

Article



A Comparative Study of Water Indices and Image Classification Algorithms for Mapping Inland Surface Water Bodies Using Landsat Imagery

Feifei Pan^{1,*}, Xiaohuan Xi² and Cheng Wang²

- ¹ Department of Geography and the Environment, University of North Texas, Denton, TX 76203, USA
- ² Key Laboratory of Digital Earth, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China; xixh@radi.ac.cn (X.X.); wangcheng@radi.ac.cn (C.W.)
- * Correspondence: feifei.pan@unt.edu; Tel.: +1-940-369-5109

Received: 16 April 2020; Accepted: 16 May 2020; Published: 18 May 2020



Abstract: A comparative study of water indices and image classification algorithms for mapping inland water bodies using Landsat imagery was carried out through obtaining 24 high-resolution (\leq 5 m) and cloud-free images archived in Google Earth with the same (or ± 1 day) acquisition dates as the Landsat-8 OLI images over 24 selected lakes across the globe, and developing a method to generate the alternate ground truth data from the Google Earth images for properly evaluating the Landsat image classification results. In addition to the commonly used green band-based water indices, Landsat-8 OLI's ultra-blue, blue, and red band-based water indices were also tested in this research. Two unsupervised (the zero-water index threshold H0 method and Otsu's automatic threshold selection method) and one supervised (the k-nearest neighbor (KNN) method) image classification algorithms were employed for conducting the image classification. Through comparing a total of 2880 Landsat image classification results with the alternate ground truth data, this study showed that (1) it is not necessary to use some supervised image classification methods for extracting water bodies from Landsat imagery given the high computational cost associated with the supervised image classification algorithms; (2) the unsupervised classification algorithms such as the H0 and Otsu methods could achieve comparable accuracy as the KNN method, although the H0 method produced more large error outliers than the Otsu method, thus the Otsu method is better than the H0 method; and (3) the ultra-blue band-based $AWEI_{nsuB}$ is the best water index for the H0 method, and the ultra-blue band-based MNDWI_{2uB} is the best water index for both the Otsu and KNN methods.

Keywords: Landsat; Google Earth; water index; unsupervised image classification; supervised image classification; relative error; overall error

1. Introduction

The Earth's inland surface water body consists of rivers, freshwater or saltwater lakes, and marshes with a total surface area of about 5.6 million km² (about 1.1% of Earth's surface area), i.e., 0.8 million km² for rivers [1], 2.1 million km² for lakes, and 2.7 million km² for marshes [2]. Although these inland surface water bodies only hold about 0.013% of Earth's total water, i.e., about 178,000 km³ [2–4], they are important compartments in the global terrestrial water cycle, and mapping inundation areas of inland surface water bodies is of great significance for flood prediction and prevention [5–8]; flood risk and damage assessments [9–17]; estimation of water storage in rivers, lakes, and reservoirs [18–20]; calculation of evaporation from wetlands and lakes/reservoirs [21]; retrieval of lake water level and river stage [22–25]; reservoir operation and management [26]; and assessment of ecological functions and health in wetlands and marshes [27,28]. In addition to the above-mentioned practical applications,

surveying inland surface water bodies can provide critical measurements/observations for improving our understanding of the water cycle and inundation dynamics at multiple spatial and temporal scales [29–31]. Given the tremendous surface area associated with the Earth's inland surface water bodies, mapping their inundation areas from space using remote sensing technique is indeed one of the most efficient approaches.

Both optical (passive) and microwave (active) remote sensing methods have been widely utilized for surveying Earth's inland surface water bodies, but both have some advantages and limitations. Microwave remote sensing is not limited by clouds, weather conditions, and sunlight, but it usually has coarse spatial resolutions, and revisit frequency is also low. Optical remote sensing of inundation areas only works under clear sky and daylight conditions, while the spatial resolutions of spaceborne optical sensors, especially those mounted on commercial satellites (e.g., Ikonos, QuickBird, WorldView, and GeoEye) could achieve the centimeter-level resolutions. Since commercial satellite imagery is not free to the public and can be acquired only through purchase, the most commonly used optical remote sensing images employed for mapping inland surface water bodies at medium resolution (e.g., 30 m) have been and will continually be collected by the Landsat series satellites (e.g., Landsat 4-5 TM, Landsat 7 ETM+, and Landsat 8 OLI) and other polar orbit satellites (e.g., Sentinel).

Most research related to mapping inland surface water bodies using Landsat imagery consists of three steps: (1) using the spectral reflectance captured by Landsat to compute one type of water index at each pixel; (2) using one type of image classification algorithm (unsupervised or supervised) to identify water and non-water pixels; and (3) using ground truth data to assess accuracy of the extracted water bodies. However, all these three steps have some unsolved issues, and inconsistent conclusions can be found across the literature [32–40]. Therefore, this study is dedicated to address these problems so that our knowledge and techniques in optical remote sensing of inland surface water bodies using Landsat imagery can be advanced.

A number of field experiments measuring various land surface features' spectral reflectance [41–43] show that turbid or algae-laden water usually has a distinct higher reflectance in the green light than in any other visible lights, and beyond the near-infrared (>0.9 μ m), their spectral reflectance approaches zero, unlike soil and vegetation which exhibit high reflectance in infrared bands. The characteristics of water spectral reflectance have promoted three commonly used water indices developed in the literature: (1) normalized difference water index (*NDWI*) [44], (2) modified normalized difference water index (*MNDWI*) [45], and (3) automated water extraction index (*AWEI*) [46]. All these water indices were derived based on spectral reflectance in the green light and near-infrared, or shortwave infrared. The *NDWI* [44] is defined as follows:

$$NDWI = (GREEN - NIR) / (GREEN + NIR)$$
(1)

