariana*, *Picea glauca*, *Larix laricina*), mosses (*Sphagnum spp.*) and Gramineae (*Carex aquatilis*) [6,12,29].

### 3. Materials and Methods

#### 3.1. Workflow

In order to obtain ground truthing observational data on the spectral characteristics of the thermokarst lakes, the KWAK study area was surveyed with a UAS in 1 September 2017 (Figure 3).

A S2 scene from 23 August 2017, the closest date with cloud-free imagery, was used for comparison with satellite data and to assess their potential for monitoring spectral reflectance of the small water bodies. The UAS scene was first used for the accurate delineation of water body boundaries and for the identification of the shaded areas within the water surface. This approach allowed the assessment of spectral mixing in sunlit vs shaded water surfaces, and accurate plotting of the water boundaries for the satellite imagery. The scene was geocoded to allow for comparisons, and the S2 imagery was analyzed at 10 and 20 m resolution.



**Figure 3.** Methodological approach for the comparison of waterbody reflectance data from the UAS survey and S2 imagery.

The reflectance values from the water bodies were extracted from the UAS survey using two methods (M1 S—all pixels and M2 S—buffer), which were applied for three types of illumination conditions: shaded areas, sunlit areas and the complete water body. The reflectance values from S2 were extracted by overlapping the water body boundaries and using also two methods (M1 S2—all pixels and M2 L2—one pixel), considering also the three types of illumination conditions. The reflectance extraction procedure was applied to the 108 lakes and ponds, and the results were compared using correlation analysis to identify the area thresholds above which S2 imagery can be used.

### 3.2. Characteristics of the Remote Sensing Data

A fixed-wing senseFly eBee UAS equipped with a multispectral Parrot Sequoia camera was used for the ground truthing survey on 1 September 2017 at 17:00 (local time). The sky was cloud-free with homogeneous illumination conditions and the flight took place at an average elevation of 110 m above the ground surface, with a longitudinal overlap of 75% and a lateral overlap of 70%. Due to the terrain characteristics, with numerous lakes separated by dense shrubs that restricted walking, no ground control points were used, and spatial accuracy was determined by the internal Global Positioning System (GPS) of the UAS. A total of 218 multispectral images and a RGB color composite were obtained, resulting in a final orthomosaic resolution of 13 cm. Processing was done using

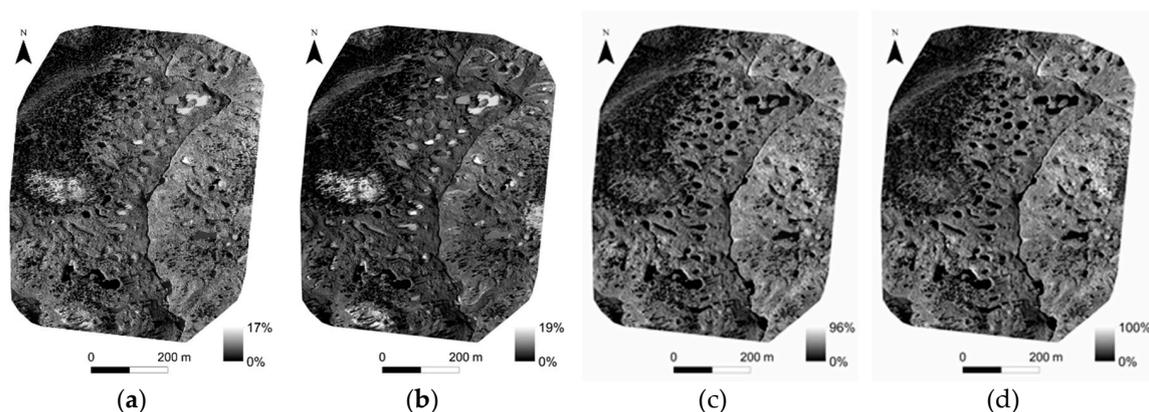
Pix4Dmapper (version 4.0). The low flight altitude greatly reduced the effects of dust, water vapor and aerosols, which obviated the need for atmospheric correction and thereby provided an advantage relative to satellite imagery [24,30].

The Parrot Sequoia camera has four spectral sensors at the green (530–570 nm), red (640–680 nm), red edge (730–740 nm) and near-infrared (770–810 nm), with a resolution of 1.2 MP. It also includes a 16 MP sensor generating true-color RGB composites. The camera has been designed for mapping and monitoring vegetation and includes a sun irradiance sensor facing upwards that allows for radiometric calibration during image collection [31]. Calibration was performed using a target for white balance before each flight and later used for the calculation of reflectance during processing.

S2 includes two satellites in sun synchronous orbits launched in 23 June 2015 and 7 March 2017. They record 13 multispectral bands with a spatial resolution from 10 to 60 m, 12 bits of spectral resolution and a revisit time of 2 days in the study area [32]. The scene from 25 August 2017 was downloaded from Earth Explorer and corrected for atmosphere effects to level 2-A using Sen2Cor (version 2.4.0). The module super-resolution in ESA's SNAP software was used for pansharpening bands 6 and 7 from 20 to 10 m following Brodu (2017) [33]. The resulting bands were named super Red-Edge for the pansharpening of B6 and super Near Infrared (NIR) for the pansharpening of B7.

### 3.3. Processing of the UAS Imagery

Very high-resolution orthomosaics were produced using Pix4Dmapper for the four spectral bands of the Sequoia camera and used for ground truthing (Figure 4). Highly homogeneous surfaces such as large water surfaces are impossible to reconstruct accurately using Structure from Motion algorithms [34]. That is because it is extremely difficult to find unique common points between neighboring images, which limits the 3-D reconstruction and mosaicking. The lakes and ponds of KWAK are, however, generally small, and individual aerial images are never fully occupied by water, which reduces the issues that large water bodies show. This was confirmed by a thorough visual inspection of the orthomosaic. Processing was conducted using positioning data from the internal GPS of the UAS, which generates reduced internal deformation of the model and a potential translation of the orthomosaic, which was calculated as 5.24 m according to the processing report. Mean square root deformations were 0.73 m (X), 0.47 m (Y) and 0.51 m (Z).



**Figure 4.** UAS survey derived reflectance for KWAK (31/08/2017, 17:00): (a) Green (530–570 nm); (b) Red (640–680 nm); (c) Red Edge (740–740 nm); (d) Near Infrared (770–810 nm).

Once analyzing the overlap between the UAS orthomosaic and the S2 imagery, we observed a northward shift of about 5 meters, probably due to the positioning accuracy of the UAS survey. Given the coarse spatial resolution of S2, we used a WorldView-4 scene with 30 cm resolution acquired in 4 July 2017, orthorectified using Rational Polynomial Coefficient (RPC) in ENVI 5.1 and the 20 m digital elevation model cdem\_33N by Natural Resources Canada [35]. After this pre-processing procedure, the WorldView-4 scene showed good concordance and overlap with the S2 imagery and

then was used to improve UAS orthomosaic geocoding. The procedure used four control points with a 1st order polynomial transformation, resulting in a Root Mean Square Error of 0.57 m and excellent visual overlap with S2 imagery was observed.

### 3.4. Delineation of the Thermokarst Lakes and Ponds

Delineation of the waterbody boundaries was essential to obtain accurate spectral values from the two remote sensing sources. We used a hybrid image classification approach [36] that was divided in three stages: i. multiresolution segmentation for the definition of the borders between water and non-water surfaces, ii. threshold definition using multispectral indices for the classification of the water bodies, and iii. non-supervised classification using K-Means Clustering for classifying the shady and sunlit areas inside the water bodies. The delineation procedure was applied using the very high-resolution mosaic obtained from the Sequoia camera.

