

## Article

# Monitoring Variations in Lake Water Storage with Satellite Imagery and Citizen Science

Sarina Little <sup>1</sup>, Tamlin M. Pavelsky <sup>1,\*</sup>, Faisal Hossain <sup>2</sup>, Sheikh Ghafoor <sup>3</sup>, Grant M. Parkins <sup>4</sup>, Sarah K. Yelton <sup>4</sup>, Megan Rodgers <sup>4</sup>, Xiao Yang <sup>1</sup>, Jean-François Crétaux <sup>5</sup>, Catherine Hein <sup>6</sup>, Mohammad Arman Ullah <sup>3</sup>, Debolina Halder Lina <sup>3</sup>, Hanne Thiede <sup>7</sup>, Darren Kelly <sup>8</sup>, Donald Wilson <sup>9</sup> and Simon N. Topp <sup>1</sup>

- <sup>1</sup> Department of Geological Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA; sarina.rinab@gmail.com (S.L.); yangxiao@live.unc.edu (X.Y.); sntopp@live.unc.edu (S.N.T.)
- <sup>2</sup> Department of Civil & Environmental Engineering, University of Washington, Seattle, WA 98195, USA; fhossain@uw.edu
- <sup>3</sup> Department of Computer Science, Tennessee Tech University, Cookeville, TN 38505, USA; sghafoor@tntech.edu (S.G.); mullah42@students.tntech.edu (M.A.U.); dlina@students.tntech.edu (D.H.L.)
- <sup>4</sup> Institute for the Environment, University of North Carolina, Chapel Hill, NC 27599, USA; parkins@unc.edu (G.M.P.); sarah.yelton@unc.edu (S.K.Y.); megan.rodgers@unc.edu (M.R.)
- <sup>5</sup> UMR5566 CNRS-CNES-IRD-Université de Toulouse III, 31400 Toulouse, France; Jean-Francois.Cretaux@legos.obs-mip.fr
- <sup>6</sup> Wisconsin Department of Natural Resources, Madison, WI 53703, USA; Catherine.Hein@wisconsin.gov
- <sup>7</sup> Citizen Scientist, Sammamish, WA 98074, USA; hanne.thiede@gmail.com
- <sup>8</sup> Citizen Scientist, Harrells, NC 28444, USA; darren.k.kelly@gmail.com
- <sup>9</sup> Citizen Scientist, Lindenhurst, IL 60046, USA; wilsondonaldm@gmail.com
- \* Correspondence: pavelsky@unc.edu; Tel.: +1-919-962-4239



**Citation:** Little, S.; Pavelsky, T.M.; Hossain, F.; Ghafoor, S.; Parkins, G.M.; Yelton, S.K.; Rodgers, M.; Yang, X.; Crétaux, J.-F.; Hein, C.; et al. Monitoring Variations in Lake Water Storage with Satellite Imagery and Citizen Science. *Water* **2021**, *13*, 949. <https://doi.org/10.3390/w13070949>

Academic Editors:  
Dariusz Wrzesiński and  
Leszek Sobkowiak

Received: 15 January 2021  
Accepted: 20 March 2021  
Published: 30 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Abstract:** Despite lakes being a key part of the global water cycle and a crucial water resource, there is limited understanding of whether regional or lake-specific factors control water storage variations in small lakes. Here, we study groups of small, unregulated lakes in North Carolina, Washington, Illinois, and Wisconsin, USA using lake level measurements gathered by citizen scientists and lake surface area measurements from optical satellite imagery. We show the lake level measurements to be highly accurate when compared to automated gauges (mean absolute error = 1.6 cm). We compare variations in lake water storage between pairs of lakes within these four states. On average, water storage variations in lake pairs across all study regions are moderately positively correlated ( $\rho = 0.49$ ) with substantial spread in the degree of correlation. The distance between lake pairs and the extent to which their changes in volume are correlated show a weak but statistically significant negative relationship. Our results indicate that, on regional scales, distance is not a primary factor governing lake water storage patterns, which suggests that other, perhaps lakes-specific, factors must also play important roles.

**Keywords:** lake water storage; water level fluctuations; remote sensing; citizen science



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Surface water stored in lakes is a key part of the global water cycle and a crucial water resource that provides drinking water, supports irrigation systems, promotes economic activity and tourism, and generates hydroelectric power. In addition to supporting human activities, lakes sustain diverse ecosystems important for many natural processes. Lakes encompass geographically diverse areas, are the lowest points in their surrounding landscapes, serve as records of past hydrologic and geologic events, and regulate surrounding climate. Because they integrate so many processes, lakes and reservoirs act as sentinels of climate change and are among the areas most threatened by climate change and other human impacts [1–4]. Lake water level, especially in endorheic (or seepage) lakes, is very

sensitive to changes in the water balance [3,5]. Studying lake levels allows us to better understand the effects of climate change on the water cycle and lake ecosystems [2,3].

Lake measurement systems, whether in situ or satellite-based, tend to focus mostly on larger lakes. Lakes smaller than 10 km<sup>2</sup> are generally poorly observed and monitored [6,7]. There are, however, many more small lakes than large lakes in the world. For example, of the >1.4 million lakes larger than 0.1 km<sup>2</sup> in the HydroLakes database [8], >1.2 million are smaller than 1.0 km<sup>2</sup>. Furthermore, as lakes decrease in waterbody size, the total shoreline length of those lakes increases [8]. As such, most land–lake interfaces occur in small lakes, in return giving these small lakes a significant role in biogeochemical cycles [6,8]. Natural lakes are not spread evenly over the globe; some areas have many lakes, while others have none. In lake-rich regions, it is possible to conceptualize lakes as either individual features or as common expressions of a shared water table. Which of these makes more sense depends on the specific question at hand but also on how correlated variations in lake water storage are within a region.

While there have been studies looking at regional patterns of change in lake water level [9,10] or surface area [11] and while many ecologists study small lake ecosystems, spatiotemporal patterns in lake water storage, especially for small lakes, remain poorly studied despite their abundance and hydrological significance. A handful of studies [12–14], mostly focused on the Tibetan Plateau, have examined controls on lake volume across multiple lakes within a region. However, our understanding remains limited regarding how the storage of water in regional groups of small lakes fluctuates, how that differs from large lakes, what controls these fluctuations, and the implication of these fluctuations on the ecosystem services they provide [6,15].

This gap in knowledge regarding small lakes is in large part caused by limitations in current lake monitoring systems [3]. To date, lake monitoring is primarily done through government-run gauge networks, studies by individual scientists, and satellite radar altimetry. In the United States, the most robust gauge network is maintained by the United States Geological Survey (USGS) and includes around 300 lakes, most of which are relatively large, man-made, or dammed reservoirs [16]. Many state and local governments have their own gauge networks, such as Lake Level Minnesota [17] and Water Data for Texas [18]. These data are often limited for scientific purposes by major temporal discontinuities (as in the case of Minnesota) or by a focus on regulated reservoirs (as in Texas). Datasets collected by individual scientists tend to be of short duration and cover a limited number of lakes. This leaves a large majority of the natural lakes in the United States unmonitored [6–8]. Similar deficiencies exist elsewhere in the world [19–21]. Furthermore, there has been an overall decline in gauging networks due to a lack of funding, lack of personnel for installation and data collection, and difficulty of installation [22]. Satellite radar altimetry can help monitor lake water level and lake water volume in areas where in situ measurements are not available, as well as provide measurements over relatively long-time spans [23]. These measurements, which are increasingly applicable to smaller lakes [24,25], remain most effective for lakes larger than 1–10 km<sup>2</sup> [26,27] and are available only for a small fraction of lakes worldwide because satellite altimeter ground tracks are widely spaced [28].

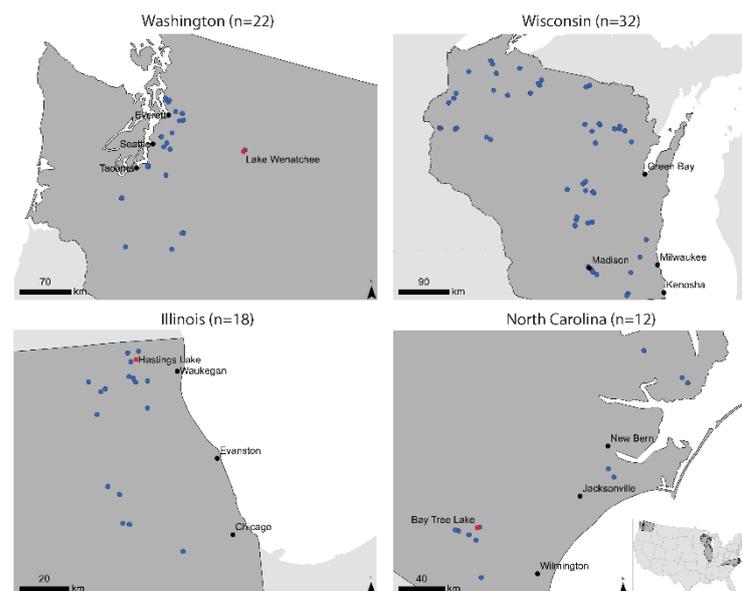
These limitations of current lake monitoring systems suggest that new strategies are needed to understand variations in small lakes. High-resolution remote sensing of inundation extent paired with measurements of lake stage by citizen scientists offers a promising alternative approach. Time series of optical satellite imagery from sensors on satellites such as Landsat and Sentinel 2 can be used to measure variations in lake area over time, an important metric that provides key information about changes in lake water storage [11,23,29–31]. However, on their own, these lake area measurements can be difficult to interpret, since they provide only indirect information about water surface elevation. Meanwhile, a new emphasis on citizen science in hydrology has shown that it is possible to work with members of the public to measure water levels across a broad range of water bodies [32–35]. By combining measurements of lake area from satellites and lake level from

citizen scientists, we can measure variations in lake volume and storage to address key science questions in lake hydrology.

In this paper, we aim to better understand the factors that control regional patterns of lake water storage. We test whether the storage of water in regional groups of small lakes varies in concert, suggesting regional-scale drivers, or if such variations are primarily driven by local factors unique to each lake. To achieve this objective, we calculate changes in lake water storage over time in groups of small natural lakes in Wisconsin, Illinois, North Carolina, and Washington, USA through the use of lake level measurements gathered by citizen scientists and lake extent measurements taken from optical satellite imagery. We compare these time series of water storage between lakes within regional groups. If time series within a group are highly correlated, then controls on lake water storage are likely regional in nature. If they are uncorrelated, then local controls specific to each lake must be relatively important.