where *GREEN* and *NIR* are spectral reflectance in the green light band and the near-infrared band, respectively. Xu (2006) showed that the *NDWI* had trouble with eliminating the build-up land noise from the extracted water bodies. Considering that the build-up land exhibits relatively higher reflectance in the shortwave infrared (1.5–3.0 µm) band than in the near-infrared (0.7–1.3 µm) band, Xu (2006) replaced the near-infrared band with the shortwave infrared (*SWIR*) band in the NDWI and referred to it as the modified normalized difference water index (*MNDWI*). Although Landsat-5 TM has two SWIR bands, i.e., band 5 (1.55–1.75 µm) and band 7 (2.09–2.35 µm), Xu (2006) found that band 5 of Landsat-5 was better than band 7, thus band 5 was used in the *MNDWI*. Landsat-8 OLI also has two SWIR bands, i.e., band 6 (1.57–1.65 µm) and band 7 (2.11–2.29 µm). In this study, we referred to the first SWIR band (i.e., band 5 in Landsat-5 and band 6 in Landsat-8) as *SWIR*₁, and the second SWIR band (i.e., band 7 in both Landsat-5 and Landsat-8) as *SWIR*₂. To differentiate from *MNDWI*, we call the *SWIR*₂-based water index *MNDWI*₂. Both *MNDWI* and *MNDWI*₂ are defined as follows:

$$MNDWI = \frac{GREEN - SWIR_1}{GREEN + SWIR_1}, \qquad MNDWI_2 = \frac{GREEN - SWIR_2}{GREEN + SWIR_2}$$
(2)

In water body classification, shadows produced by mountains, trees, buildings, and river banks can contaminate satellite imagery classification of water bodies. To remove the impact of shadows, Feyisa et al. (2014) proposed the *AWEI* using five bands, given as follows:

$$AWEI_{ns} = 4(GREEN - SWIR_1) - (0.25NIR + 2.75SWIR_2)$$
(3)

$$AWEI_s = BLUE + 2.5GREEN - 1.5(NIR + SWIR_1) - 0.25SWIR_2$$

$$\tag{4}$$

where *BLUE*, *GREEN*, *NIR*, *SWIR*₁, and *SWIR*₂ are spectral reflectance in the blue light, green light, near infrared, shortwave infrared 1, and shortwave infrared 2 bands, respectively. To use the AWEI, Feyisa et al. (2014) suggested the following criteria: (1) for the areas without high albedo surfaces such as snow cover, and where shadows are the main factor causing errors in the extracted water bodies, $AWEI_s$ alone is sufficient to identify water; (2) if there is no shadow, $AWEI_{ns}$ alone is sufficient; (3) if both high albedo surfaces and shadow/dark surfaces are present, Equations (3) and (4) are used sequentially; and (4) without any shadow/dark surfaces and high albedo surfaces, either one alone can be used.

According to Equations (1)–(4), unsurprisingly, we can find that the spectral reflectance in the green band is the key variable in these three commonly used water indices, given the relatively high reflectance in green light associated with turbid or algae-laden water measured in the fields [41,42]. Thus, in this paper, we referred to these commonly used water indices as the green band-based water indices. However, if we pay attention to the measured spectral reflectance of clear water in Meaden and Kapetsky (1991), we can find that the reflectance in blue light is actually the highest among all visible lights. On the other hand, if water contains a certain amount of sediments, the spectral reflectance in red light should be the highest [41]. It seems that these kinds of questions have not been paid much attention in the literature; therefore, one goal of this study is to evaluate performances of all water indices including the green (commonly used), ultra-blue (only Landsat 8 has this ultra-blue band), blue, and red band-based water indices.

One key purpose of utilizing water indices in the extraction of water bodies from satellite imagery is to simplify the image classification by defining the zero-water index value as the threshold to differentiate water and non-water pixels. However, this single zero-water index threshold method may not work well, and studies showed that a dynamic or automatic selected threshold method such as the Otsu method [47] was better than the zero-water index threshold method [32]. No matter the single threshold method or the automatic selected threshold method, they all belong to the unsupervised image classification. Compared to the unsupervised image classification, the supervised image classification should perform better because human intervention and input of training data could assist computers with identifying water and non-water pixels, although computational efficiency of the supervised methods is usually lower than that of the unsurprised methods. Obviously, there is a tradeoff between computational efficiency and accuracy of image classification, thus the following questions should be answered: (1) Which image classification algorithm is proper for a particular research study? (2) How should one choose an image classification method? Therefore, the second goal of this study is to address these questions and provide recommendations regarding selection of image classification methods through comparing performances of different image classification methods.

Accuracy assessment is the critical final step in image classification that also has some critical issues to be addressed, such as (1) how to collect ground truth data to validate the image classification results and (2) how to properly compare the computed accuracy (e.g., the Kappa coefficient, relative error, overall error, omission error, and commission error) among tested sites. Collecting ground truth data for validating extracted water bodies from Landsat imagery is time-consuming and labor-intensive, especially to draw a general conclusion. Multiple Landsat images across the globe are necessary for the accuracy assessment, which make it even more difficult to facilitate a ground campaign to collect all ground truth data for validating all the selected Landsat images across the globe. Considering the difficulty in collecting ground truth data, this study took advantage of high-resolution satellite

imagery (spatial resolution ≤ 5 m) archived in Google Earth and utilized the Google Earth imagery as the "alternate" ground truth data for evaluating Landsat imagery classification results.

In this study, we targeted the three critical issues discussed above and conducted a systematic study to shed new light on the problems and provided our recommendations for the remote sensing community with regard to the best water index(es) and image classification method(s) in terms of the accuracy of extracted water bodies from Landsat imagery. The remainder of this paper is organized into four sections: Section 2 describes the study area and data sources. Section 3 introduces the methods employed in this study. Results and discussion are presented in Section 4. Conclusions are given in Section 5.