#### 3.4.1. Segmentation

Segmentation, the most important step in object-based classifications for delineating objectives with spectral homogeneity [36], was conducted using the multiresolution algorithm of eCognition 9.0. We used the four Sequoia spectral bands and attributed a double weight to the near-infrared, which allowed for a better discrimination of the lake surface, since it shows high reflectance values in the vegetated areas that border the lakes. This allowed for the application of the edge-based function, which also requires three other parameters: scale, shape and compactness. The values for these parameters were obtained iteratively by analyzing results through visual inspection and then selecting the best resulting combination (scale: 5; shape: 0.0004; compactness: 0.5).

The weighting in the multiresolution segmentation algorithm of scale, shape and compactness, is one of its main challenges [36]. The segmentation procedure showed that the objects could be only be classified accurately with very low values, resulting in a slow iterative analysis process. The selection of the shape and compactness parameters was conducted independently using an extreme segmentation scenario with a fixed scale parameter of 5.

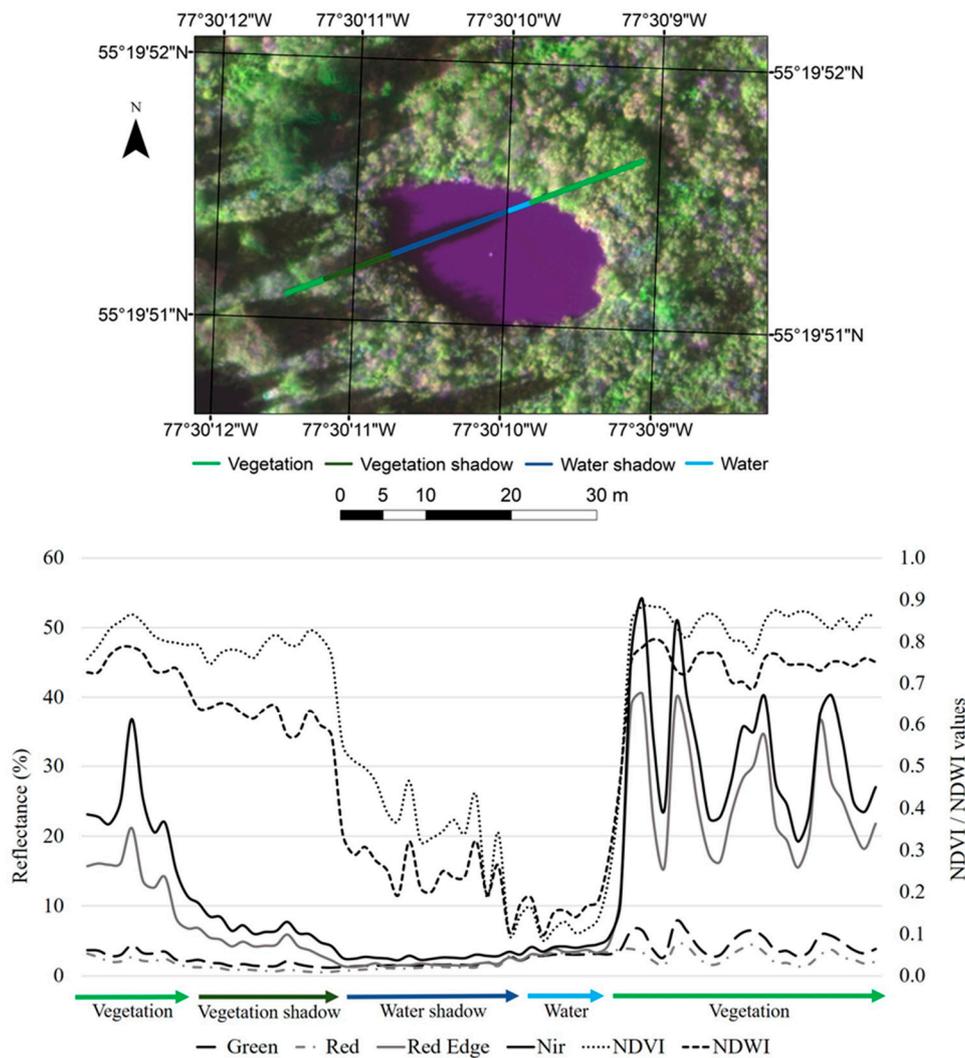
#### 3.4.2. Classification

The Normalized Difference Water Index (NDWI) and the Normalized Difference Vegetation Index (NDVI) were compared to determine which would be best for lake classification. The NDWI (1) identified open water, while NDVI (2) provided an inverse picture of lake presence by identifying the vegetation outside the lakes.

$$NDWI = \frac{Green - Near\ Infrared}{Green + Near\ Infrared} \quad (1)$$

$$NDVI = \frac{Near\ Infrared - Red}{Near\ Infrared + Red} \quad (2)$$

The use of spectral index thresholds following the segmentation reduced the salt-and-pepper effects (misclassification of sporadic pixels into a specific class area due to increased spectral variability of pixels), which is especially relevant in very high-resolution imagery [36–38]. Figure 5 shows the spectral bands, NDWI and NDVI indexes profiles along a transect located inside pond number 52, where is possible to observe that the results were best with the NDWI, which provided a good identification of the water polygons and showed less influence of shadowing artefacts.