## 2. Study Areas

For this study, we selected groups of natural lakes in North Carolina, Washington, Illinois, and Wisconsin (Figure 1). We chose these lakes based on the feasibility of gauge installation, ease of access, and lack of active controls on water levels. Because we were interested in observing natural patterns, it was important to exclude any actively human-controlled lakes or reservoirs. However, some of the lakes have structures that, while not actively managed, were designed to generally keep water levels higher than would naturally occur. The one exception is Horsepen Lake, in North Carolina, which experienced minor changes to its outlet following Hurricane Florence in 2018. We retained it in our analysis, though removing it would not substantially affect our results. We included both drainage lakes and seepage lakes and did not explicitly attempt to differentiate between them. The lake regions selected are of varying geographies and contain lakes with diverse characteristics, such as area, depth, and proximity to other study lakes (Table 1) [36]. A table of lake properties can be found in the Appendix A (Table A1). We worked with citizen scientists to collect water level data for lakes in North Carolina, Washington, and Illinois as part of the Lake Observations from Citizen Scientists and Satellites (LOCSS) Project [37], and we obtained data for lakes in Wisconsin from an existing network operated by the Wisconsin Department of Natural Resources [38].



**Figure 1.** Lake gauge locations in each study region. Blue dots indicate study lakes, while red dots represent lakes.

**Table 1.** Comparison of lake level monitoring attributes between states.

	North Carolina	Washington	Illinois	Wisconsin
Number of Lakes	12	22	18	32
Lake Areas (km <sup>2</sup> ) Mean (Min–Max)	24.54 (0.89–103)	5.72 (0.06–89)	0.41 (0.06–0.98)	3.85 (0.05–39.61)
Distance to other study Lakes (km) Mean (Min–Max)	115.32 (0.3–249)	89.72 (1–206)	31.85 (0.1–83)	181.62 (0.8–437)
Satellite Images Available (1 January 2015–1 February 2020)	2497	2303	1289	3056
Number of Lake Level Measurements	2238	1296	827	960
Lake Level Starting Dates	18 April 2017	10 September 2018	13 May 2019	1 April 2015

### 3. Methods

This project combines lake water levels collected by citizen scientists (Section 3.1) with same-day surface water areas calculated from Landsat and Sentinel 2 imagery (Section 3.2) to determine the change in lake water storage over time (Section 3.3). We then compared the time series of lake water storage between lakes within each regional group to determine if lakes fluctuate in coherence (Section 3.4) as well as determine if the distance between lakes drives coherence (Section 3.5).

#### 3.1. Measuring Lake Water Levels

##### 3.1.1. Data Acquisition

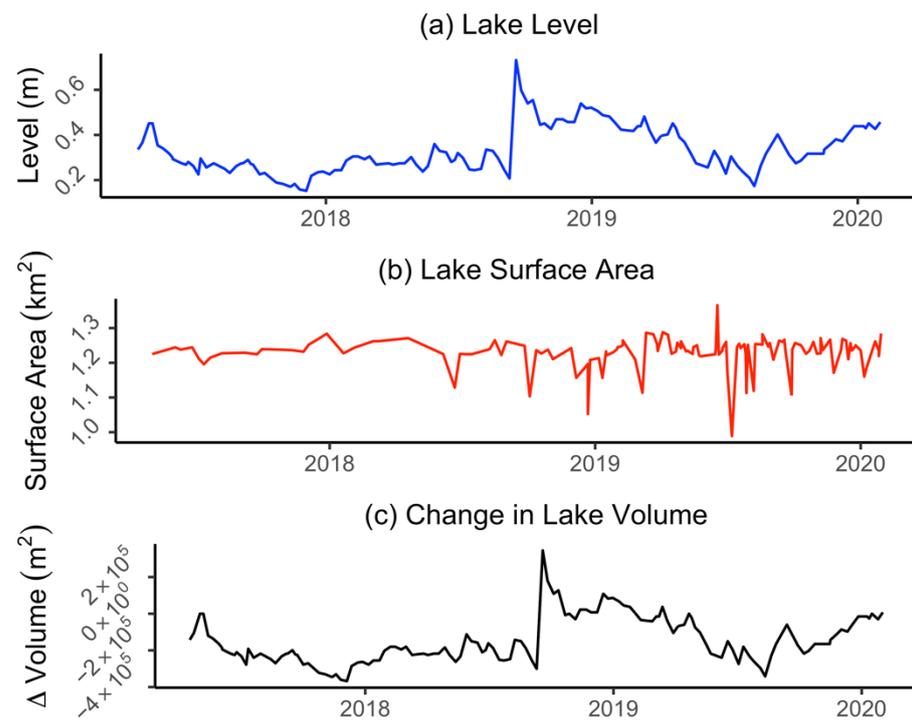
Two citizen science projects provide the lake level data used in this study. For North Carolina, Washington, and Illinois, the lake level data come from measurements collected through the LOCSS Project. LOCSS is a NASA-funded project, begun in 2017, that combines lake level data reported by a network of citizen scientists with satellite images, from which lake area is determined, to better understand how the volume of water in natural lakes is changing over time [39]. To collect these measurements, the LOCSS project installed water level gauges into natural lakes. On top of the gauge is a sign with instructions, a unique gauge ID, and a phone number (Figure 2). Citizen scientists passing by the gauge read the lake level and text in the measurement, along with the gauge ID, to the phone number. Alternatively, citizen scientists provide measurements on datasheets, via the LOCSS website [37], or occasionally via other means such as email or a phone call. The measurements are recorded and displayed in real-time on the LOCSS website. The record begins on 18 April 2017 for the North Carolina lakes, on either 10 September 2018 or 11 June 2019 for the Washington lakes, and on 13 May 2019 for the Illinois lakes, depending on the lake gauge installation date. All US LOCSS data are collected from gauges incremented in units of 0.02 feet (~6 mm). Because it is possible to estimate between two increments, we generally receive measurements to the nearest 0.01 foot (~3 mm). The data submitted by the citizen scientists occur mostly at random time intervals, with some gauges having more readings than others. We were interested in capturing measurements on or near Landsat and Sentinel 2 overpass dates, however, to complete our volumetric analysis, it is also worth noting that citizen scientists can sign up for email notifications and the LOCSS newsletter on the LOCSS website ([www.locss.org](http://www.locss.org), accessed 1 March 2021), which encourages and reminds them to make measurements on or close to Landsat and Sentinel 2 overpass dates.



**Figure 2.** An example of a LOCSS lake gauge with a sign on top for texting in measurements. Figure (a) depicts a LOCSS lake gauge sign and what a citizen scientist would read. Figure (b) depicts a LOCSS lake gauge located in Lake Howard, WA.

The lake level data from Wisconsin come from multiple sources but are dominated by volunteer lake level monitoring efforts coordinated by the Wisconsin Department of Natural Resources (WDNR) and partners (USGS, The North Lakeland Discovery Center, county-level land and water departments, Table 1) [38]. Like the LOCSS project, volunteers in Wisconsin take readings on a staff gauge and various professionals install/remove and survey the staff gauges each spring and fall. Unlike LOCSS, individual volunteers are responsible for monitoring a specific lake at least monthly, but ideally weekly, during the ice-free season and are either trained to enter their data into WDNR’s Surface Water Integrated Monitoring System or mail in a paper copy of their datasheet. The earliest data used for this study start from on 1 April 2015, but start dates vary by lake. For this study, we analyzed data collected until 1 February 2020. The LOCSS data collection method was, in part, inspired by the CrowdHydrology text messaging system [32,34] and the Wisconsin network protocols [38], so this study is partially indebted to their previous work.

An example of the lake level time series for Bay Tree Lake in eastern North Carolina shows the typical format of the data (Figure 3a). Outliers were automatically filtered out of the North Carolina, Washington, and Illinois lakes based upon the maximum and minimum values on the gauge boards. Measurements that exceeded the maximum possible value on a gauge were automatically removed. There are eight outliers in North Carolina (out of 2238 total measurements), 33 (out of 1296) in Washington, and none (out of 827) in Illinois. In Wisconsin, no outlier removal was required because the data were preprocessed by Wisconsin’s DNR. In total, 5287 lake level measurements were used in this study.



**Figure 3.** Example time series of (a) lake water level, (b) lake surface area, and (c) change in lake volume in Salters Lake, located in eastern North Carolina.

### 3.1.2. GPS Processing

For three lakes in North Carolina and one in Washington, there are multiple gauges on the same lake. We focused on studying change in lake level for the whole lake, not the individual gauges, so for these cases, the lake level records were merged together to create one record of change in lake level. To do this accurately, the scales of the lake levels were matched based on GPS elevation data taken in the field during gauge installation. GPS data were recorded at a rate of one measurement per second using a high-precision Septentrio PolaRx-5 receiver (Leuven, Belgium) that was floated on a small raft [40] for one hour immediately adjacent to the gauge.

This elevation was paired with the lake level measurement that was taken on that day. If the height of a gauge were to change due to the effects of ice, human impacts, or other factors, a new GPS elevation would be collected once the gauge was reinstalled or fixed. Depending on the elevation and lake level measurement pair taken on that day the old and new record could be matched. To match up multiple gauges on one lake, we used the GPS and lake level pairing for each gauge to determine what the lake levels were at the same elevations. We then calculated the difference between the levels at the same elevation and added or subtracted that value from the lake level data record for each gauge.

### 3.1.3. Validation

One important concern regarding citizen science data is data accuracy, especially when collected by those without specific training [41]. To address this concern and validate our lake level data, we installed Levellogger Junior Edge pressure transducers (Manufactured by Solinst, Georgetown, ON, Canada), corrected for atmospheric pressure variations using a Solinst Barologger located within at least 40 km. We collected coincident citizen science and pressure transducer measurements at 14 gauges in North Carolina, 7 gauges in Washington, and 12 gauges in Illinois. The pressure transducers recorded measurements every 15 min for periods of 6–10 months. The pressure transducer data and citizen science data were then adjusted to the same units and scale, matched by date and time, and regressed against each other to determine accuracy.