2. Study Areas and Data

The number of high-resolution satellite images archived in Google Earth generally is much less than the number of Landsat images over a specified area because high-resolution satellite (e.g., Ikonos, QuickBird, WorldView, and GeoEye) operators do not provide free daily images to Google Earth. To match the image acquisition date between Landsat imagery and high-resolution satellite imagery over a particular water body (such as lakes and rivers), manually searching high-resolution satellite images archived in Google Earth and Landsat imagery is needed. The strategy taken in this study is to first find high-resolution satellite imagery (click "Show historical imagery" button on Google Earth) over a selected lake, then go to the United States Geological Survey (USGS) EarthExplorer website (earthexplorer.usgs.gov) to search around the same date (±1 day) for a Landsat image without cloud cover over the selected lake.

The method for building the ground truth data based on Google Earth images is described in Section 3. Through searching for high-resolution satellite images archived in Google Earth and Landsat images at the USGS EarthExplorer website, this study selected 24 lakes across the globe as shown in Figure 1. The acquisition date, pixel cell size, water surface elevation (WSE), latitudinal range, and longitudinal range of each Google Earth image are listed in Table 1, along with the acquisition date, and path and row indices of the corresponding Landsat-8 OLI scene. For each Landsat-8 OLI scene, both Precision and Terrain corrected (L1TP) Level-1 (scaled and calibrated digital number) and Level-2 (computed surface reflectance) data were downloaded from the USGS EarthExplorer website. Landsat image processing is also described in Section 3.



Figure 1. Geographic locations of 24 selected lakes across the globe.

6:10			Landsat-8 OLI					
Site	Date	Cell	WSE *	Latitudinal Range	Longitudinal Range	Date	Path	Row
Atitlan	2013/12/04	3.5 m	1558 m	14.7274–14.7590°N	91.1405–91.1855°W	2013/12/04	20	50
Baikal	2013/07/22	4.0 m	450 m	53.0140-53.0593°N	107.0193-107.1232°E	2013/07/21	133	23
Balkhash	2014/10/10	4.0 m	338 m	46.3157-46.3551°N	74.8289–74.9075°E	2014/10/10	151	28
Bansagar	2014/02/20	3.5 m	324 m	24.0759–24.1072°N	80.9680-81.0148°E	2014/02/20	143	43
Beaver	2014/03/19	2.0 m	336 m	36.3524-36.3667°N	93.9478-93.9707°W	2014/03/20	26	35
Brantley	2016/03/12	3.0 m	983 m	32.5583-32.5811°N	104.3742-104.4090°W	2016/03/12	31	37
Brown	2016/04/19	2.0 m	56 m	27.4832–27.4978°S	153.4223-153.4450°E	2016/04/19	89	79
Buchanan	2014/01/13	4.0 m	304 m	30.7729-30.8028°N	98.4219–98.4667°W	2014/01/13	28	39
Burton	2014/10/22	3.0 m	569 m	34.8247-34.8504°N	83.5408-83.5842°W	2014/10/22	18	36
Caspian	2016/08/03	3.5 m	-29 m	42.5993-42.6273°N	47.7777–47.8241°E	2016/08/03	168	30
Chao	2017/07/27	3.0 m	5 m	31.5708-31.5945°N	117.5209-117.5598°E	2017/07/28	121	38
Chelan	2014/07/14	3.0 m	336 m	48.0300-48.0541°N	120.3585-120.4081°W	2014/07/14	46	26
Chuzenji	2017/07/10	4.0 m	1271 m	36.7150-36.7544°N	139.4568-139.5245°E	2017/07/10	107	35
Issykkul	2013/08/31	4.0 m	1603 m	42.5661-42.6077°N	78.1267–78.2045°E	2013/08/31	148	30
Mohave	2015/01/13	3.0 m	198 m	35.4921-35.5166°N	114.6591-114.6962°W	2015/01/13	39	35
Murray	2016/01/28	3.0 m	228 m	34.0811-34.1062°N	97.0776–97.1164°W	2016/01/28	27	36
Ohrid	2015/07/14	3.0 m	690 m	41.0126-41.0444°N	20.6104-20.6684°E	2015/07/14	186	31
Okeechobee	2017/02/11	5.0 m	2 m	26.9826-27.0357°N	80.9090-80.9756°W	2017/02/11	15	41
Sakakawea	2016/08/01	3.0 m	560 m	47.5413-47.5680°N	101.7566-101.8110°W	2016/08/01	33	27
Salton	2016/10/13	3.5 m	-70 m	33.4696-33.5003°N	115.9332-115.8825°W	2016/10/14	39	37
Sélingué	2014/01/26	3.0 m	345 m	11.5978–11.625°N	8.1443-8.1826°W	2014/01/27	199	52
Tanganyika	2017/06/30	3.0 m	768 m	4.8932-4.9134°S	29.5851-29.6130°E	2017/07/01	172	63
Titicaca	2013/08/31	4.0 m	3819 m	15.5053–15.5372°S	69.8433–69.8889°W	2013/09/01	2	71
Trichonida	2013/09/28	4.0 m	11 m	38.5043-38.5481°N	21.6065-21.6836°E	2013/09/28	184	33

Table 1. Characteristics of 24 selected Google Earth and Landsat-8 OLI images.

* WSE: water surface elevation.

3. Methods

3.1. Data Processing

Figure 2 is a flowchart illustrating steps for building the alternate ground truth data from Google Earth (GE) images, extracting and processing Landsat image data, classifying GE and Landsat images, and comparing GE and Landsat image classification results. Each saved GE image was first georeferenced through selecting the World WGS84 as the Geographic Coordinate System and entering coordinates of four corners using the georeferencing function in ArcGIS. Then, each georeferenced GE image was projected from the Geographic projection to the Universal Transverse Mercator (UTM) projection to match the projection of the corresponding Landsat image. The corresponding Landsat-8 OLI image was subsequently clipped to the same spatial extent of the GE image and then resampled into the same grid cell size as the GE image using the bilinear resampling method embedded in ArcGIS.