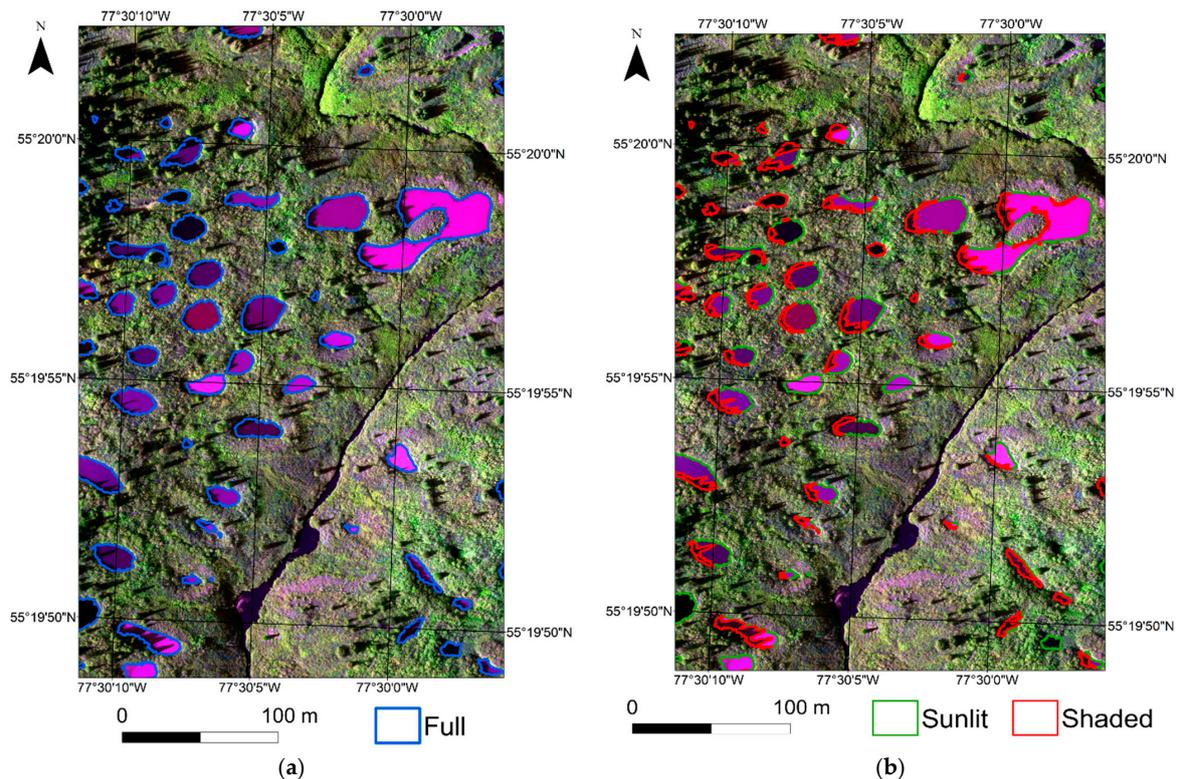


**Figure 5.** Reflectance, NDWI and NDVI variation across Pond-52 showing the effects of different surface types.

### 3.4.3. Clustering

Once the thermokarst lakes and ponds were delineated, the orthomosaicked Sequoia bands were cut using their boundaries, in order to analyze only the water surfaces. The following step was a K-Means Clustering classification using an ArcGis 10.4 tool (K-Means Clustering) for the identification of the shady areas inside the water surfaces [37]. The best clustering was obtained using a spectral resolution of 19 and a spatial resolution of 15, which allowed for a balanced classification between segmentation and accuracy in delineation of the shady and sunlit areas.

Figure 6 shows the final process of pond delimitation and the subsequent classification of shaded areas of the water bodies. The shaded areas with illumination from the southwest affected 5,948 m<sup>2</sup> of water surface, which is about 26% of the total lake and pond area. The shadows had an average size of 47 m<sup>2</sup> and a standard deviation of 44 m<sup>2</sup>, with only 12 of a total of 109 ponds showing fully sunlit surfaces.



**Figure 6.** Results of the thermokarst lake and pond classification in a part of the KWAK study area: (a) Results of the delineation using the multiresolution algorithm and NDWI threshold classification; (b) Identification of the shaded and sunlit areas following the K-Means Clustering. Background orthomosaic is a Green (red channel), Red (green channel) and Red Edge (blue channel) color composite obtained from the UAS survey.

### 3.5. Limitations and Base Assumptions

Despite the same general waveband overlaps, the Sequoia and S2 sensors show differences in central wavelength, width and radiometric depth. This is a limitation of the current approach and needs to be further assessed. We assume that the differences would have generated noise that may decrease the quality of the correlation but that would not affect the overall comparison. While S2 has a depth of 12 bits, the Sequoia has 10 bits. In this approach, we compared the Sequoia and S2 bands using the ones with closest central wavelength match, as shown in Table 1.

**Table 1.** Comparison of the characteristics of the Parrot Sequoia and Sentinel-2.

	Parrot Sequoia	Sentinel-2
Acquisition date	01/09/2017	23/08/2017
Acquisition time (local)	17:00	10:30
Spectral resolution	10 bits	12 bits
Spatial resolution	0.13 m	B3,4–10 m, B6,7–20 m
Green band	530–570 nm	543–578 nm (B3)
Red band	640–680 nm	650–680 nm (B4)
Red Edge band	730–740 nm	733–748 nm (B6)
Super Red Edge band	-	733–748 nm (Super B6)
NIR band	770–810 nm	773–793 nm (B7)
Super NIR	-	773–793 nm (Super B7)

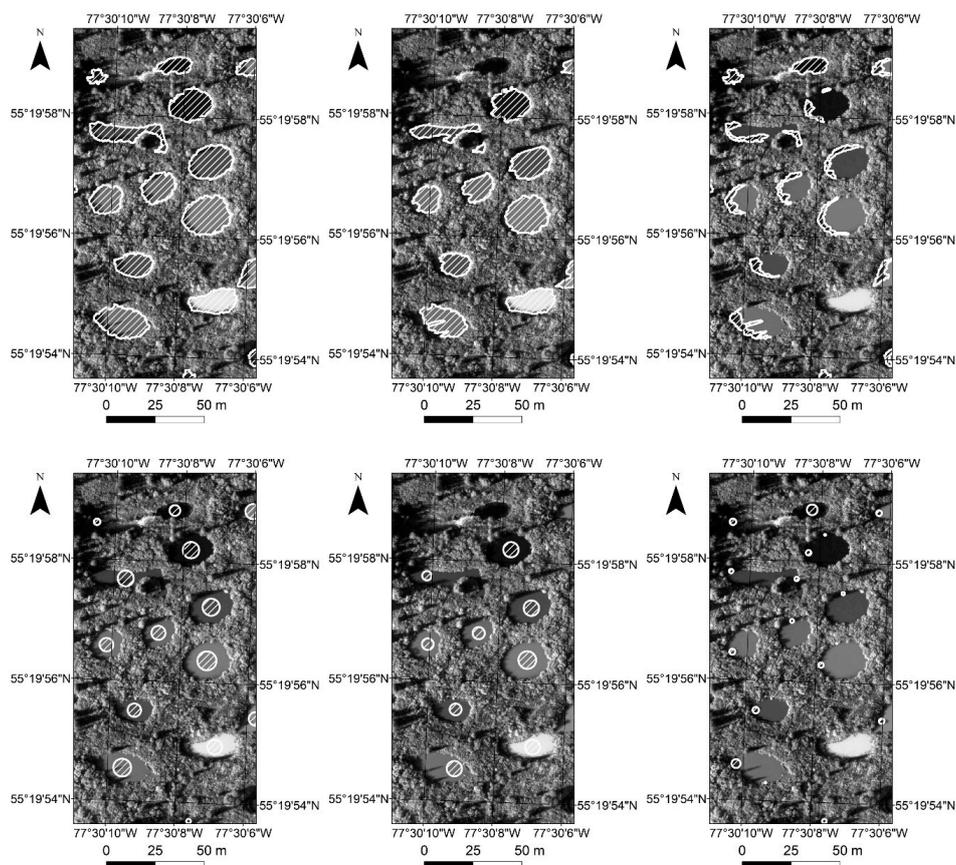
The lack of synchronicity between the UAS survey and the satellite imagery has two effects:

- sun position, since S2 satellite scenes are obtained at 10:30 local time, while the field survey

was conducted at 17:00, hence with different shadow and water surface reflection, and ii. possible changes in water spectral characteristics due to the difference on dates of collection, which is 9 days. We neglected these two effects, although we acknowledge that they should be mitigated in future assessments. The reason for not being able to accommodate these issues relates to logistical constraints for the UAS survey, limited by helicopter time and by difficulties in selecting the survey date, together with the high frequency of cloudy days in the study area. The 9-day difference is small and we assume that there were no significant spectral differences between the dates. The lakes are small, constrained by small heavily vegetated catchments and encroached by high shrubs, which limits short-term water color variability induced by wind effects or water inflow (the snowpack in the valley had totally melted >3 months earlier).