### 3.2. Measuring Lake Surface Area

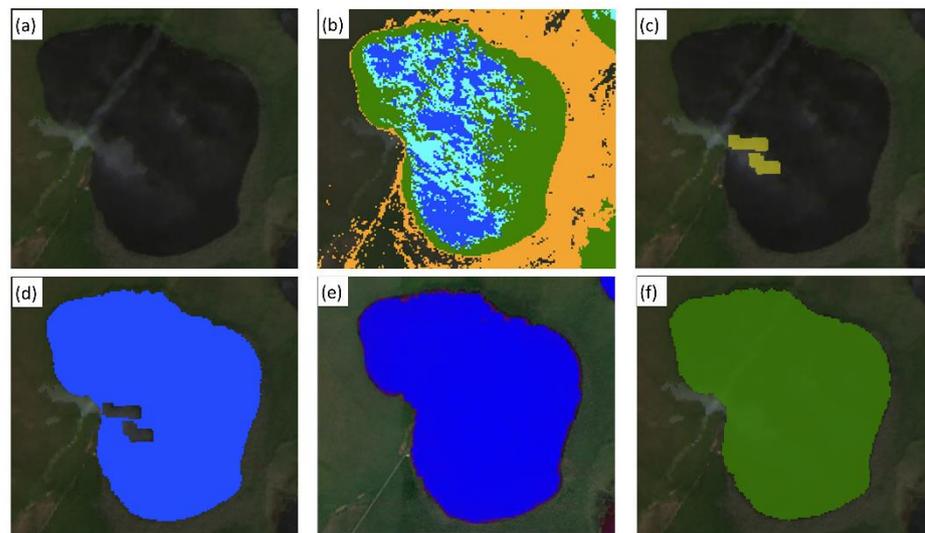
#### 3.2.1. Data Acquisition

To monitor change in lake surface extent, we used Landsat 8 Optical Land Imager (OLI) and Sentinel 2 MultiSpectral Imager (MSI) images. The Landsat satellites have a temporal resolution of 16 days and a spatial resolution of 30 m. The Sentinel 2 satellites have a temporal resolution of five days and a spatial resolution of 10 m. The use of both Landsat and Sentinel 2 sensors greatly increases image availability. We acquired the images and calculated the lake surface area using code written in Google Earth Engine (GEE) [42,43]. The satellite images were selected using lake polygons from the National Hydrography Dataset (NHD) [44]. Lake polygons do not exist in the NHD database for Horsepen Lake in North Carolina or for Deep Quarry, Harrier, and Herrick lakes in Illinois. In these cases, lake polygons are hand drawn in GEE. Clouds in images lead to commission errors, false positives, and overestimation, so images with more than 30% cloud cover were filtered out. After this filter was applied, we retained 8590 satellite images during the timeframe of our study.

#### 3.2.2. Water Mask

For all images in this collection, we calculated lake surface area using the Dynamic Surface Water Extent (DSWE) method [45] on both the Landsat and Sentinel 2 imagery. In order to use DSWE for the Sentinel 2 imagery, we first applied a transformation function to the Sentinel 2 imagery to approximate reflectance values for equivalent Landsat bands [46,47]. The DSWE algorithm then classified each pixel into one of five categories: high probability of water, moderate probability of water, low probability of water, not water, or wetland (Figure 4b). Open water pixels with a high or moderate level of confidence and wetland pixels were selected to create the water mask over each lake (Figure 4d). We included wetland pixels due to the large amounts of inundated vegetation in some lakes. To allow the inclusion of images in which a lake is partially obscured by clouds, we employed a cloud filling algorithm. We identified cloudy areas over lakes using the CFmask algorithm for Landsat [48] and the built-in quality band for Sentinel 2. It fills in the cloud-created gaps in the lake mask (Figure 4c) within the lake polygon boundary using the Surface Water Occurrence (SWO) Value developed by Pekel et al. [28]. The SWO value represents the frequency, between 1984 and 2015, that surface water was observed by Landsat. We used the mean water occurrence value from the shoreline pixels along the DSWE water mask as a dynamic threshold for water occurrence to fill in the cloud-obscured areas within the lake and designate those pixels as water (Figure 4e). Using a function written in Google Earth Engine (GEE), we then calculated the lake surface area for each lake image in the time series (Figure 3f).

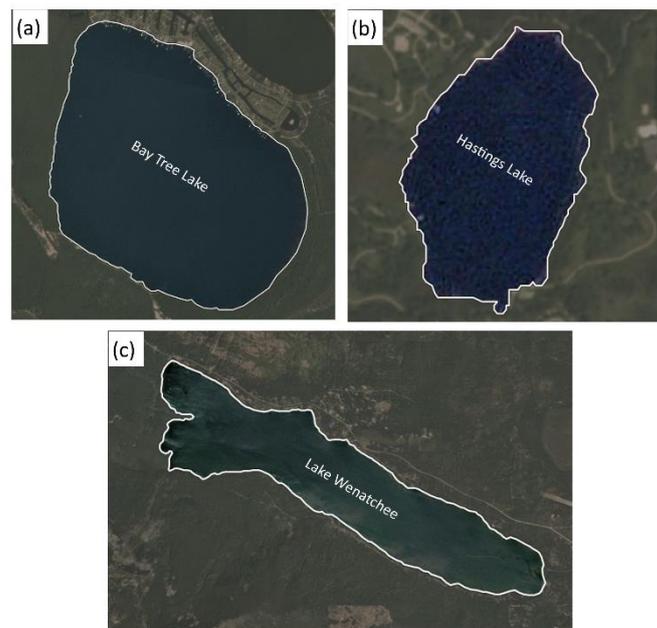
Values  $>\pm 25\%$  from the median surface area value in the lake surface area time series were treated as outliers and removed. This threshold was chosen by manually checking lakes with apparent changes in area. In lakes with very high variations in surface area, this threshold might require adjustment, but the surface areas of lakes studied here are sufficiently stable so the application of this threshold is unlikely to result in the removal of correct observations. After removing the outliers, the number of usable satellite images dropped to 7961. These images were matched to the lake level data to calculate variations in lake volume (See Section 3.3).



**Figure 4.** Representation of our surface area classifications over Great Lake in Eastern North Carolina taken on 19 January 2017: (a) Landsat satellite image; (b) DSWE classifications with high probability of water (blue), moderate probability of water (cyan), low probability of water (orange), inundated vegetation or wetland (green), no water (no color); (c) cloud mask; (d) initial water mask; (e) Pekel surface water occurrence values ranging from low occurrence (red) to high occurrence (dark blue); (f) complete water mask with cloudy areas filled in.

### 3.2.3. Validation

To evaluate the accuracy of our automated lake surface extent calculations from the Landsat and Sentinel 2 imagery, we used high-resolution satellite data from Planet (San Francisco, CA, USA), PlanetScope and RapidEye imagery, to provide a second estimate of lake surface extent [49]. Planet imagery was collected using 170 CubeSat satellites and provided daily temporal resolution [49]. It has three satellite types: PlanetScope with 3.7 m resolution, RapidEye with 6.5 m resolution, and SkySat with 72 cm resolution [49]. For this analysis, we used the PlanetScope and RapidEye images because they are freely available to researchers, and we conducted the analysis on the Planet imagery platform, which includes a built-in tool that automatically calculates the area of a polygon. While Cooley et al. [11,50] developed automated surface area calculation methods using Planet imagery, here we manually delineated surface areas primarily due to the small scale of the analysis and potential errors associated with automated detection, such as errors caused by clouds, shadows, and wetlands. We chose three lakes for the validation: Bay Tree Lake in eastern North Carolina, Hastings Lake in Illinois, and Lake Wenatchee in Washington. These lakes were chosen because they represent different test cases. Bay Tree Lake poses little difficulty when calculating lake surface area. Hastings Lake, however, is one of our smallest lakes, leading to potential errors. Lake Wenatchee, meanwhile, experiences both topographic shadow and adjacent snow cover. We selected five dates from each lake's surface area record that also had a corresponding lake level measurement and found the corresponding Planet image on that date. In three cases, images were matched plus or minus a day due to a gap in the Planet record or clouds. Some dates had two usable images due to overlapping satellite orbits. In these cases, both were used and tested. For the analysis, we first manually drew the lake boundary on each image, creating a lake polygon (Figure 5). Next, we calculated the surface area over that polygon using the Planet interface tool. Everything inside of the polygon is considered water, so the lake surface area is equivalent to the polygon area. Additionally, we compared the same-day area estimates from Sentinel 2 and Landsat for the subset of lakes/days when both sensors captured images ( $n = 209$ ).



**Figure 5.** Examples of lake surface area validation and Planet imagery in (a) a simple case, Bay Tree Lake on 21 July 2017, (b) a case with a small lake, Hastings Lake on 24 November 2019, and (c) a case with topographic shadow, Lake Wenatchee on 2 June 2019. The white line is the manually delineated lake shoreline. Everything inside of that boundary is considered water. Surface areas of (a) 5.88 km<sup>2</sup>, (b) 0.29 km<sup>2</sup>, and (c) 9.68 km<sup>2</sup> were calculated.

### 3.3. Measuring Lake Water Storage

To calculate change in lake water storage on dates without coincident satellite images, we developed rating curves between lake level and lake volume. We calculated the initial lake water volume using the lake level measurements collected by the citizen scientists and the surface area measurements derived from the Landsat and Sentinel 2 imagery. We then used a rating curve to estimate variations in volume from lake level alone.

#### 3.3.1. Calculation of Lake Water Storage

To calculate change in lake storage, or volume, over time, the lake water level measurements for each of the lakes were combined with lake surface area measurements. If a lake level and a surface area measurement fell on the same day, they were automatically paired. If there were two lake level measurements on the same day, then a mean surface area was used. If there was no lake level taken on the day a surface area measurement was taken, then the nearest lake level measurement  $\pm 1$  day was used as a pairing. If there was a lake level measurement from the day before as well as one for the day after the image was collected, we averaged the lake level measurements. If there was no matching lake surface area measurement for a lake level measurement or vice versa, those data were not used in calculating volume variations.