Among the downloaded Landsat-8 OLI Level-1 images in eight bands, the band 8 image was only used for correcting the possible errors in the georeferenced Google Earth image. Two steps were taken in this study to reduce such errors: (1) first, visually inspecting the images to identify a couple of benchmark pixels in both Google Earth image and the resampled (i.e., with the same pixel cell size as the corresponding Google Earth image) Landsat-8 OLI band 8 image, and then determining the average shift in the pixel distance and applying the average pixel shift distance to correct the Google Earth image; and (2) computing the spatial correlation between the Google Earth image and the resampled Landsat-8 OLI band 8 image over a range of shifts in *x* and *y* directions to determine the optimal shifts in *x* and *y* directions that are associated with the maximum spatial correlation coefficient. Then, use the optimal shifts to correct the Google Earth image.



Figure 2. Flowchart of steps for collecting and processing Google Earth and Landsat image data for evaluating different water indices and image classification methods.

3.2. Water Index

As discussed in Section 1, the commonly used water indices are referred as the green band-based water indices. In this study, in addition to these green band-based water indices, we also compared the performances of the other three sets of water indices, i.e., ultra-blue band, blue band, and red band-based water indices. To keep the commonly used notations of various water indices, this study added a subscript to each water index term for representing the visible band used in the water index as follows:

$$NDWI_X = \frac{X - NIR}{X + NIR}, \quad MNDWI_X = \frac{X - SWIR_1}{X + SWIR_1}, \quad MNDWI_{2X} = \frac{X - SWIR_2}{X + SWIR_2}$$
(5a)

$$AWEI_{nSX} = 4(X - SWIR_1) - (0.25NIR + 2.75SWIR_2)$$
(5b)

$$AWEI_{sX} = BLUE + 2.5X - 1.5(NIR + SWIR_1) - 0.25SWIR_2$$
(5c)

where *X* (*uB*, *B*, *G*, *R*) stands for the ultra-blue, blue, green, or red band used in computing water index.

This study utilized two types of Landsat data for computing water indices: surface reflectance given in the Level-2 products (hereafter SR water index) and top-of-atmosphere (TOA) spectral

reflectance R_{λ} computed from the digital number (DN) of each Landsat-8 imagery pixel using the following equation (hereafter TR water index):

$$R_{\lambda} = (DN_{\lambda} \times M + A) / \cos\theta \tag{6}$$

where $M = 2 \times 10^{-5}$ and A = -0.1 are rescaling factors for converting the digital number to reflectance in band λ , and θ is the solar zenith angle in degrees which is given in the metadata file of each Landsat scene.

3.3. Image Classification Methods

3.3.1. Unsupervised Image Classification

The simplest unsupervised image classification method is to select a single threshold to differentiate water and non-water pixels. Without computing any water index, a simple density slice method can be used to determine a threshold from the histogram of an image, i.e., choosing the digital number associated with the valley of the histogram of the image as the threshold. Although this approach is simple to carry out, it is subject to uncertainty and errors if a histogram does not show a distinct valley. Using the computed water indices for image classification, instead of selecting a threshold based on the histogram, a zero-water index threshold method is usually chosen for extracting water bodies. This method can improve the efficiency of image classification, but it is also subject to uncertainty and errors, because a threshold value of the zero-water index might not achieve the most accurate extraction of the water body. Therefore, this study evaluated the accuracy of the extracted water bodies based on the zero-water index threshold method (hereafter the H0 method).

In addition to the H0 method, a nonparametric and unsupervised automatic threshold selection method proposed by Otsu (1979) was evaluated in this study. The principle of Otsu's method is to maximize the following objective function f:

$$f = P_W P_{NW} (\mu_W - \mu_{NW})^2 \tag{7}$$

where P_W and P_{NW} are probabilities of water pixels and non-water pixels, respectively, and μ_W and μ_{NW} are mean water index values of classified water pixels and non-water pixels, respectively. The optimal water index threshold is determined through searching the water index threshold (WIT) between -1 and 1 with an interval of 0.01 for maximizing the objective function shown in Equation (7). All terms on the right hand of Equation (7) are computed as follows:

$$P_{W} = \frac{n_{W}}{n}, \ P_{NW} = \frac{n_{NW}}{n}, \ \mu_{W} = \frac{\sum_{i=1}^{n_{W}} WI_{i}}{n_{W}}, \ \mu_{NW} = \frac{\sum_{i=1}^{n_{NW}} WI_{i}}{n_{NW}}$$
(8)

where WI_i is the water index of pixel *i*; and *n*, n_W , and n_{NW} are numbers of total pixels, pixels with WI > WIT, and pixels with $WI \le WIT$, respectively.

3.3.2. Supervised Image Classification

Given the relative simplicity of identifying water and water-land boundaries with human visual inspection, supervised classification might be a good choice for fulfilling the task of water pixel classification by inputting a training dataset for classification. There are several supervised classification methods, such as maximum likelihood, Gaussian mixture, minimum distance, nearest neighbor, *k*-nearest neighbor, etc. This study chose the *k*-nearest neighbor classifier (hereafter the KNN method) to be evaluated because of its simplicity and effectiveness [48].

Applying the KNN method to determine if an unknown pixel *x* belongs to water class or non-water class, we first need to compute the spectral distance between the pixel *x* and each training pixel in *m*-dimensional spectral space as follows:

$$d_{i} = \left[\sum_{j=1}^{m} \left(R_{x,j} - R_{i,j}\right)^{2}\right]^{1/2}$$
(9)

where *i* is the index of training pixels, *n* is the number of training pixels, *j* is the band index, *m* is the number of bands to be used in image classification, $R_{x,j}$ is the spectral reflectance of the pixel *x* to be classified in band *j*, and $R_{i,j}$ is the spectral reflectance of the training pixel *i* in band *j*. All the computed spectral distances between the unknown pixel *x* and all the training pixels will be ranked from the lowest to the highest. Based on *k* ranked spectral distances, the final step is to determine which class the pixel *x* belongs to, and *k* is the number of the nearest training pixels to be considered in image classification. There are two questions that must be answered before we can accomplish the final step: (1) What is the suitable *k* value? (2) What is the proper method for classifying unknown pixels? In this study, since we are interested in identifying water-body class, there are actually only two classes to be determined, water and non-water, thus the training pixels either belong to water class or non-water class. Therefore, we set *k* to be the number of training water pixels (*n*_w).