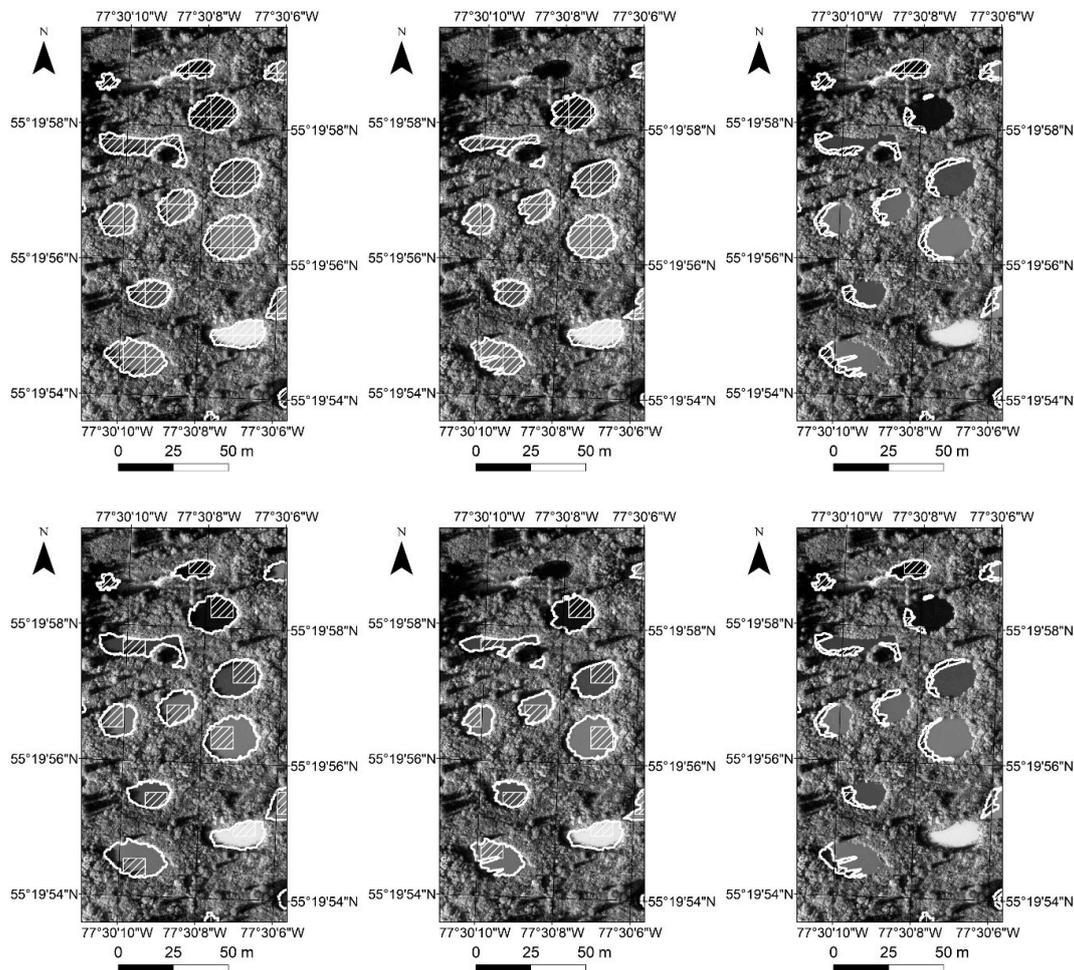
### 3.6. Methods for Extraction of Sensor Reflectance

Reflectance of small lakes and ponds obtained from satellite imagery, especially close to the lake margins, show a spectral mixing effect due to the presence of vegetation outside the lake, but also due to shadow effects from trees and tall shrubs that project into the water body surface. In order to better assess and mitigate this effect, the water bodies were classified in three areal categories: i. full-water body (FWB), ii. shaded water (SHW) and iii. sunlit water (SLW). Two methods were used for the extraction of the reflectance values from each of these areas in order to test for their quality: i. the average value of the pixels inside the area (M1 S), and ii. the average value of the pixels inside a circle in the center of the lake with an area of 20% of the lake area (M2 S) (Figure 7). This latter value was selected as the largest circle that fitted all lakes and ponds.



**Figure 7.** Methods used for the extraction of surface reflectance from the lake and pond surfaces from the Parrot Sequoia imagery. Upper row: average of the whole area of the marked polygons (M1 S), lower row: average of 20% area at the center of the target (M2 S). Left column: full-water body (FWB), Center column: sunlit water body (SLW), right column: shaded water body (SHD).

The extraction of the reflectance values from S2 imagery used the same two methods applied to the Sequoia imagery, which were named M1 S2 and M2 S2. Similarly, the FWB, SHW and SLW areas were used as reference. In M1 S2 the full intersected pixel values were used for reference, while in M2 S2 only the value of the pixel with largest intersected area was considered (Figure 8).

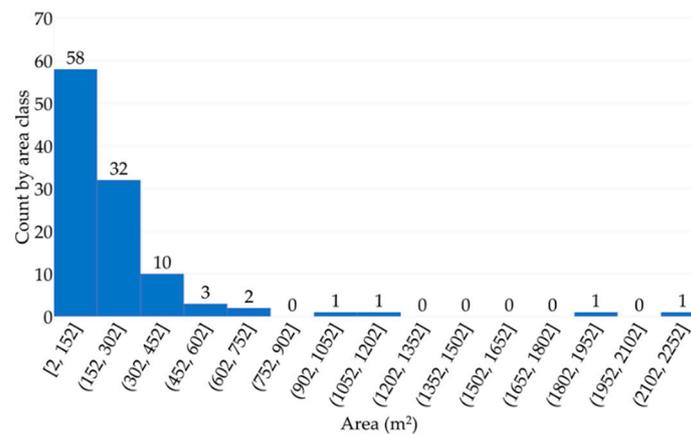


**Figure 8.** Methods used for the extraction of surface reflectance from the lake and pond surfaces from the Sentinel-2 imagery. Upper row: averaging the whole area (M1 S2), lower row: considering the value of the largest pixel (M2 S2). Left column: full-water body (FWB), Center column: sunlit water body (SLW), right column: shaded water body (SHD).

## 4. Results

### 4.1. Characteristics of the Themokarst Lakes and Ponds

The identification and delineation procedure showed the presence of 109 lakes and ponds in the KWAK area, covering 23,188 m<sup>2</sup> in a total surveyed area of 468,991 m<sup>2</sup>, which corresponded to 5% of the total study area. The lakes had an average area of 212 m<sup>2</sup> and a large standard deviation of 309 m<sup>2</sup>. The size distribution was positively skewed, with 50% of the water bodies below 150 m<sup>2</sup> and 80% below 300 m<sup>2</sup> (Figure 9).

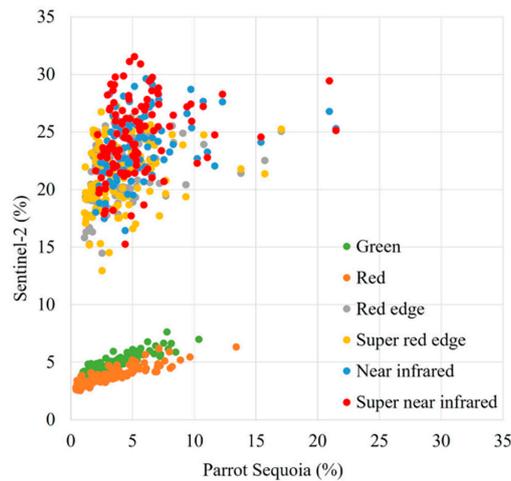


**Figure 9.** Thermokarst lake and pond size distribution in the KWAK study area.

#### 4.2. Evaluation of the Potential of Sentinel-2 for the Characterization of Small Water Bodies

The comparison between the reflectance values obtained by UAS survey using the Parrot Sequoia camera (here considered, the ground truth), with the values measured at the same lakes using the S2 scene is presented in Table 2. Two reflectance extraction methods from the Sequoia were tested, as well as two methods for S2. The effect of the shadows over the lake surface was assessed by analyzing the full-water body surface (FWB), the shaded water areas (SHW) and the sunlit water areas (SLW) separately. The analysis was made using Pearson correlation coefficients (R), considering values of  $p < 0.05$  and  $p < 0.01$  for evaluating the statistical significance of the relationships. The analysis was done by band comparison, as shown in Table 2. The objective was not to validate the accuracy of the satellite reflectance measures, but rather to identify if there was a good agreement between measures from lower to higher reflectance across the water surfaces, independently of the value agreement. This is presented as a first validation step to identify the minimum lake size for which S2 can detect and monitor the water reflectance.