Lake volume change was then calculated for each date with a lake level and area pairing based on a linear equation that assumes lake volume change can be approximated by trapezoidal volume:

$$V = \frac{h}{2} \times (B_1 + B_2) \quad (1)$$

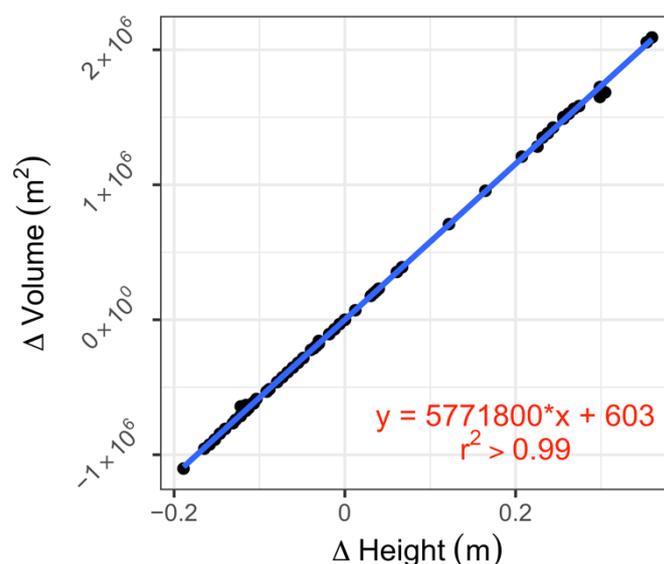
where  $V$  is volume,  $h$  is height,  $B_1$  is one base measurement of the trapezoid and  $B_2$  is the other base measurement of the trapezoid. This basic equation can then be applied to capture the change in water volume for lakes [51]:

$$\Delta V \left( \frac{t_i}{t_{i0}} \right) = \frac{[B(t_i) + B(t_{i0})]}{2} \times [h(t_i) - h(t_{i0})] \quad (2)$$

where  $\Delta V\left(\frac{t_i}{t_{i0}}\right)$  is change in volume from  $t_i$ , time of image, to  $t_{i0}$ , time of the first image in the time series.  $B(t_i)$  is the lake water surface area at the time of the image, while  $B(t_{i0})$  is the lake water surface area at the time of the first image in the time series. The variable  $h(t_i)$  is the height at time of image and  $h(t_{i0})$  is the height of the lake in the first image in the time series. Applying this equation to each level and area pair leaves us with a time series of variations in lake water volume for each lake (Figure 3c). Absolute volume cannot be calculated because lake bathymetry is unknown. We worked under the assumption that lake bathymetry remains constant throughout the study period. It is also worth noting that the trapezoidal hypothesis most likely only works for a certain range of lake level to area pairings since area grows quadratically; however, in this study, the hypsometries found are very linear, confirming the applicability of the trapezoidal hypothesis. Note that in Figure 3 the water level and storage change data are highly correlated, while the inundated area data are not. Because variations in lake area are so small (standard deviation 0.04 km<sup>2</sup>, ~3% of mean area), nearly all variations in storage are explained by increases or decreases in water level. In addition, some of the observed variation in area for this lake is likely due to errors in satellite retrievals.

### 3.3.2. Hypsometric Curve

In order to maximize data usage for each lake, we used a hypsometry approach to estimate variations in lake volume. This approach has been used by others [23,52–54]. Ogilvie et al. (2018) [52] noted that for most small lakes, change in lake water volume primarily functions as a response to changes in lake water level; we made the same observation when looking at the degree of variation of change in lake water levels, areas, and volumes in each region. To develop a level-volume change rating curve for each lake, we regressed the calculated variations in lake water volume against the corresponding lake level measurements. A linear equation was determined for this relationship (Figure 6). The equation was then applied to all lake level measurements in the lake's record to calculate a volume variation time series.

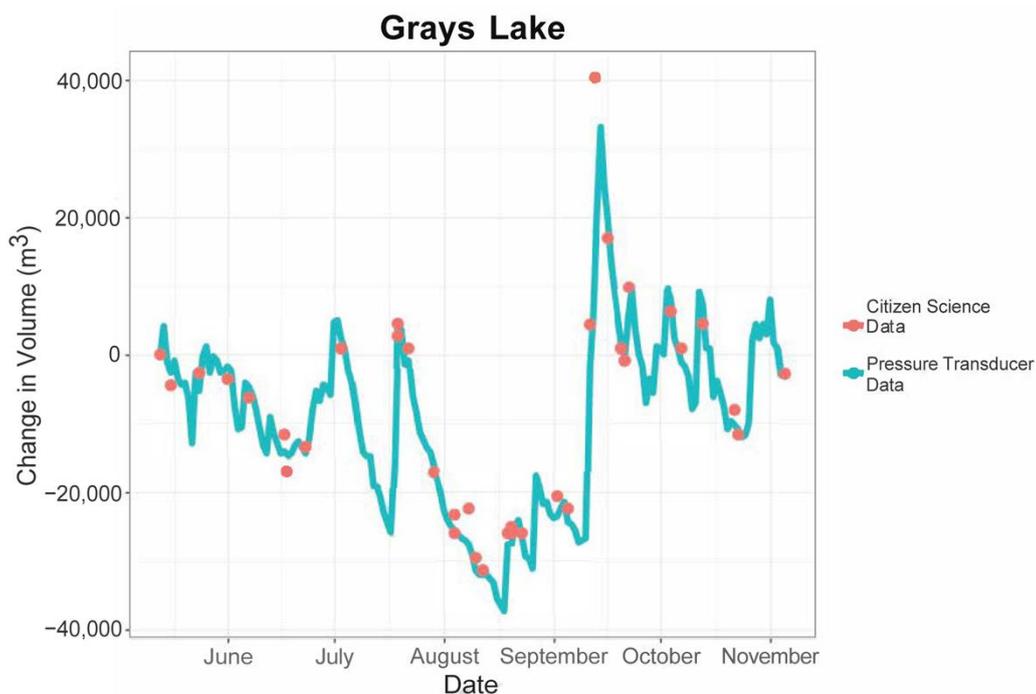


**Figure 6.** Example of the linear relationship determined between change in height and calculated change in volume to create a rating curve at Bay Tree Lake.

### 3.3.3. Propagation of Lake Height Errors into Volume Variations

Ideally, we would like to estimate errors in volume change calculations. However, doing so robustly would require bathymetric measurements that we did not have for most lakes. Instead, we evaluated how the error in the lake height measurements propagates into volume estimates. To do so, we also calculated lake volume changes using lake

level data collected from the pressure transducers using the same steps described in Sections 3.3.1 and 3.3.2. The change in volume time series for both the pressure transducer data and the citizen science data were overlaid on the same plot. The time series show similar variations, though some of the high-frequency variability is not captured by the less temporally dense citizen science data (Figure 7). We then calculated Pearson's Correlation Coefficient ( $r$ ) between the pressure transducer data and the citizen science data time series.



**Figure 7.** Time series of change in lake water volume calculated from the citizen science data versus the change in lake water volume calculated from pressure transducer data in Grays Lake, located in Illinois. This demonstrates the accuracy of the citizen science data.

### 3.4. Correlations between Change in Lake Water Storage

#### 3.4.1. Data Acquisition

Once variations in lake water storage are calculated for all lakes in all regions, we calculated the Spearman's correlation coefficient ( $\rho$ ) [55] of those changes between all pairs of lakes in each region. Lake pairs were simply identified as all possible combinations of study lakes within a study region. We chose to use Spearman's  $\rho$  because it is less sensitive than a Pearson's correlation to the assumption that relationships between paired lake water levels are linear. We also calculated Pearson's correlations (not reported here) and noted only minor differences with Spearman's  $\rho$ . Observations for each lake were matched by date, or  $\pm 1$  day if no same-date match. If there were fewer than 10 paired measurements, a correlation coefficient was not calculated, and the lake to lake pair was not used. We used a threshold of  $n = 10$  to avoid a case where a very small number of observations led to a spurious correlation. However, we noted that including all lakes does not substantially change our results. This analysis was repeated for all lake pairs in a region, no matter how close or distant, and completed separately for each of the four regions.

#### 3.4.2. Validation

In order to evaluate whether paired lake correlations based on sparse citizen science data are useful, we compared them against paired lake correlations based on much more temporally dense pressure transducer measurements. If correlation coefficients are similar

using the two different methods, then we have higher confidence in the analysis based on citizen science measurements.

### 3.5. Spatial Analysis

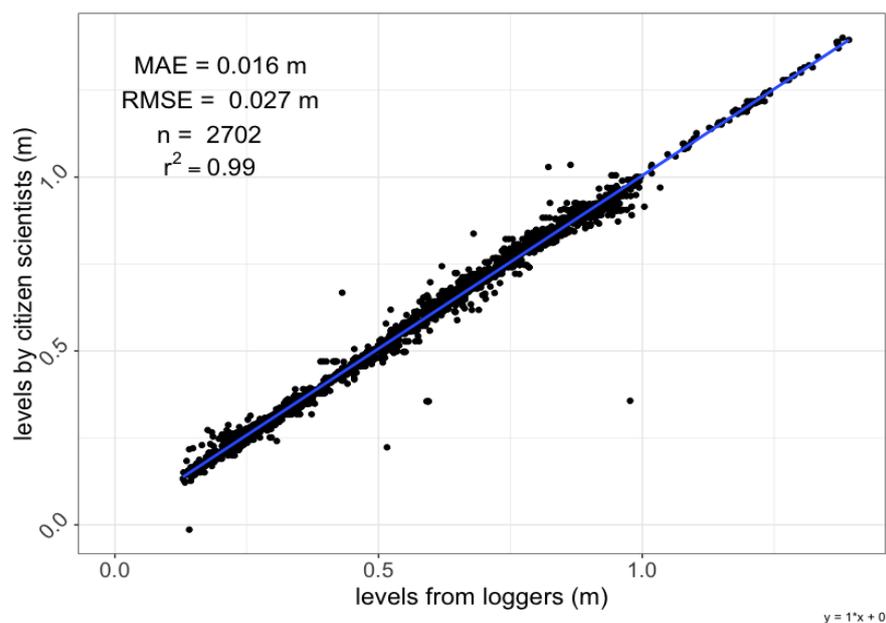
To assess whether correlations in paired lake water storage are controlled by distance, we regressed the correlation coefficients described in Section 3.4.1 against the distance between the pair of lakes associated with each correlation. We assessed whether distance is a major control on correlation using Spearman's  $\rho$  between paired lake distance and paired lake correlation coefficient.

## 4. Results

### 4.1. Validation Results

#### 4.1.1. Citizen Science Data

Comparison of lake level measurements made by citizen scientists against pressure transducers show that measurements by citizen scientists are highly accurate (Figure 8). Across 2702 corresponding lake level measurements, we observe an  $R^2$  value of  $>0.99$ , a mean absolute error (MAE) of 1.6 cm, and a root mean squared error (RMSE) of 2.7 cm. The measurement error of the pressure transducers is 0.8 cm. There are several ( $n = 9$ ) outliers in the citizen science data, and we attributed most of these to data entry errors by citizen scientists.



**Figure 8.** A scatterplot depicting the accuracy of citizen science measurements as compared to data collected by pressure transducers. The accuracy of citizen science measurements is very high, with MAE of only 1.6 cm. The accuracy of the pressure transducers themselves is  $\sim 0.8$  cm.