With regard to the second question, two possible approaches can be used to solve this problem. (1) Count the numbers of the nearest neighbors that belong to the water class (k_w) and the non-water class (k_{nw}) among the k ranked nearest training pixels ($k = k_w + k_{nw}$). If k_w is greater than k_{nw} , then the unknown pixel belongs to the water class, otherwise it belongs to the non-water class. (2) Compute the average spectral distance (d_w) to the k_w nearest water pixels and the average distance (d_{nw}) to the k_{nw} nearest non-water pixels. If d_w is less than d_{nw} , the unknown pixel belongs to the water class, otherwise it belongs to the non-water class. However, these two methods are subject to uncertainty associated with the selected k value, because a small variation in the selected k value could result in a different image classification result. To eliminate such uncertainty, in this study, we proposed to compute the sum of the inverse distances of each class among the identified k ranked nearest training pixels. For the water-body identification problem, there are only two classes, water or non-water, thus we only need to compute two sums of the inverse distances. To avoid the division by zero (i.e., as the spectral distance is zero), we first identify the minimum non-zero spectral distance (d_{min}) among the distances from pixel x to all training pixels. If the computed spectral distance is zero, we set the inverse spectral distance to be $2/d_{min}$, otherwise the inverse spectral distance is $1/d_i$, where d_i is the spectral distance from pixel x to training pixel *i*. If the sum of the spectral distances to k_w nearest water training pixels is greater than that to k_{nw} nearest non-water training pixels, pixel x is a water pixel, otherwise it is a non-water pixel.

3.4. Assessment of Image Classification Results

To evaluate the Landsat image classification results, first, a polygon covering a portion of the lake water body and a portion of land on each GE image were defined (Figure 3, as an example), and the KNN method was then used for classifying the water body and land inside the predefined polygon. The reason for choosing the KNN method is that each GE image is an RGB image and the digital numbers (DN) in the three channels (RGB) do not necessarily correspond to the same bands as the Landsat, thus each digital number (DN) could not be converted into the spectral reflectance, and furthermore, each GE image does not contain any near-infrared or shortwave infrared band images which are required for computing water index. Therefore, the H0 method is not applicable for GE images. On the other hand, as a supervised image classification method, the KNN method with input of training data would achieve a higher accuracy than other unsupervised image classification algorithms.



Figure 3. The left panel shows a Google Earth (GE) image of Lake Atitlan overlaid by the defined polygon (white), the identified water-land boundary (green), and the buffer zone boundary (red). The right panel shows the Landsat-8 OLI band 5 image over the same area overlaid by the polygon (white) and the buffer zone boundary (red).

Digital numbers (DN) in three RGB channels of GE images were directly used to compute the digital number distances for identifying water or non-water pixels using the KNN method as described in Section 3.3.2. The classification results were checked and corrected if man-made structures, boats, or clouds appeared in the identified water body areas through human visual inspection. Both the high spatial resolution associated with GE images and visual inspection and correction ensured a relatively high accuracy of the GE image classification results.

After the extraction of the water body inside the predefined polygon, the water-land boundary was identified for defining a buffer zone with a width of 300 m (each side has a perpendicular distance of 150 m to the water-land boundary). The buffer zones were the domains where the image classification results were evaluated. The predefined polygon over each lake was the domain where the optimal water index threshold was determined by the Otsu method, and water and non-water training pixels were selected from for the KNN method. For example, the GE image overlaid by the predefined polygon (in white) and the identified water-land boundary (in green) and the buffer zone (in red) are shown in the left panel of Figure 3, and the corresponding Landsat-8 OLI band 5 image overlaid by the predefined polygon (in white) and the buffer zone (in red) are shown in the right column of Figure 3. The GE and Landsat-8 OLI band 5 images overlaid by the defined polygons and buffer zones for all 24 lakes are shown in the Supplementary Material section.

Using the GE image classification results as the alternate ground truth data, two measures were employed in this study for assessing the Landsat image classification results: (1) relative error (RE) of the extracted water body area, and (2) overall error (OE) of the Landsat image classification:

$$RE = \frac{(n_{Lw} - n_{Gw})}{n_{Gw}} \times 100\%, \quad OE = 100\% - \frac{(n_{Lw|Gw} + n_{Ln|Gn})}{n} \times 100\%$$
(10)

where n_{Lw} is number of water pixels classified by the Landsat, n_{Gw} is number of water pixels classified by the GE image, $n_{Lw|Gw}$ is number of water pixels classified by the Landsat given that they are classified as water pixels by the GE image, $n_{Ln|Gn}$ is number of non-water pixels classified by the Landsat given that they are classified as non-water pixels by the GE image, and n is number of total pixels inside the buffer zone. These two errors are computed inside the defined buffer zone for each study area. The computed relative errors can reveal if the Landsat-extracted water body areas are overestimated (i.e., positive REs) or underestimated (i.e., negative REs). The overall errors are computed as 100% minus the overall accuracy of Landsat image classification results, as shown in Equation (10), and the overall image classification accuracies are evaluated based on error or confusion matrices [48], thus the computed overall errors can reveal the omission or commission errors in the Landsat image classifications of water and non-water pixels.