The results showed that the correlations were generally better in the visible bands (Green and Red) of S2. In these bands the correlations were always positive and very strong or strong. The Red Edge and NIR bands showed weak positive or non-significant correlations in all methods and extraction methods. It is noteworthy that the correlation improved from the shaded water surfaces (SHW) to the sunlit surfaces (SLW), and even more to the full-water body surfaces (FWB), reflecting the effects of spectral mixing in the pixel footprint from the satellite sensors. The values were generally closer in the SLW and FWB with non-significant correlations in the SHW. The pansharpening applied to the Red Edge and NIR did not show an improvement compared to the Sequoia data. The correlations were always weak positive or non-significant, and in 18 out of 24 cases, the correlation is worse than with the original resolution. This shows that the pansharpening algorithm may improve contrast but does not accurately solve sub-pixel reflectance. Figure 10 shows the scatter plot of the reflectances obtained using the methods giving the highest correlation coefficient: the average value of the pixels inside a circle in the center of the lake with an area of 20% from Sequoia (M2 S), with the value of the full intersected pixels from S2 (M1 S2), using the full-water bodies. These were the two methods selected for assessing the water body size detection threshold in S2.



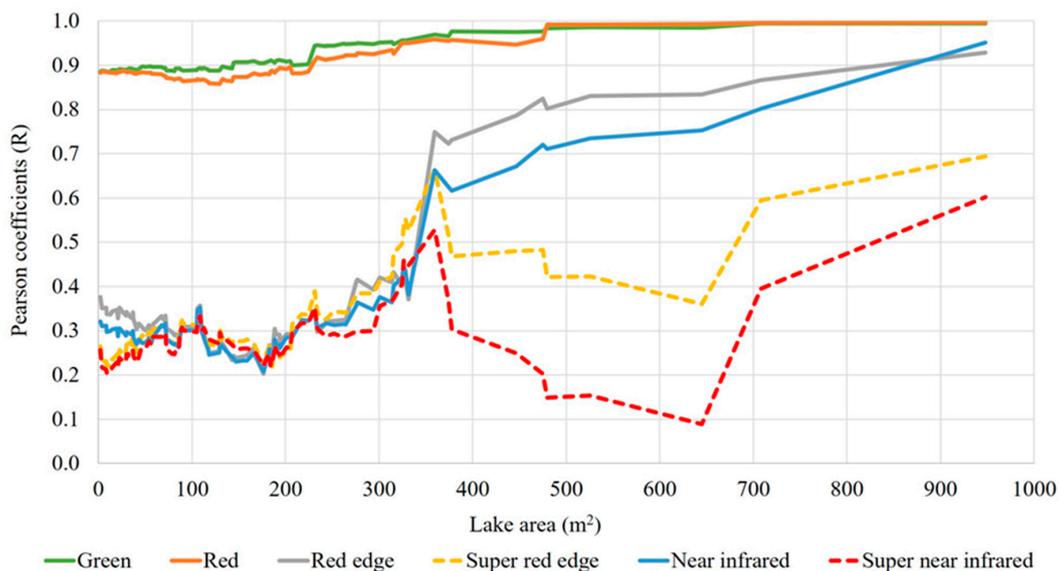
**Figure 10.** Reflectance of Parrot Sequoia (M2 S) and Sentinel-2 (M1 S2) for the full-water bodies.

**Table 2.** Correlation between reflectance from the green, red, red-edge, super red edge, near-infrared and super near-infrared bands from the Parrot Sequoia survey, the Sentinel-2 scene. The values are Pearson correlation coefficients (R).  shows values significant at  $p \leq 0.01$ .  shows values significant at  $p \leq 0.05$ .

		Bands	Sentinel-2	
			M1 S2	M2 S2
Method 1 Sequoia (M1 Sequoia)	Full-Water Body (FWB)	Green	<b><u>0.89</u></b>	<b><u>0.83</u></b>
		Red	<b><u>0.88</u></b>	<b><u>0.83</u></b>
		Red-edge	<b><u>0.33</u></b>	<b><u>0.29</u></b>
		Super red edge	<b><u>0.22</u></b>	<b><u>0.20</u></b>
		Near-infrared	<b><u>0.29</u></b>	<b><u>0.26</u></b>
		Super near-infrared	<b><u>0.22</u></b>	<b><u>0.19</u></b>
	Sunlit Water Surface (SLW)	Green	<b><u>0.90</u></b>	<b><u>0.86</u></b>
		Red	<b><u>0.87</u></b>	<b><u>0.83</u></b>
		Red-edge	<b><u>0.27</u></b>	<b><u>0.26</u></b>
		Super red edge	<b><u>0.27</u></b>	<b><u>0.24</u></b>
		Near-infrared	<b><u>0.24</u></b>	<b><u>0.24</u></b>
		Super near-infrared	<b><u>0.26</u></b>	<b><u>0.25</u></b>
	Shady Water Surface (SHW)	Green	<b><u>0.82</u></b>	<b><u>0.78</u></b>
		Red	<b><u>0.74</u></b>	<b><u>0.70</u></b>
		Red-edge	-0.05	-0.02
		Super red edge	-0.11	-0.09
		Near-infrared	-0.04	0.00
		Super near-infrared	-0.08	-0.06
Method 2 Sequoia (M2 Sequoia)	Full-Water Body (FWB)	Green	<b><u>0.89</u></b>	<b><u>0.83</u></b>
		Red	<b><u>0.88</u></b>	<b><u>0.82</u></b>
		Red-edge	<b><u>0.38</u></b>	<b><u>0.31</u></b>
		Super red edge	<b><u>0.27</u></b>	<b><u>0.23</u></b>
		Near-infrared	<b><u>0.32</u></b>	<b><u>0.27</u></b>
		Super near-infrared	<b><u>0.26</u></b>	<b><u>0.23</u></b>
	Sunlit Water Surface (SLW)	Green	<b><u>0.91</u></b>	<b><u>0.87</u></b>
		Red	<b><u>0.87</u></b>	<b><u>0.81</u></b>
		Red-edge	<b><u>0.28</u></b>	<b><u>0.27</u></b>
		Super red edge	<b><u>0.28</u></b>	<b><u>0.24</u></b>
		Near-infrared	<b><u>0.23</u></b>	<b><u>0.23</u></b>
		Super near-infrared	<b><u>0.26</u></b>	<b><u>0.24</u></b>
	Shady Water Surface (SHW)	Green	<b><u>0.75</u></b>	<b><u>0.73</u></b>
		Red	<b><u>0.63</u></b>	<b><u>0.60</u></b>
		Red-edge	-0.04	-0.04
		Super red edge	-0.05	-0.03
		Near-infrared	-0.04	-0.02
		Super near-infrared	-0.04	-0.02