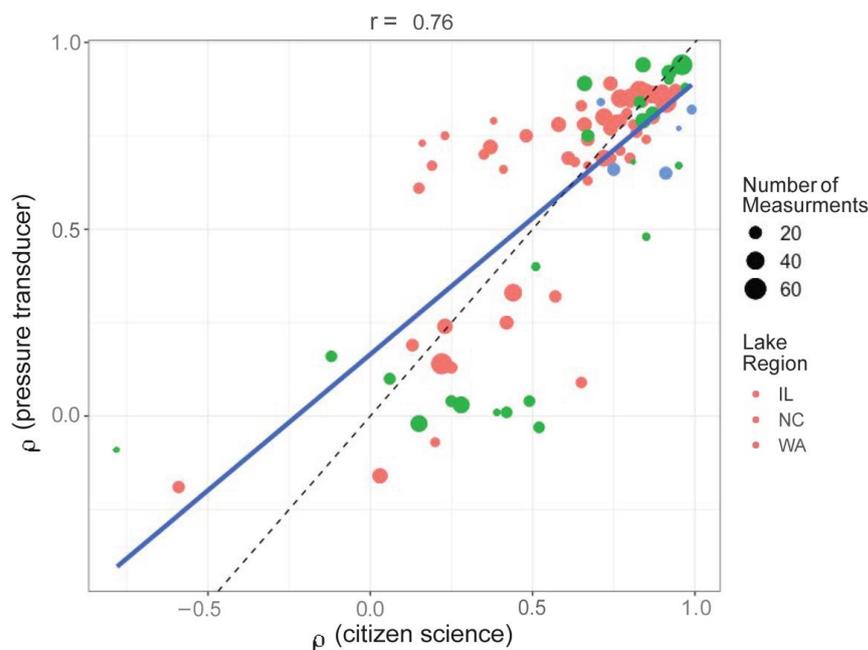
#### 4.1.2. Lake Surface Area

Comparison against manual classifications based on high-resolution Planet imagery suggests that automated Landsat and Sentinel 2 classifications are generally accurate. For Bay Tree Lake, the average surface area delineated from Planet imagery is 98.95% of the surface area delineated from Landsat or Sentinel 2 imagery (mean absolute error (MAE):  $0.066 \text{ km}^2$ ); for Hastings Lake, it is 80.19% (MAE:  $0.059 \text{ km}^2$ ); and for Lake Wenatchee, it is 103% (MAE:  $0.373 \text{ km}^2$ ). In Bay Tree Lake, the most straight-forward case, the percent difference between the Landsat or Sentinel 2 imagery and the Planet imagery is minimal. Lake Wenatchee, a more difficult case due to the topographic shadow that covers the lake at times as a result of the lake's location in the Cascade Mountains, has a slightly larger percent difference. The values of topographically shadowed pixels in the imagery

are similar to those of the water, leading to a modest overestimation of surface water in our automatic Landsat and Sentinel 2 classifications compared to Planet imagery. Hastings Lake, our representative small lake case, has the greatest percent difference. We see an underestimation of lake surface area in our automated classifications. This is most likely due to the coarse spatial resolution of the imagery compared to the lake area, especially in the case of Landsat. In general, we see a larger difference between Landsat and Sentinel 2 derived areas in small lakes because of Sentinel 2's higher spatial resolution. This difference is generally negligible. The 209 coincident Landsat and Sentinel 2 observations that occurred during our study period show a mean difference of  $-0.8\%$  ( $\sigma = 5\%$ ) in their predicted areas. We note that because the underestimation typical of small lakes is relatively consistent in time, it will likely have a small impact on the time series of water storage used for lake correlations and will, instead, represent a systematic offset. We also note that, during our study period, we observed minimal change in lake area in all study lakes. Thus, in the cases studied here, changes in lake volume are almost entirely driven by variations in water surface elevation.

#### 4.1.3. Lake Volume Correlations

In order to test propagation of errors in lake heights from citizen scientists into lake volume calculations, we compared the correlation coefficients of change in lake volume between pairs of lakes derived from the citizen science measurements against those calculated from the automated water level loggers. We see that they compare very well and are similar, though not identical (Figure 9). The Pearson's correlation coefficient is 0.76 between paired lake correlation coefficients from the two different data sources. That, plus the relative similarity of the best fit and one-to-one lines, suggests that, in general, correlation coefficients from temporally sparse citizen science data are likely reliable. The average difference in correlation coefficients calculated for the same pair of lakes using data from citizen scientists and from pressure transducers is 0.23. There is no relationship between the number of measurements collected and the difference between citizen science-based and pressure transducer-based correlation coefficients in any of the lake regions.

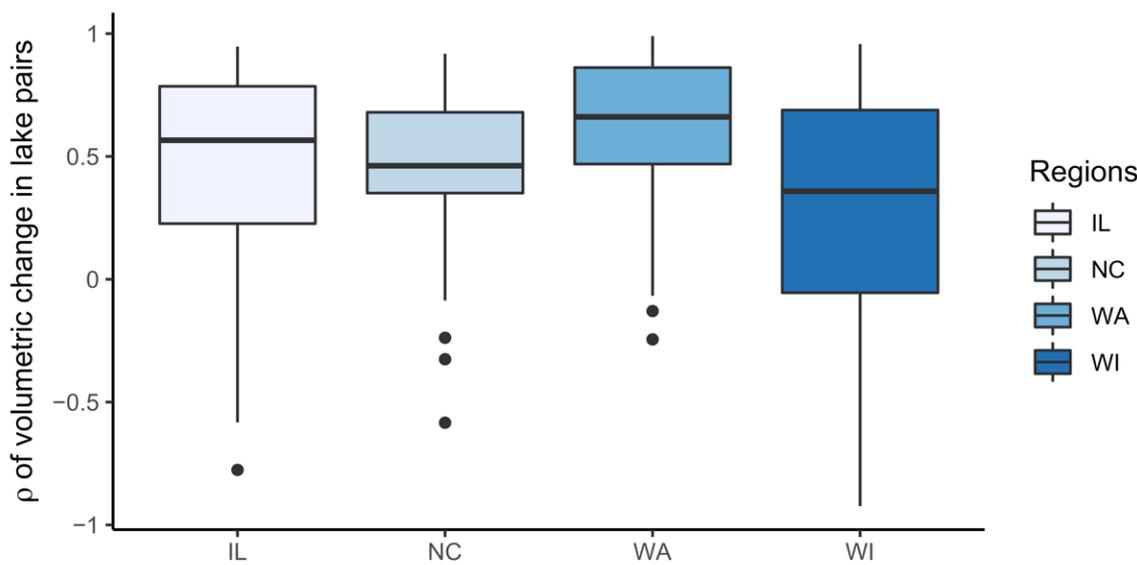


**Figure 9.** Correlation coefficients between lake pairs in a region of the citizen-scientist-based change in volume versus the correlation coefficients between lake pairs in a region of the pressure-transducer-based change in volume. A strong correlation between the two is seen. The dashed line is a one-to-one line and the blue line is the line of best fit.

## 4.2. Variations in Lake Water Storage

### 4.2.1. Regional Coherence of Changes in Lake Water Volume

On average, water storage variations in pairs of lakes within each of the study regions are positively correlated (average  $\rho = 0.48$ , Figure 10). The distributions of correlations are similar for all four regions, though interestingly they are somewhat higher in Washington, the only region with notable topography, and lowest in North Carolina and Wisconsin. There is a substantial spread in the degree of correlation, with some pairs of lakes highly correlated and others uncorrelated.

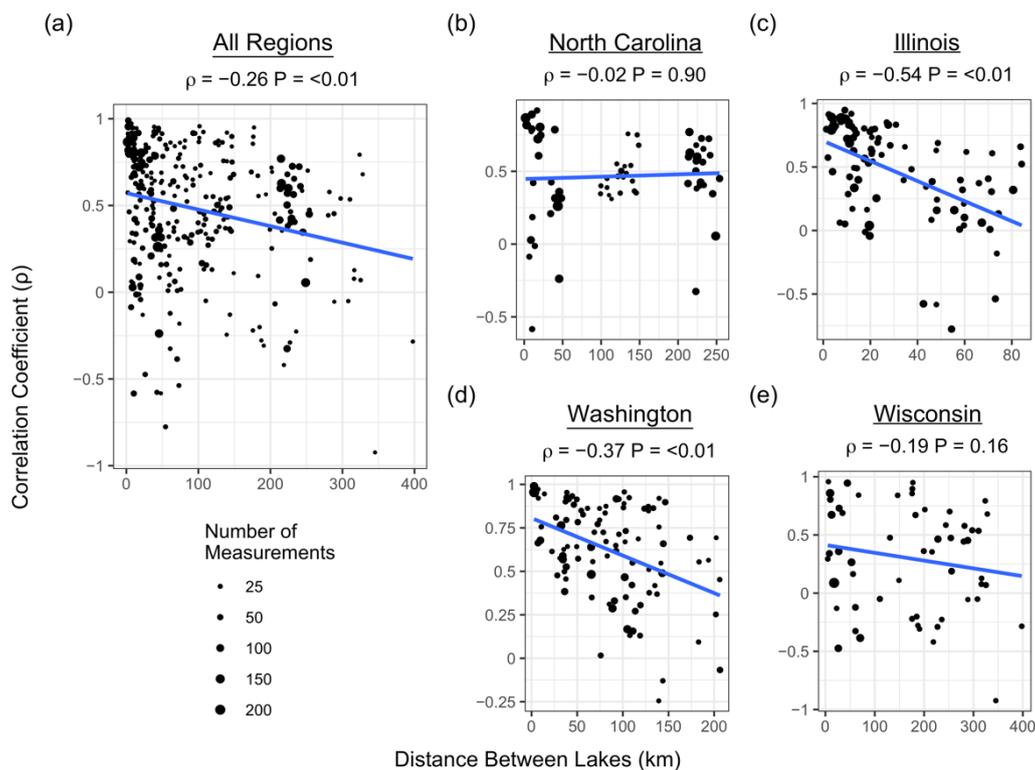


**Figure 10.** The distribution of correlation coefficients ( $\rho$ ) between change in lake water volume for all lake pairs in each region. The boxes represent the distribution of the correlation coefficients to the upper and lower quartiles. The whiskers represent the full range of the correlation coefficients to the minimum and maximum, with the points being outliers in the data set.