4. Results and Discussion

This study tested four sets (i.e., ultra-blue, blue, green, and red band-based) of five different water indices (Equations (5a–c)) with three different image classification algorithms (i.e., the H0 method, the Otsu method, and the KNN method). In addition to the Level-1 Landsat data, the Level-2 Landsat surface reflectance data for each scene were also used in computing water indices for image classification and ultimately for evaluating the Landsat image classification results over 24 selected lakes across the globe (Figure 1). Therefore, in total, $4 \times 5 \times 3 \times 2 \times 24 = 2880$ Landsat image classification results were evaluated in this study. The relative errors of the Landsat-extracted water body areas and overall image classification errors of these 2880 cases are listed in the Supplementary Materials section. Given the large number of cases, the boxplot was employed in this study for illustrating the results. All boxplots in this paper show the 5th percentile, 25th percentile, 50th percentile (i.e., medium), 75th percentile, and 95th percentile values, and dots represent data points outside the range of the 5th–95th percentile.

4.1. Impact of Different Landsat Products on Water Classification Results

To assess the impact of different Landsat products on the Landsat-extracted water bodies, three boxplots of water indices versus relative errors of the Landsat-extracted water body areas compared against the water body areas identified from the GE images corresponding to three different image classification methods (H0, Otsu, and KNN) using the water indices computed from the TOA reflectance (i.e., TR water index) and the surface reflectance (i.e., SR water index) are shown in Figures 4 and 5, respectively. Comparisons of these two figures indicate that, as the surface reflectance was used for computing water index, the Landsat-extracted water body areas were generally underestimated, especially as the H0 method was used for image classification; all medians of the relative errors are negative, as shown in Figure 5.



Figure 4. Boxplots of relative errors of the Landsat-extracted water body areas versus water indices computed from the top-of-atmosphere (TOA) reflectance corresponding to three different image classification methods (H0, Otsu, and *k*-nearest neighbor (KNN)).



Figure 5. Boxplots of relative errors of the Landsat-extracted water body areas versus water indices computed from the surface reflectance corresponding to three different image classification methods (H0, Otsu, and KNN).

To demonstrate the cause for such underestimations associated with the SR water indices, Figure 6 shows the computed NDWIs over Chuzenji Lake using the TOA reflectance and the surface reflectance. The TR water indices for the water body inside the buffer zone (white polygon) are greater than zero, while the SR water indices for the water body inside the buffer zone (red polygon) are less than zero, thus the H0 method underestimated the water body area inside the buffer zone if the surface reflectance was used for computing water index.



Figure 6. Normalized difference water indices (NDWI) over Chuzenji Lake computed using the TOA reflectance (**left**) and the surface reflectance (**right**). The white polygon in the left panel and the red polygon in the right panel represent the same buffer zone.

Theoretically speaking, using the surface reflectance to compute water index should yield a higher accuracy in image classification than the TR water index. However, comparisons of relative errors in the Landsat-extracted water body areas and overall image classification errors between the TR water index and SR water index for the three different image classification methods listed in Table 2 all show that the SR water index led to about 75% cases (out of 480 cases) with a worse accuracy than the TR water index, no matter which of three image classification methods (i.e., H0, Otsu, and KNN) was used. The results suggest that the water indices computed based on the Landsat current version Level-2 surface reflectance products might be subject to higher errors and uncertainties than the TR water indices in some regions; therefore, in the remainder of this paper, without specifying, all water indices were computed from the TOA reflectance (i.e., TR water indices).

Error –	TR WI vs. SR WI for H0			TR WI vs. SR WI for Otsu			TR WI vs. SR WI for KNN		
	Better	Worse	Same	Better	Worse	Same	Better	Worse	Same
RE	353 (74%)	127 (26%)	0 (0%)	375 (78%)	104 (22%)	1 (0%)	360 (75%)	115 (24%)	5 (1%)
OE	348 (73%)	132 (27%)	0 (0%)	377 (78%)	94 (20%)	9 (2%)	360 (75%)	117 (24%)	3 (1%)

Table 2. Comparisons of relative errors and overall errors between the TOA reflectance (TR) water index and the surface reflectance (SR) water index for three different image classification methods.

4.2. Comparisons of Three Image Classification Algorithms

In addition to the boxplots of the relative errors of the Landsat-extracted water body areas versus the TR water index for three different image classification algorithms illustrated in Figure 4, the boxplots of the overall Landsat image classification errors versus the TR water index are plotted in Figure 7.



Figure 7. Boxplots of overall image classification errors versus the TR water index corresponding to three different image classification methods (H0, Otsu, and KNN).

Both Figures 4 and 7 suggest that the H0 method did not perform well compared to the Otsu and KNN methods because of the larger error outliers produced by the H0 method, especially as the MNDWI₂ type water index was utilized in the H0 method for image classification. According to Figures 4 and 7, the water index MNDWI₂ employed in the H0 method for image classification produced larger relative errors and overall image classification errors than any other water indices, no matter which visible light band (i.e., ultra-blue, blue, green, or red) was used in the water index MNDWI₂. Unlike the H0 method, as the Otsu and KNN methods were employed for image classification, the water index MNDWI2 yielded comparable relative errors in the Landsat-extracted water body areas and overall image classification errors as the other water indices. These results indicate that if the H0 method is used for classifying water and non-water pixels, the water index MNDWI₂ should be avoided, i.e., SWIR₂ should not be used in the MNDWI. Actually, as Xu (2006) first proposed the MNDWI, SWIR₁ rather than SWIR₂ was chosen for computing the MNDWI. The reason for the overestimated water body areas is that, in some cases, reflectance in SWIR₂ band for non-water pixels was less than the visible band reflectance of these non-water pixels, thereby producing positive MNDWI₂ for these non-water pixels, thus the H0 method based on the water index WI_{3x} overestimated water body areas, such as in Brown Lake and Lake Okeechobee (results listed in Tables ST1 and ST4 in the Supplementary Materials section).