#### 4.3. Identification of the Minimum Water Body Area Allowing for Spectral Characterization by Sentinel-2 Imagery

In order to evaluate the effects of lake area on the values of the Pearson correlation coefficients between the ground truth survey (UAS with Parrot Sequoia) and the S2 scene, we ranked the water bodies per area and conducted consecutive correlations using all lakes above a certain lake area threshold. As expected, a general increase in  $R$  was found as the threshold of lake area increased, as shown in Figure 11. This occurred in all bands, except in those which were pansharpened. In the visible spectrum, there was always a strongly positive correlation with the ground truthing, with a value close to 1 for waterbodies of about 500 m<sup>2</sup>. Such a behavior did not occur at the Red Edge and NIR bands, which showed a slight reduction in  $R$  with increasing lake area up to 200 m<sup>2</sup>, reflecting spectral mixture between water and the surrounding vegetation. The increase after that threshold should be linked to the increase of the water body to surrounding terrain ratio in the satellite pixel, as lake size increased.  $R$  increased regularly after 200 m<sup>2</sup>, became strongly positive above 350 m<sup>2</sup> ( $R > 0.7, p < 0.01$ ), and further increased thereafter.



**Figure 11.** Variation of the Pearson correlation coefficient ( $R$ ) with a moving threshold of water body area in KWAK for Sentinel-2 imagery and Parrot Sequoia reflectance measurements.

The super-resolution algorithm showed similar correlations to the original bands in the small waterbodies, increasing until the 350 m<sup>2</sup> threshold, but for larger lakes, showed a sudden decrease, not reaching statistically significant values thereafter.

Although the high versatility, adaptability and flexibility of the UAS systems, these cannot compete with satellite imagery in terms of spatial coverage [24]. Therefore, using a WorldView-4 scene of 25 km<sup>2</sup> around the KWAK site (Figure 12), we delineated 999 lakes and ponds, with 109 of them showing more than 350 m<sup>2</sup>. This means that S2 may allow for monitoring around 11% of the lakes and 73% of the total lake area. Based on a WorldView-2 scene for a subarctic region further northward (site BGR, located in the Sheldrake River Valley in the vicinity of Umiujaq), the representativeness of the minimum water body size class increased to 30% of all lakes and 85% of the total lake area (Figure 13).

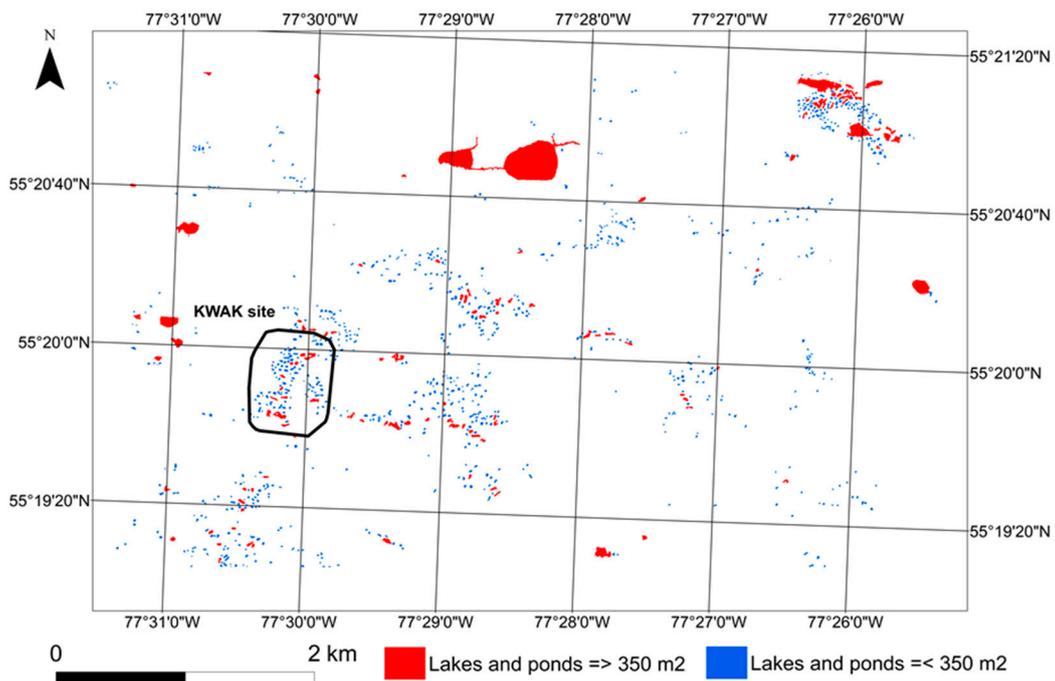


Figure 12. Lakes and ponds around the KWAK site.

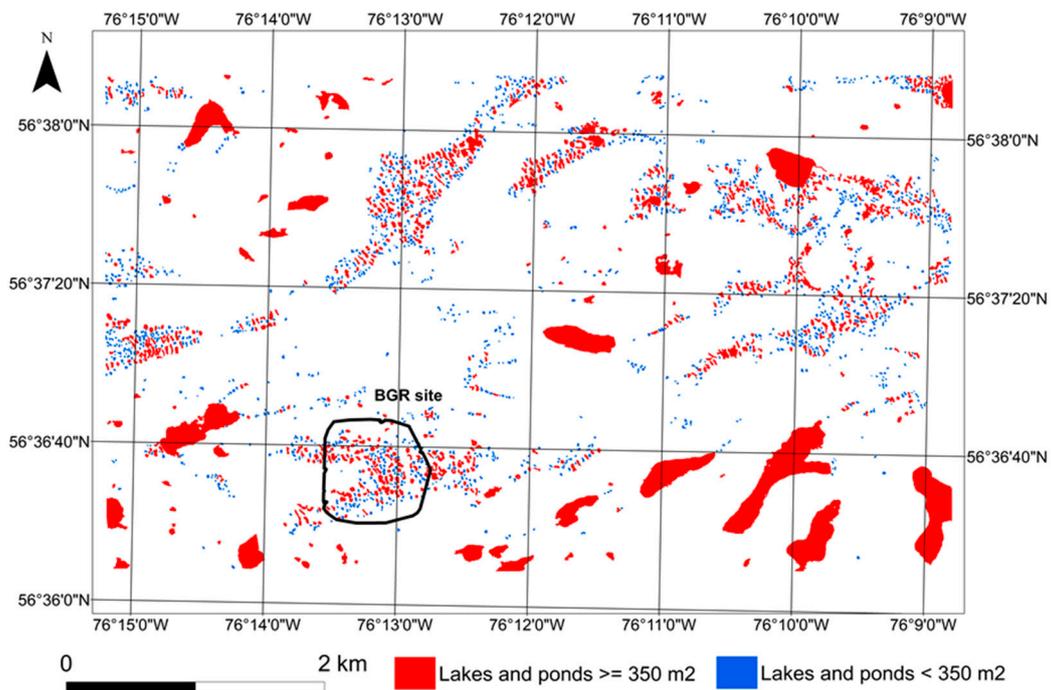


Figure 13. Lakes and ponds around the BGR site.