### 4.2.2. Correlations and Relationship with Distance

There is a weak negative relationship between distance and paired lake correlation in two regions (Washington and Illinois) and no significant relationship in the other two regions (North Carolina and Wisconsin) (Figure 11b–e). When lakes in all regions are combined, there is a weak but statistically significant negative relationship between distance and paired lake correlation (Figure 11a).

We tested whether the correlations with distance are influenced by the number of paired measurements and found no relationship in all of the regions combined or in any region individually. We also tested whether the variability in paired lake correlation increases with paired lake distance. There is a slight but significant ( $p < 0.01$ ) negative relationship in North Carolina and a positive relationship in Illinois. Results for other regions, and for all regions together, show no significant relationship.



**Figure 11.** Relationship between the Spearman's correlation coefficient ( $\rho$ ) of change in lake water storage between paired lakes in a region and the distance between those two lakes: all regions (a), North Carolina (b), Illinois (c), Washington (d), and Wisconsin (e). Overall, there is a negative linear relationship between correlation and distance. This trend is different region by region.

## 5. Discussion

First and foremost, this study demonstrates that two still-evolving approaches, citizen science and optical remote sensing of lake area, can be combined to accurately monitor changes in lake water storage over time. Citizen science lake level data are nearly as accurate as pressure transducer data, with the primary difference being that the pressure transducer data capture a near-continuous time series. The lake surface areas automatically calculated from Landsat and Sentinel 2 imagery are quite similar to the areas manually calculated from Planet imagery in most cases. For simple cases, like Bay Tree Lake (Figure 5a), lake surface areas compare very well to manually digitized maps made from same-day, high-resolution Planet imagery. For cases where topographic shadows are important, like Lake Wenatchee (Figure 5c), lake surface area is overestimated in the automated classifications, likely due to topographic shadow pixels misclassified as water. For small lakes, like Hastings Lake (Figure 5b), we see an underestimation of lake surface area in our automated classifications, likely due to the fact that a large portion of the lake is covered by mixed water-land pixels. In general, the Planet imagery has a finer spatial resolution than that of the Landsat and Sentinel 2 imagery, meaning that Planet imagery can capture more heterogeneity and detect subtle changes in inundation extent not apparent in Landsat or Sentinel 2 imagery [11]. It is important to note that there could still be errors in the Planet surface area delineations due to user error. Ultimately, there is no available dataset that represents absolute truth when measuring surface area through satellite imagery [51]. Overall, our results suggest that citizen science data and optical satellite imagery accurately capture changes in lake water storage.

The study presented here begins to provide a baseline in the level of correlation we may expect to see in the changes of lake water storage between lakes in a region, at least on timescales of 1–2 years. Perhaps surprisingly, the average correlation between

pairs of lakes within our four study regions is only moderately positive ( $\rho = 0.48$ ) and highly variable. This result suggests that measurements made in one lake provide some regional information, but, at least in the regions we examined, there is also a great deal of lake-to-lake variability in water storage time series. This result contrasts with findings by Watras et al. (2014) [10], in four lakes in northern Wisconsin, who found that, over the last 70 years, water level variations were controlled by a near decadal oscillation (~13 years), independent of lake or aquifer hydrology and that the oscillation has been driven by net atmospheric flux as well as level-dependent outflow. In contrast, Euliss and Mushet (1996) [9], who studied the controlling factors of 36 small wetland waterbodies in the prairie-pothole region in North and South Dakota, found that during one spring and summer season, water level fluctuations were influenced by lake-to-lake variability. Variations were greater in wetlands with more agricultural activity, rather than grasslands, and water level fluctuations in semi-permanent wetlands were more stable due to the inputs of groundwater, whereas seasonal and temporary wetlands solely depended on runoff. Perales et al. [56] studied changes in lake water storage across 47 lakes in Northern Wisconsin during a five-year drought between 2005 and 2010. They found that most lakes lost water during this period, but the degree of loss varied considerably; much of this variability could be explained by lake-specific factors, including soil characteristics, elevation, the presence of littoral wetlands, and whether the lake was a drainage or seepage lake. Correlations between surface areas in small lakes and their sub-seasonal dynamics have also been studied by Cooley et al. (2019) [11] across Alaska and the Canadian Shield. They noted that, as a whole, lake surface area declined across all study regions, with some distinct localized exceptions, which primarily reflected negative net summer atmospheric flux in North America in 2017. They also note that changes in surface area are primarily driven by lake level fluctuation, not shoreline vegetation growth. As the body of work exploring controls on lake storage fluctuations remains limited, further work needs to be done to determine what localized and regional factors control observed patterns. Unlike Watras et al. [10], our study spans up to two years, and hence we do not capture decadal-scale patterns. The patterns we see are most likely yearly cyclical patterns and sub-seasonal dynamics, like in Cooley et al. [11], or localized weather events and also likely represent a small fraction of the total change in storage lakes would experience at longer time scales. A future study would benefit from at least 10 years of data, which would allow better comparison to long-term studies such as the work of Watras et al. [10].

We observed only a limited correlation between paired lake Spearman's  $\rho$  and paired lake distance. The hypothesis that as the distance between lakes increases, the correlation of change in lake water storage between those two lakes decreases is an intuitive one. However, the results from this study provide only limited evidence for this relationship. When aggregated together across all regions, our data show a significant but weak correlation ( $\rho = -0.26$ ). Two regions, Illinois and Washington, show slightly stronger negative relationships between correlation and distance ( $\rho = -0.37$  and  $-0.54$ , respectively), while North Carolina and Wisconsin show no significant relationship between correlation and distance. We also tested for any relationship between the variation in paired lake correlation (as measured by the absolute value of residuals) and distance but found no statistically significant relationship in any region. All of this suggests that, when thinking about the spatial synchrony between the change in lake water storage of the lakes, there are other factors in play that are more important than distance [57].

With such limited evidence supporting distance as a primary driver of the spatial synchrony of lake volume change, other local factors must be in play that matter more in governing lake volume change. Watras et al. [10] observed stronger lake-to-lake correlations than we did, which may reflect the possibility that regional drivers are more important at longer time scales. However, they also examined a much smaller number of lakes (focused mostly on one pair in Northern Wisconsin, along with the Great Lakes and groundwater well records). The Perales study referenced above [56] from Northern Wisconsin found that lake and watershed characteristics (e.g., seepage versus drainage lakes, conductivity,

elevation, soil permeability, percent wetland buffer) were good predictors of water level change, with very different magnitudes of response to drought among some neighboring lakes. In addition, the lakes in our study experienced some major and somewhat unusual events during the study period—for example, the large spike in water levels in Salters Lake (Figure 3) is the result of Hurricane Florence, which impacted some North Carolina lakes much more than others, pointing to the local influence of the storm occurring at shorter time scales. Our results suggest that, at least in the places and over the time periods studied here, both regional and lake-specific factors control variations in lake water storage.

While the results of this study represent the first analysis of lake-to-lake water storage correlation across many lakes, there is still a great deal that we do not understand about the drivers of lake water levels. Exploring what these drivers might be is of great interest but is beyond the scope of this paper. Many relevant questions will be addressed after the launch of the upcoming SWOT satellite mission in 2022 [58]. Since SWOT will measure lakes as small as 250 m by 250 m, it will provide a time series of variations in water storage for millions of lakes globally. Continuing measurements from LOCSS and other sources based on citizen science and satellite imaging will provide key ground-truthing data for SWOT, especially in areas where on-the-ground measurements are limited.

Overall, this study addresses a gap in understanding regional patterns of lake water storage. We suggest that change in lake water volume in small lakes can be monitored accurately by combining measurements of water level from citizen scientists with optical satellite imaging of lake area. Assessment of relationships in water volume between pairs of lakes, on average, produce moderate correlations, suggesting that both regional patterns and lake-specific factors are important drivers of variations in lake water storage. To our knowledge, this is the first study to assess the degree of correlation in water storage time series among many small lakes in different regions. It presents the first results that can guide future hypotheses and data collection efforts. We hope that continued data collection by citizen scientists and satellites can further our understanding of natural lakes, including the usually overlooked small ones, especially in a time of rapid environmental change.

**Author Contributions:** Conceptualization, S.L., T.M.P., F.H., S.G., G.M.P., S.K.Y., M.R. and C.H.; Data curation, S.L., X.Y., C.H. and S.N.T.; Formal analysis, S.L. and X.Y.; Funding acquisition, T.M.P., F.H., S.G., G.M.P., S.K.Y. and J.-F.C.; Investigation, S.L., H.T., D.K. and D.W.; Methodology, S.L., T.M.P., G.M.P., X.Y., C.H. and S.N.T.; Project administration, T.M.P.; Software, S.G., M.A.U. and D.H.L.; Supervision, T.M.P.; Validation, S.L.; Visualization, S.L.; Writing—original draft, S.L.; Writing—review and editing, T.M.P., F.H., S.G., G.M.P., S.K.Y., M.R., X.Y., J.-F.C., C.H., D.H.L. and S.N.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by a grant from the NASA Citizen Science for Earth Systems Program (#80NSSC18M0099), PI: Tamlin Pavelsky. S. Little was partially supported by the Preston Jones and Mary Elizabeth Frances Dean Martin Fellowship Fund of the University of North Carolina Department of Geological Sciences.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data and code used within this analysis is publicly available on GitHub at <https://github.com/GlobalHydrologyLab/CitSci>.

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

## Appendix A

Table A1. Individual Lake Characteristics.