According to Figures 4 and 7, differences in the relative errors of the Landsat-extracted water body areas and overall Landsat image classification errors between the Otsu method and the KNN method are insignificant. Table 3 presents comparisons of relative errors in the Landsat-extracted water body and overall Landsat image classification errors among three image classification algorithms (i.e., H0, Otsu, and KNN). Surprisingly, the numbers of smaller REs or OEs produced by the H0 method are slightly greater than those produced by the Otsu and KNN methods. According to Table 3, as a supervised image classification method, the KNN method is unsurprisingly better (in terms of the numbers of smaller REs and OEs) than the Otsu method, which is an unsupervised method, but surprisingly, it is not better than the H0 method. However, larger error outliers produced by the H0 as shown in Figures 4 and 7 indicate that the H0 method is more sensitive to the water index than the other two methods (Otsu and KNN). These results suggest that: (1) it is not necessary to use some supervised image classification methods for identifying water bodies from Landsat imagery given the high computational cost associated with the supervised image classification methods; (2) the unsupervised image classification algorithms such as the Otsu and H0 methods could yield comparable accuracy in the Landsat-extracted water body areas and image classification of water and non-water classes as the KNN method; and (3) although the zero-water index threshold (i.e., the H0 method) worked better in slightly more than 50% of cases compared to the automatic threshold determined by the Otsu method, the Otsu method produced less large error outliers than the H0 method as shown in Figures 4 and 7. Therefore, if there is no preference when selecting water index for classifying water and non-water classes, the Otsu method is preferable to the H0 method.

Error —	H0 vs. Otsu			H0 vs. KNN			Otsu vs. KNN		
	Better	Worse	Same	Better	Worse	Same	Better	Worse	Same
RE	255	224	1	252	227	1	210	266	4
OE	255	224	1	258	220	2	220	252	8

Table 3. Comparisons of relative errors and overall errors among three image classification methods

4.3. Comparisons of Twenty Water Indices

The results presented in Section 4.2 demonstrate that the accuracies of Landsat-extracted water body areas and Landsat classifications of water and non-water pixels depend on the water index used for classifying water and non-water classes, no matter which image classification method is used, especially for the H0 method, which is very sensitive to water index. That is probably one reason for some debate on the performances of various water indices in the literature. In this paper we reviewed eight relevant studies [32-39] and found that 43.75% (3.5/8) studies claimed that MNDWI was the best, 37.5% (3/8) studies claimed that NDWI was the best, and 18.75% (1.5/8) studies claimed that AWEI was the best, in terms of accuracy of Landsat image classifications. None of the water indices exhibited outstanding superiority (i.e., more than 50%) to others, which is also the case in this study. According to the relative errors of the Landsat-extracted water body areas listed in ST1 (for the H0 method) and ST2 (for the Otsu method) for five commonly used green band-based water indices, i.e., NDWI, MNDWI, MNDWI₂, AWEI_{ns}, and AWEI_s, the corresponding percentages of rank-one in terms of accuracy among 24 lakes are 12.5%, 25%, 25%, 25%, 25%, and 25% for the H0 method, and 8.33%, 12.5%, 25%, 25%, and 29.17% for the Otsu method, respectively.

To compare the performances of twenty water indices tested in this study, means of absolute relative errors (MARE) of Landsat-extracted water body areas and overall errors (MOE) of Landsat image classifications over 24 tested lakes for three different image classification methods (i.e., H0, Otsu, and KNN) were computed and listed in Table 4. If we only focus on the commonly used green band-based water indices, according to Table 4, we can find that AWEI_{sG} produced both the lowest MARE and MOE for the H0 method, AWEI_{nsG} produced both the lowest MARE and MOE for the Otsu method, and MNDWI_{2G} produced both the lowest MARE and MOE for the KNN method. However, if we consider all twenty water indices, Table 4 shows that the ultra-blue band-based AWEI_{nsuB} is the best water index for the H0 method, and the ultra-blue band-based MNDWI_{2uB} is the best water index for both the Otsu and KNN methods, because they produced the smallest MAREs and MOEs compared to all other water indices for the same image classification algorithm. None of the red band-based water indices, which is probably due to the fact that none of the 24 selected lakes in this study had high sediments loads leading to high reflectance in red light.

X47. (X 1		H	D	Ots	su	KNN		
water	Index	MARE (%)	MOE (%)	MARE (%)	MOE (%)	MARE (%) MOE (%		
	NDWI _{uB}	7.41	4.53	7.67	4.42	6.82	4.10	
Ultra-blue	MNDWI _{uB}	16.81	8.26	7.24	4.18	6.39	3.98	
band	MNDWI _{2uB}	51.61	24.94	6.64	3.98	5.49	3.89	
based	AWEI _{nsuB}	4.86	3.59	6.80	4.34	7.65	4.81	
	AWEI _{suB}	9.30	5.71	7.12	4.51	7.12	4.59	
	NDWI _B	5.97	3.89	8.56	4.76	7.70	4.44	
D1 1 1	MNDWI _B	8.33	5.10	8.04	4.49	7.60	4.32	
Blue band based	MNDWI _{2B}	42.24	20.40	7.29	4.19	6.30	4.05	
	AWEI _{nsB}	6.35	3.90	6.99	4.40	7.54	4.80	
	AWEI _{sB}	7.40	4.64	7.27	4.60	7.22	4.65	
Green band based	NDWI _G	7.18	4.36	9.83	5.27	9.13	4.98	
	MNDWI _G	6.90	4.37	9.65	5.12	8.98	4.86	
	MNDWI _{2G}	34.25	16.67	8.53	4.66	7.56	4.33	
	AWEI _{nsG}	9.09	4.92	7.21	4.49	7.79	4.93	
	AWEI _{sG}	6.41	4.09	7.97	4.93	7.92	4.99	
	NDWI _R	12.02	6.42	12.01	6.57	11.01	6.27	
D 11 1	MNDWI _R	8.63	4.89	12.30	6.30	11.79	6.08	
Ked band	MNDWI _{2R}	23.73	12.20	11.46	5.96	10.92	5.75	
based	AWEI _{nsR}	13.27	6.77	7.78	4.79	8.18	5.13	
	AWEI _{sR}	6.99	4.25	9.41	5.65	8.95	5.49	