## 5. Discussion

With a frequent revisit time, S2 offers new unique possibilities for monitoring lake surfaces, but is limited by pixel size, which varies from 10 to 20 m. For the characterization of small water bodies using remote sensing the main challenge is related to the relationship between the pixel footprint of the sensor in relation to the water body boundaries. If the pixel is completely inside the water body, the surface characteristics can be properly assessed, but if the pixel footprint cuts across the lake boundary, then the reflectance is a mixture between the water and the neighboring surface properties.

For the accurate monitoring of small water bodies, a precise map with their boundaries is needed and for precise reflectance monitoring, the pixel footprint must entirely lie within the water body and best, away from its margins. Hence, in a perfect fit, for pixels central to a lake with a square shape, the minimum monitorable lake size would be equal to pixel size, which would be from 100 to 400 m<sup>2</sup> (depending on band resolution). Knowing that this situation is impossible in natural conditions, leaving a margin of 1 pixel around the monitored pixel would be the safest approach, which would increase the minimum square lake monitoring threshold to 900 to 3600 m<sup>2</sup>. This would preclude monitoring a vast number of the small lakes and ponds that play an important role in the Global Climate System. The only option for monitoring would be then to use commercial very high-resolution satellite imagery, which has a high price and limited availability [24]. The analyses in the present study aimed at evaluating the possibility of using a size threshold between the optimal threshold, and the best theoretically possible threshold, which would open a new range of possibilities to track the seasonal spectral dynamics of a subset of small lakes and ponds. The present approach used full-pixel values as a first step for a future sub-pixel spectral unmixing approach. It is important to note that only the lake area is analyzed here, independently of lake shape, which also plays a role in spectral mixing.

Inside the water body, factors such as spectral differences induced by different lake properties or by shadowing effects may also affect the spectral signature measured at the pixel site. To these, the typical effects influencing water reflectance measured at the sensor and related to sun angle, wind effects need to be accounted for [30]. Therefore, there are external factors that require further attention. These notably include the reflection from neighboring features and shadowing from high shrubs and trees, which are spatially and temporally variable and are especially problematical in high latitudes due to the high solar zenith angle. These factors are not yet accounted for in models of atmospheric and topographic correction and hence further research using ultra-high-resolution UASs and their ability to generate digital surface models will be of particular value. Other subjects needing further attention are the effects of lake morphometry, aquatic plants, tall vegetation surrounding the waterbodies and wind, and the potential for applying sub-pixel analysis and spectral unmixing techniques.

## 6. Conclusions

The detailed comparison of water reflectance from 109 thermokarst lakes and ponds obtained from a UAS survey with Sentinel-2 data allowed the determination of a lower boundary for monitoring waterbodies using remote sensing imagery. Sentinel-2 bands showed a significant increase in correlation with UAS reflectance measurements for waterbody sizes greater than 350 m<sup>2</sup>, a threshold where external effects may be reduced. This indicates the potential for using Sentinel-2 to monitor a substantial fraction of the total water body area in subarctic regions.

The approach made here was based on the premise that the use of Sentinel-2 for monitoring small lakes and ponds should be preceded by an accurate delineation of the lake boundaries, with very high-resolution imagery. We did this in the present study using an UAS. For the regional scale, lake boundaries may be easily obtained from very high-resolution imagery, such as that provided by WorldView-2, although regular updates may be required for permafrost landscapes that are in rapid transition.

**Author Contributions:** Conceptualization, P.F. and G.V.; Data curation, P.F. and G.V.; Formal analysis, P.F., G.V. and D.F.; Funding acquisition, G.V. and W.F.V.; Investigation, P.F., G.V., J.C., D.F. and W.F.V.; Methodology, P.F., G.V., J.C. and W.F.V.; Project administration, G.V. and W.F.V.; Resources, G.V., and W.F.V.; Supervision, G.V. and W.F.V.; Validation, P.F., G.V. and D.F.; Writing – original draft, P.F. and G.V.; Writing – review and editing, P.F., G.V., J.C., D.F. and W.F.V.

**Funding:** This research was funded by the project Shrubifly - Remote sensing characterization of vegetation formation changes in thaw lake catchments (Whapmagoostui—Kuujjuarapik, Hudson Bay, subarctic Canada), funded by the Portuguese Science and Technology Foundation (FCT) under the Portuguese Polar Program (PROPOLAR) and the ZEPHYRUS research group of the CEG/IGOT—Universidade de Lisboa (UID/GEO/00295/2019). Additional funding was provided by Natural Sciences and Engineering Research

Council of Canada (NSERC), the Fonds de Recherche du Québec-Nature et Technologies (FQRNT); the Canada First Research Excellence Fund (CFREF) program Sentinel North (project BOND), and the Network of Centres of Excellence ArcticNet.

**Acknowledgments:** The image analyses were conducted at the GEOMODLAB at CEG/IGOT. The manager and personnel of the CEN field station at Whapmagoostui-Kuujuarapik are thanked for their support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Schuur, E.A.G.; McGuire, A.D.; Schädel, C.; Grosse, G.; Harden, J.W.; Hayes, D.J.; Hugelius, G.; Koven, C.D.; Kuhry, P.; Lawrence, D.M.; et al. Climate change and the permafrost carbon feedback. *Nature* **2015**, *520*, 171–179. [[CrossRef](#)] [[PubMed](#)]
- Vincent, W.F.; Lemay, M.; Allard, M. Arctic permafrost landscapes in transition: Towards an integrated Earth system approach. *Arct. Sci.* **2017**, *3*, 39–64. [[CrossRef](#)]
- Wauthy, M.; Rautio, M.; Christoffersen, K.S.; Forsström, L.; Laurion, I.; Mariash, H.L.; Peura, S.; Vincent, W.F. Increasing dominance of terrigenous organic matter in circumpolar freshwaters due to permafrost thaw. *Limnol. Oceanogr.* **2018**, *3*, 186–198. [[CrossRef](#)]
- Laurion, I.; Vincent, W.F.; MacIntyre, S.; Retamal, L.; Dupont, C.; Francus, P.; Pienitz, R. Variability in greenhouse gas emissions from permafrost thaw ponds. *Limnol. Oceanogr.* **2010**, *55*, 115–133. [[CrossRef](#)]
- Matveev, A.; Laurion, I.; Deshpande, B.N.; Bhiry, N.; Vincent, W.F. High methane emissions from thermokarst lakes in subarctic peatlands. *Limnol. Oceanogr.* **2016**, *61*, S150–S164. [[CrossRef](#)]
- Bouchard, F.; Francus, P.; Pienitz, R.; Laurion, I.; Feyte, S. Subarctic thermokarst ponds: Investigating recent landscape evolution and sediment dynamics in thawed permafrost of northern Québec (Canada). *Arct. Antarct. Alp. Res.* **2014**, *46*, 251–271. [[CrossRef](#)]
- Polishchuk, Y.M.; Bryksina, N.A.; Polishchuk, V.Y. Remote analysis of changes in the number of small thermokarst lakes and their distribution with respect to their sizes in the cryolithozone of western Siberia. *Izv. Atmos. Ocean. Phys.* **2015**, *51*, 999–1006. [[CrossRef](#)]
- Anthony, W.K.; Deimling, T.S.; Nitze, I.; Frolking, S.; Emond, A.; Daanen, R.; Anthony, P.; Lindgren, P.; Jones, B.; Grosse, G. 21st-century modeled permafrost carbon emissions accelerated by abrupt thaw beneath lakes. *Nat. Commun.* **2018**, *9*, 3262. [[CrossRef](#)] [[PubMed](#)]
- Kokelj, S.V.; Jorgenson, M.T. Advances in thermokarst research. *Permafr. Periglac. Process.* **2013**, *24*, 108–119. [[CrossRef](#)]
- Breton, J.; Vallières, C.; Laurion, I. Limnological properties of permafrost thaw ponds in northeastern Canada. *Can. J. Fish. Aquat. Sci.* **2009**, *66*, 1635–1648. [[CrossRef](#)]
- Watanabe, S.; Laurion, I.; Chokmani, K.; Pienitz, R.; Vincent, W.F. Optical diversity of thaw ponds in discontinuous permafrost: A model system for water color analysis. *J. Geophys. Res.* **2011**, *116*, 1–17. [[CrossRef](#)]
- Bouchard, F.; Francus, P.; Pienitz, R.; Laurion, I. Sedimentology and geochemistry of thermokarst ponds in discontinuous permafrost, subarctic Quebec, Canada. *J. Geophys. Res.* **2011**, *116*, 1–14. [[CrossRef](#)]
- Payette, S.; Delwaide, A.; Caccianiga, M.; Beauchemin, M. Accelerated thawing of subarctic peatland permafrost over the last 50 years. *Geophys. Res. Lett.* **2004**, *31*, 1–4. [[CrossRef](#)]
- Ropars, P.; Boudreau, S. Shrub expansion at the forest–tundra ecotone: Spatial heterogeneity linked to local topography. *Environ. Res. Lett.* **2012**, *7*, 015501. [[CrossRef](#)]
- Kuhn, M.; Lundin, E.J.; Giesler, R.; Johansson, M.; Karlsson, J. Emissions from thaw ponds largely offset the carbon sink of northern permafrost wetlands. *Sci. Rep.* **2018**, *8*, 9535. [[CrossRef](#)]
- Muster, S.; Heim, B.; Abnizova, A.; Boike, J. Water body distributions across scales: A remote sensing based comparison of three Arctic tundra wetlands. *Remote Sens.* **2013**, *5*, 1498–1523. [[CrossRef](#)]
- Tian, B.; Li, Z.; Zhang, M.; Huang, L.; Qiu, Y.; Li, Z.; Tang, P. Mapping thermokarst lakes on the Qinghai–Tibet plateau using nonlocal active contours in Chinese GaoFen-2 multispectral imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 1687–1700. [[CrossRef](#)]
- Verpoorter, C.; Kutser, T.; Seekell, D.A.; Tranvik, L.J. A global inventory of lakes based on high-resolution satellite imagery. *Geophys. Res. Lett.* **2014**, *41*, 6396–6402. [[CrossRef](#)]