Lake Name	Region	Approx. Size (km <sup>2</sup> )	Date of First Lake Level Measurement	Date of Last Lake Level Measurement Used	Number of Lake Level Measurements	Number of Satellite Images (1 January 2015–1 February 2020)
Bay Tree Lake	NC	5.8	18 April 2017	20 January 2020	325	130
Catfish Lake	NC	3.8	19 September 2017	22 May 2019	37	196
Great Lake	NC	11.3	19 September 2017	10 January 2020	42	127
Horsepen Lake	NC	1.2	21 September 2017	20 January 2020	201	130
Jones Lake	NC	0.9	18 April 2017	1 February 2020	149	130
Lake Mattamuskeet West	NC	57.6	1 June 2017	27 January 2020	137	129
Phelps Lake	NC	64.5	1 June 2017	25 January 2020	295	100
Salters Lake	NC	1.3	18 April 2017	1 February 2020	137	173
Singletary Lake	NC	2.3	18 April 2017	30 January 2020	97	130
Lake Mattamuskeet East	NC	105.7	1 June 2017	27 January 2020	134	129
Lake Waccamaw	NC	35.8	18 April 2017	19 January 2020	260	130
White Lake	NC	4.3	18 April 2017	25 December 2019	122	130
Beaver Lake	WA	0.3	10 September 2018	25 January 2020	56	75
Walupt Lake	WA	1.5	12 June 2019	30 October 2019	15	109
Bosworth Lake	WA	0.4	12 September 2018	28 January 2020	25	122
Lake Cassidy	WA	0.5	12 September 2018	26 October 2019	36	122
Coldwater Lake	WA	3.0	11 June 2019	1 February 2020	54	75
Deep Lake	WA	0.2	10 June 2019	28 November 2019	39	74
Dog Lake	WA	0.2	13 June 2019	30 October 2019	9	144
Echo Lake	WA	0.06	12 September 2018	30 October 2019	50	132
Fish Lake	WA	2.0	12 June 2019	24 January 2020	24	95
Lake Howard	WA	0.1	11 June 2019	10 December 2019	10	155
Lake Ki	WA	0.4	11 June 2019	25 November 2019	8	154
Lake Lawrence	WA	1.3	12 September 2018	1 February 2020	66	135
Leech Lake	WA	0.2	13 June 2019	5 December 2019	30	144
Phantom Lake	WA	0.3	10 September 2018	31 January 2020	166	100
Crabapple Lake	WA	0.1	11 June 2019	26 August 2019	8	134
Lake Roesiger	WA	1.4	10 September 2018	28 January 2020	29	121
Lake Sammamish	WA	19.6	11 September 2018	30 January 2020	144	98
Steel Lake	WA	0.2	10 June 2019	31 January 2020	110	99
Flowing Lake	WA	0.5	10 September 2018	28 January 2020	46	122
North Lake	WA	0.2	10 June 2019	1 February 2020	50	99
Lake Wenatchee	WA	9.7	10 September 2018	25 January 2020	60	95
Lake Martha	WA	0.2	13 September 2018	26 January 2020	74	154

Table A1. Cont.

Lake Name	Region	Approx. Size (km <sup>2</sup> )	Date of First Lake Level Measurement	Date of Last Lake Level Measurement Used	Number of Lake Level Measurements	Number of Satellite Images (1 January 2015–1 February 2020)
Diamond Lake	IL	0.6	16 May 2019	23 November 2019	26	63
Lake Defiance	IL	0.3	15 May 2019	26 December 2019	56	63
Gages Lake	IL	0.6	13 May 2019	23 November 2019	26	91
Hastings Lake	IL	0.3	14 May 2019	11 January 2020	97	112
Herrick Lake	IL	0.05	23 May 2019	31 January 2020	89	88
Harrier Lake	IL	0.08	29 May 2019	21 November 2019	28	63
Lake Killarney	IL	0.2	15 May 2019	11 December 2019	22	63
Highland Lake	IL	0.4	14 May 2019	21 September 2019	23	91
East Loon Lake	IL	0.8	14 May 2019	15 January 2020	41	87
West Loon Lake	IL	0.7	14 May 2015	15 January 2020	NA	87
Lily Lake	IL	0.4	15 May 2019	17 December 2019	26	88
McCullom Lake	IL	1.0	15 May 2019	5 November 2019	16	102
Maple Lake	IL	0.2	16 May 2019	11 January 2020	26	102
Deep Quarry Lake	IL	0.1	29 May 19	5 November 2019	38	88
Round Lake	IL	1.0	14 May 2019	26 December 2019	64	91
Silver Lake	IL	0.2	16 May 19	1 January 2020	33	88
Timber Lake	IL	0.3	14 May 2019	26 December 2019	45	87
Grays Lake	IL	0.3	13 May 19	11 January 2020	42	91
Archibald Lake	WI	1.6	26 April 2016	07 May 2019	92	84
Axhandle Lake	WI	0.3	13 July 2015	5 October 2017	29	92
Bass Lake	WI	0.3	13 August 2015	20 October 2016	39	58
Bear Lake	WI	0.1	22 June 2017	3 November 2017	12	107
Big Twin Lake	WI	0.3	15 July 2016	04 May 2019	59	84
Clear Lake	WI	0.3	21 June 2017	30 November 2017	24	73
Crystal Lake	WI	0.1	14 July 2017	21 September 2017	10	76
Deep Lake	WI	0.1	17 June 2017	06 July 2017	3	128
Des Moines Lake	WI	0.9	18 April 2017	2 November 2017	28	152
Duck Lake	WI	0.5	5 July 2016	23 September 2019	78	137
Grindle Lake	WI	0.2	9 November 2015	10 August 2019	91	112
Horseshoe Lake	WI	1.6	18 August 2015	6 June 2019	69	159
Kentuck Lake	WI	4.0	05 July 2019	5 July 2019	1	101
Kilby Lake	WI	0.2	20 May 2017	1 October 2017	40	120
Lake Five	WI	0.4	23 May 2017	3 October 2017	19	119
Long Lake	WI	1.1	18 August 2015	19 August 2019	57	159
Loon Lake	WI	0.1	14 August 2015	20 October 2016	20	89
Mann Lake	WI	1.0	21 May 2018	12 May 2018	1	104

Table A1. Cont.

Lake Name	Region	Approx. Size (km <sup>2</sup> )	Date of First Lake Level Measurement	Date of Last Lake Level Measurement Used	Number of Lake Level Measurements	Number of Satellite Images (1 January 2015–1 February 2020)
Moose Lake	WI	0.5	6 August 2016	29 September 2018	5	94
Parker Lake	WI	0.2	10 June 2017	16 September 2017	14	128
Paya Lake	WI	0.4	2 November 2015	20 September 2019	64	88
Phantom Lake	WI	0.2	19 August 2015	6 November 2015	14	58
Poplar Lake	WI	0.5	30 June 2018	11 September 2019	9	159
Sand Lake	WI	3.6	18 April 2017	26 July 2017	15	120
South Neva Lake	WI	0.1	30 August 2015	21 November 2016	23	152
Spur Lake	WI	0.5	29 July 2019	24 September 2019	12	105
Stratton Lake	WI	0.3	11 July 2017	21 September 2017	10	76
Summit Lake	WI	1.1	5 April 2019	6 September 2019	17	137
Twin Lakes	WI	0.05	11 June 2017	1 October 2017	16	124
Underwood Lake	WI	0.2	24 August 2015	7 May 2019	9	165
Wheeler Lake	WI	1.1	26 August 2015	7 May 2019	46	84
Wolf Lake	WI	0.1	22 June 2017	3 November 2017	13	98

## References

1. MEA. *Millennium Ecosystem Assessment Global Assessment Reports*; Island Press: Washington, DC, USA, 2005.
2. Adrian, R.; O'Reilly, C.M.; Zagarese, H.; Baines, S.B.; Hessen, D.O.; Keller, W.; Livingstone, D.M.; Sommaruga, R.; Straile, D.; Van Donk, E.; et al. Lakes as sentinels of climate change. *Limnol. Oceanogr.* **2009**, *54*, 2283–2297. [\[CrossRef\]](#)
3. Williamson, C.; Saros, J.; Vincent, W.; Smol, J. Lakes and reservoirs as sentinels, integrators, and regulators of climate change. *Limnol. Oceanogr.* **2009**, *54*, 2273–2282. [\[CrossRef\]](#)
4. IPCC. *Intergovernmental Panel on Climate Change Fifth Assessment Report*, 5th ed.; The Intergovernmental Panel on Climate Change: Washington, DC, USA, 2014.
5. Wang, J.; Song, C.; Reager, J.; Yao, F.; Familietti, J.; Sheng, Y.; MacDonald, G.; Brun, F.; Schmied, H.M.; Marston, R.A.; et al. Recent global decline in endorheic basin water storages. *Nat. Geosci.* **2018**, *11*, 929–932. [\[CrossRef\]](#)
6. Downing, J.A.; Prairie, Y.T.; Cole, J.J.; Duarte, C.M.; Tranvik, L.J.; Striegl, R.G.; McDowell, W.H.; Kortelainen, P.; Caraco, N.F.; Melack, J.M.; et al. The global abundance and size distribution of lakes, ponds, and impoundments. *Limnol. Oceanogr.* **2006**, *51*, 2388–2397. [\[CrossRef\]](#)
7. Stanley, E.H.; Collins, S.M.; Lotting, N.R.; Oliver, S.K.; Webster, K.E.; Cheruvilil, K.S.; Soranno, P.A. Biases in lake water quality sampling and implications for macroscale research. *Limnol. Oceanogr.* **2019**, *64*, 1572–1585. [\[CrossRef\]](#)
8. Messenger, M.L.; Lehner, B.; Grill, G.; Nedeva, I.; Schmitt, O. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nat. Commun.* **2016**, *7*, 13603. [\[CrossRef\]](#)
9. Euliss, N.H.; Mushet, D.M. Water-level fluctuation in wetlands as a function of landscape condition in the prairie pothole region. *Wetlands* **1996**, *16*, 587–593. [\[CrossRef\]](#)
10. Watras, C.J.; Read, J.S.; Holoman, K.D.; Liu, Z.; Song, Y.-Y.; Watras, A.J.; Morgan, S.; Stanely, E.H. Decadal oscillation of lakes and aquifers in the upper Great Lakes region of North America: Hydroclimatic implications. *Geophys. Res. Lett.* **2014**, *41*, 456–462. [\[CrossRef\]](#)
11. Cooley, S.W.; Smith, L.C.; Ryan, J.C.; Pitcher, L.H.; Pavelsky, T.M. Arctic-Boreal Lake Dynamics Revealed Using CubeSat Imagery. *Geophys. Res. Lett.* **2019**, *46*, 2111–2120. [\[CrossRef\]](#)
12. Lei, Y.; Yao, T.; Bird, B.; Yang, K.; Zhai, J.; Sheng, Y. Coherent lake growth on the central Tibetan Plateau since the 1970s: Characterization and attribution. *J. Hydrol.* **2013**, *483*, 61–67. [\[CrossRef\]](#)
13. Zhang, G.; Yao, T.; Shum, C.K.; Yi, K.; Yang, K.; Xie, H.; Feng, W.; Bolch, T.; Wang, L.; Behrangi, A.; et al. Lake volume and groundwater storage variations in Tibetan Plateau's endorheic basin. *Geophys. Res. Lett.* **2017**, *44*, 5550–5560. [\[CrossRef\]](#)
14. Qiao, B.; Zhu, L.; Yang, R. Temporal-spatial differences in lake water storage changes and their links to climate change throughout the Tibetan Plateau. *Remote Sens. Environ.* **2019**, *222*, 232–243. [\[CrossRef\]](#)