19. Pekel, J.-F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. [[CrossRef](#)]
20. Muster, S.; Roth, K.; Langer, M.; Lange, S.; Aleina, F.C.; Bartsch, A.; Morgenstern, A.; Grosse, G.; Jones, B.; Sannel, A.B.K.; et al. PeRL: A circum-Arctic permafrost region pond and lake database. *Earth Syst. Sci. Data* **2017**, *9*, 317–348. [[CrossRef](#)]
21. Hese, S.; Grosse, G.; Pöcking, S. Object based thermokarst lake change mapping as part of the ESA data user element (DUE) permafrost. *ISPRS* **2010**, XXXVIII-4/C7, 1682–1777.
22. Du, Y.; Zhang, Y.; Ling, F.; Wang, Q.; Li, W.; Li, X. Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. *Remote Sens.* **2016**, *8*, 354. [[CrossRef](#)]
23. Toming, K.; Kutser, T.; Laas, A.; Sepp, M.; Paavel, B.; Nöges, T. First experiences in mapping lake water quality parameters with Sentinel-2 MSI imagery. *Remote Sens.* **2016**, *8*, 640. [[CrossRef](#)]
24. Manfreda, S.; McCabe, M.F.; Miller, P.E.; Lucas, R.; Madrigal, V.P.; Mallinis, G.; Dor, E.B.; Helman, D.; Estes, L.; Ciruolo, G.; et al. On the use of unmanned aerial systems for environmental monitoring. *Remote Sen.* **2018**, *10*, 641. [[CrossRef](#)]
25. Bhiry, N.; Delwaide, A.; Allard, M.; Bégin, Y.; Filion, L.; Lavoie, M.; Nozais, C.; Payette, S.; Pienitz, R.; Saulnier-Talbot, E.; Vincent, W.F. Environmental change in the Great Whale River region, Hudson Bay: Five decades of multidisciplinary research by Centre d'études nordiques (CEN). *Écoscience* **2011**, *18*, 182–203. [[CrossRef](#)]
26. McManus, K.M.; Morton, D.C.; Masek, J.G.; Wang, D.; Sexton, J.O.; Nagol, J.; Ropars, P.; Boudreau, S. Satellite-based evidence for shrub and graminoid tundra expansion in northern Quebec from 1986 to 2010. *Glob. Chang. Boil.* **2012**, *18*, 2313–2323. [[CrossRef](#)]
27. Dufour-Tremblay, G.; Vriendt, L.D.; Lévesque, E.; Boudreau, S. The importance of ecological constraints on the control of multi-species treeline dynamics in Eastern Nunavik, Québec. *Am. J. Bot.* **2012**, *99*, 1638–1646. [[CrossRef](#)] [[PubMed](#)]
28. Tremblay, B.; Lévesque, E.; Boudreau, S. Recent expansion of erect shrubs in the Low Arctic: Evidence from Eastern Nunavik. *Environ. Res. Lett.* **2012**, *7*, 035501. [[CrossRef](#)]
29. Bégin, P.N.; Vincent, W.F. Permafrost thaw lakes and ponds as habitats for abundant rotifer populations. *Arct. Sci.* **2017**, *3*, 354–377. [[CrossRef](#)]
30. Assmann, J.J.; Kerby, J.T.; Cunliffe, A.M.; Myers-Smith, I.H. Vegetation monitoring using multispectral sensors—Best practices and 3 lessons learned from high latitudes. *NRC J. Unmanned Veh. Syst.* **2019**, *7*, 54–75. [[CrossRef](#)]
31. MicaSense. *Sequoia User Guide*; Drones Parrot SAS: Paris, France, 2018; pp. 4–33.
32. ESA. *Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services*; European Space Agency: Noordwijk, The Netherlands, 2012; pp. 15–45.
33. Brodu, N. Super-resolving multiresolution images with band-independent geometry of multispectral pixels. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4610–4617. [[CrossRef](#)]
34. Pix4D. *Pix4Dmapper 4.0 User Manual*; Pix4D SA: Lausanne, Switzerland, 2017.
35. Canadian Digital Elevation Model. Available online: <https://open.canada.ca/data/en/dataset/7f245e4d-76c2-4caa-951a-45d1d2051333> (accessed on 8 January 2008).
36. Kelly, M.; Blanchard, S.D.; Kersten, E.; Koy, K. Terrestrial remotely sensed imagery in support of public health: New avenues of research using object-based image analysis. *Remote Sen.* **2011**, *3*, 2321–2345. [[CrossRef](#)]
37. Dey, V. A Supervised Approach for the Estimation of Parameters of Multiresolution Segmentation and Its Application in Building Feature Extraction from VHR Imagery. Master's Thesis, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, NB, Canada, September 2011.
38. Jawak, S.D.; Kulkarni, K.; Luis, A.J. A review on extraction of lakes from remotely sensed optical satellite data with a special focus on cryospheric lakes. *Adv. Remote Sens.* **2015**, *4*, 196–213. [[CrossRef](#)]