15. Hanson, P.C.; Carpenter, S.R.; Cardille, J.A.; Coe, M.T.; Winslow, L.A. Small lakes dominate a random sample of regional lake characteristics. *Freshw. Biol.* **2007**, *52*, 814–822. [CrossRef]
16. U.S. Geological Survey. *Updated Daily. USGS Water Data for the Nation: U.S. Geological Survey National Water Information System Database*; U.S. Geological Survey: Reston, VA, USA, 2020. Available online: <https://waterdata.usgs.gov/nwis> (accessed on 1 December 2018).
17. Lake Level Minnesota. Available online: <https://www.dnr.state.mn.us/climate/waterlevels/lakes/index.html> (accessed on 27 January 2019).
18. How to Get the Numbers. Available online: <https://waterdatafortexas.org/reservoirs/methodology> (accessed on 10 February 2020).
19. Shiklomanov, A.I.; Lammers, R.B.; Vörösmarty, C.J. Widespread decline in hydrological monitoring threatens Pan-Arctic research. *Eos Trans. Am. Geophys. Union* **2002**, *83*, 13–16. [CrossRef]
20. IAHS Ad Hoc Group on Global Water Data Sets. Global water data: A newly endangered species. *Eos Trans. Am. Geophys. Union* **2001**, *8*, 54–58. [CrossRef]
21. Stokstad, E. Scarcity of rain, stream gages threatens forecasts. *Science* **1999**, *285*, 1199–1200. [CrossRef]
22. Fekete, B.M.; Roberts, R.D.; Kumagai, M.; Nachtnebel, H.-P.; Odada, E.; Zhulidov, A.V. Time for in situ renaissance. *Science* **2015**, *14*, 685–686. [CrossRef]
23. Crétaux, J.-F.; Abarca-del-Río, R.; Bergé-Nguyen, M.; Arsen, A.; Drolon, V.; Clos, G.; Maisongrande, P. Lake Volume Monitoring from Space. *Surv. Geophys.* **2016**, *37*, 269–305. [CrossRef]
24. Baup, F.; Frappart, F.; Maubant, J. Combining high-resolution satellite images and altimetry to estimate the volume of small lakes. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 2007–2020. [CrossRef]
25. Kleinerherenbrink, M.; Naeije, M.; Slobbe, C.; Egido, A.; Smith, W. The performance of CryoSat-2 fully-focused SAR for inland water-level estimation. *Remote Sens. Environ.* **2020**, *237*, 111589. [CrossRef]
26. Arsen, A.; Crétaux, J.; Abarca-del-Río, R. Use of SARAL/AltiKa over Mountainous Lakes, Intercomparison with Envisat Mission. *Mar. Geod.* **2015**, *38*, 534–548. [CrossRef]
27. Hughes, D.A. Comparison of satellite rainfall data with observations from gauging station networks. *J. Hydrol.* **2006**, *327*, 399–410. [CrossRef]
28. Alsdorf, D.E.; Birkett, C.M.; Dunne, T.; Melack, J.; Hess, L. Water level changes in a large Amazon lake measured with spaceborne radar interferometry and altimetry. *Geophys. Res. Lett.* **2001**, *28*, 2671–2674. [CrossRef]
29. Kite, G.; Pietroniro, A. Remote sensing of surface water. *Remote Sens. Hydrol. Water Manag.* **2000**, *42*, 217–238.
30. Peckel, 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] [PubMed]
31. Sheng, Y.; Song, C.; Wang, J.; Lyons, E.A.; Knox, B.R.; Cox, J.S.; Gao, F. Representative lake water extent mapping at continental scales using multi-temporal Landsat-8 imagery. *Remote Sens. Environ.* **2016**, *185*, 129–141. [CrossRef]
32. Lowery, C.S.; Fienen, M.N. CrowdHydrology: Crowdsourcing Hydrologic Data and Engaging Citizen Scientists. *Groundwater* **2012**, *51*, 151–156. [CrossRef]
33. Buytaert, W.; Zulkafli, Z.; Grainger, S.; Acosta, L.; Alemie, T.C.; Bastiaensen, J.; Bievre, B.D.; Bhusal, J.; Clark, J.; Dewulf, A.; et al. Citizen Science in hydrology and water resources: Opportunities for knowledge generation, ecosystem service management, and sustainable development. *Front. Earth Sci.* **2014**, *2*, 26. [CrossRef]
34. Lowery, C.S.; Fienen, M.N.; Hall, D.D.; Stepenuck, K.F. Growing Pains of Crowdsourced Stream Stage Monitoring Using Mobile Phones: The Development of CrowdHydrology. *Front. Earth Sci.* **2019**, *7*, 164–169. [CrossRef]
35. Strobl, B.; Etter, S.; van Meerveld, I.; Seibert, J. The CrowdWater game: A playful way to improve the accuracy of crowdsourced water level class data. *PLoS ONE* **2019**, *14*, e0222579. [CrossRef]
36. McDonald, C.P.; Rover, J.A.; Stets, E.G.; Striegl, R.G. The regional abundance and size distribution of lakes and reservoirs in the United States and implications for estimates of the global lake extent. *Limnol. Oceanogr.* **2012**, *57*, 597–606. [CrossRef]
37. Lake Observations by Citizen Scientists and Satellites. Available online: [www.locss.org](http://www.locss.org) (accessed on 15 August 2018).
38. Citizen Lake Monitoring Network. Available online: <https://dnr.wi.gov/lakes/clmn/> (accessed on 26 October 2018).
39. Pavelsky, T.; Ghafoor, S.; Hossain, F.; Parkins, G.; Yelton, S.; Little, S.; Topp, S.; Rodgers, M.; Yang, X. Monitoring the World's Lakes: Progress from Citizen Science and Remote Sensing. 2019. Available online: <https://pubs.awma.org/flip/EM-Nov-2019/pavelsky.pdf> (accessed on 1 March 2021).
40. Pitcher, L.H.; Smith, L.C.; Cooley, S.W.; Zaino, A.; Carlson, R.; Pettit, J.; Gleason, C.J.; Minear, T.; Fayne, J.V.; Harlan, M.E.; et al. Advancing field-based GNSS surveying for validation of remotely sensed water surface elevation products. *Front. Earth Sci.* **2020**, *8*, 278. [CrossRef]
41. Cohen, J.P. Citizen science: Can volunteers do real research? *BioScience* **2008**, *58*, 192. [CrossRef]
42. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]
43. Kumar, L.; Mutanga, O. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. *Remote Sens.* **2018**, *10*, 1509. [CrossRef]
44. Simley, J.D.; Carswell, W.J., Jr. *The National Map—Hydrography: U.S. Geological Survey Fact Sheet 2009–3054*; U.S. Department of the Interior, U.S. Geological Survey: Washington, DC, USA, 2009.

45. Jones, J.W. Improvised Automated Detection of Subpixel-Scale Inundation—Revised Dynamic Surface Water Extent (DSWE) Partial Surface Water Tests. *Remote Sens.* **2019**, *11*, 374. [[CrossRef](#)]
46. Zhang, H.K.; Roy, D.P.; Yan, L.; Li, Z.; Huang, H.; Vermote, E.; Skakun, S.; Roger, J.C. Characterization of Sentinel-2A and Landsat-8 top of atmosphere, surface, and nadir NRDF adjusted reflectance and NDVI differences. *Remote Sens. Environ.* **2018**, *215*, 482–494. [[CrossRef](#)]
47. Ahmad, S.K.; Hossain, F.; Eldardiry, H.; Pavelsky, T.M. A Fusion Approach for Water Area Classification using Visible, Near Infrared and Synthetic Aperture Radar for South Asian Conditions. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 2471–2480. [[CrossRef](#)]
48. Foga, S.; Scaramuzza, P.L.; Guo, S.; Zhu, Z.; Dilley, R.D., Jr.; Beckmann, T.; Laue, B. Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sens. Environ.* **2017**, *194*, 379–390. [[CrossRef](#)]
49. Planet Team. *Planet Application Program Interface: In Space for Life on Earth*; Planet Team: San Francisco, CA, USA, 2017; Available online: <https://api.planet.com> (accessed on 1 March 2021).
50. Cooley, S.W.; Smith, L.C.; Stephan, L.; Mascaro, J. Tracking Dynamic Northern Surface Water Change with High-Frequency Planet CubeSat Imagery. *Remote Sens.* **2017**, *9*, 1306. [[CrossRef](#)]
51. Quéllec, M.; Crétaux, J.-F. *Progress in Lake Water Storage Algorithms*; SWOT Science Team: Pasadena, CA, USA, 2018.
52. Ogilvie, A.; Belaud, G.; Massuel, S.; Mulligan, M.; Le Goulven, P.; Calvez, R. Surface water monitoring in small water bodies: Potential and limits of multi-sensor Landsat time series. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 4349–4380. [[CrossRef](#)]
53. Medina, C.; Gomez-Enri, J.; Alonso, J.J.; Villares, P. Water volume variations in Lake Izabal (Guatemala) from in situ measurements and ENVISAT Radar Altimeter (RA-2) and Advanced Synthetic Aperture Radar (ASAR) data products. *J. Hydrol.* **2009**, *382*, 34–48. [[CrossRef](#)]
54. Liebe, J.; van de Giesen, N.; Andrenini, M. Estimation of small reservoir storage capacities in a semi-arid environment. *Phys. Chem. Earth* **2005**, *30*, 448–454. [[CrossRef](#)]
55. Spearman, C. The Proof and Measurement of Association between Two Things. *Am. J. Psychol.* **1904**, *15*, 72–101. [[CrossRef](#)]
56. Perales, K.M.; Hein, C.L.; Lottig, N.R.; Zanden, M.J.V. Lake water level response to drought in a lake-rich region explained by lake and landscape characteristics. *Can. J. Fish. Aquat. Sci.* **2020**, *77*, 1836–1845. [[CrossRef](#)]
57. Lottig, N.R.; Tan, P.-N.; Wagner, T.; Cheruvilil, K.S.; Soranno, P.A.; Stanley, E.H.; Scott, C.E.; Stow, C.A.; Yuan, S. Macroscale patterns of synchrony identify complex relationships among spatial and temporal ecosystem drivers. *Ecosphere* **2017**, *8*, e02024. [[CrossRef](#)]
58. Biancamaria, S.; Lettenmaier, D.P.; Pavelsky, T.M. The SWOT Mission and Its Capabilities for Land Hydrology. *Surv. Geophys.* **2016**, *37*, 307–337. [[CrossRef](#)]