Table 4. Means of absolute relative errors (MARE) and overall errors (M	OE).
---	----	----

5. Conclusions

This paper addressed three important issues related to extraction of water bodies from Landsat imagery: How to collect ground truth data across the globe for validating Landsat image classification results? Which water indices (among NDWI, MNDWI, AWEI) and which image classification

(unsupervised or supervised) methods are the best for extracting water bodies from Landsat images? First, this study took advantage of high-resolution satellite images archived in Google Earth to obtain 24 high-resolution (\leq 5 m) and cloud-free images and each image covers a portion of 24 selected lakes across the globe, and then a method was developed to generate the alternate ground truth data from the Google Earth images for properly evaluating the Landsat image classification results.

With regard to the computed water indices for identifying water pixels from Landsat imagery, in addition to the commonly used green band-based water indices (i.e., NDWI, MNDWI, and AWEI), Landsat-8 OLI's ultra-blue, blue, and red band-based water indices were also tested in this research, thus a total of 20 types of water indices were evaluated. Both Level-1 Landsat images (used for computing the top-of-atmosphere reflectance) and Level-2 Landsat surface reflectance data were utilized for computing water indices that are referred to as TR and SR water indices, respectively. With regard to the image classification, two unsupervised methods i.e., the single zero-water index threshold method (i.e., the H0 method), and Otsu's automatic threshold selection method, and the supervised KNN method were employed for conducting the image classification. Through comparing a total of $24 \times 20 \times 3 \times 2 = 2880$ image classification results with the alternate ground truth data derived from the Google Earth images, the following conclusions are drawn:

- (1) The top-of-atmosphere reflectance computed from the Level-1 Landsat image data are better than the current Level-2 Landsat surface reflectance products for computing water indices, because the water indices computed based on the Landsat current version Level-2 surface reflectance products might be subject to higher errors and uncertainties than the TR water indices in some regions.
- (2) It is not necessary to use some supervised image classification methods for identifying water bodies from Landsat imagery given the high computational cost associated with the supervised image classification methods. The unsupervised image algorithms such as the Otsu and H0 methods could yield comparable accuracy in the Landsat-extracted water body areas and image classification of water and non-water classes as the KNN method.
- (3) Although the zero-water index threshold (i.e., the H0 method) worked better in slightly more than 50% cases compared to the automatic threshold determined by the Otsu method, the Otsu method produced less large error outliers than the H0 method. Therefore, if there is no preference when selecting water index for classifying water and non-water classes, the Otsu method is preferable to the H0 method.
- (4) Among five commonly used green band-based water indices, AWEI_s produced both the lowest mean absolute relative errors (MARE) in the Landsat-extracted water body areas and mean overall errors in the Landsat image classifications (MOE) for the H0 method, AWEI_{ns} produced both the lowest MARE and MOE for the Otsu method, and MNDWI₂ produced both the lowest MARE and MOE for the KNN method.
- (5) Comparisons among twenty water indices over 24 lakes across the globe showed that the ultra-blue band-based AWEI_{nsuB} is the best water index for the H0 method, and the ultra-blue band-based MNDWI_{2uB} is the best water index for both the Otsu and KNN methods.

In this study, none of the red band-based water indices showed any improvement in extracting water features compared to the green band-based water indices, which is probably due to the fact that none of the 24 selected lakes had high sediments loads leading to high reflectance in red light. Due to limited numbers of high-resolution satellite images archived in Google Earth that can be used for assessing the Landsat water body mapping results, the evaluations of different visible band-based water indices (including both TR and SR water indices) and three image classification algorithms were only carried out on 24 individual lakes with low turbidity, less vegetation cover, and single image acquisition dates. The performances of various water indices and image classification algorithms might differ from the results presented in this study if multiple Landsat images with different acquisition dates over a single lake are used for mapping flooded areas, because as water level declines, a number

of issues such as mixed water-vegetation-sediment pixel, vegetation cover, and turbidity will arise, which deserve further research.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-4292/12/10/1611/s1. Figure SF1: (left panel): 24 Google Earth images overlaid by the predefined polygon (in white) and the identified water-land boundary (in green) and the buffer zone (in red); (right panel): 24 Landsat-8 OLI band 5 images overlaid by the predefined polygon and the identified buffer zone. Table ST1: Relative errors (%) of Landsat-8 OLI water classification results using the zero-water index threshold (H0) method. Table ST2: Relative errors (%) of Landsat-8 OLI water classification results using the Otsu method. Table ST3: Relative errors (%) of Landsat-8 OLI water classification results using the KNN method. Table ST4: Overall errors (%) of Landsat-8 OLI water/land classification results using the zero-water index threshold (H0) method. Table ST5: Overall errors (%) of Landsat-8 OLI water/land classification results using the Otsu method. Table ST6: Overall errors (%) of Landsat-8 OLI water/land classification results using the KNN method. Table ST7: Relative errors (%) of Landsat-8 OLI SR water classification results using the zero-water index threshold (H0) method. Table ST8: Relative errors (%) of Landsat-8 OLI SR water classification results using the Otsu method. Table ST9: Relative errors (%) of Landsat-8 OLI SR water classification results using the KNN method. Table ST10: Overall errors (%) of Landsat-8 OLI SR water/land classification results using the zero-water index threshold (H0) method. Table ST11: Overall errors (%) of Landsat-8 OLI SR water/land classification results using the Otsu method. Table ST12: Overall errors (%) of Landsat-8 OLI SR water/land classification results using the KNN method.

Author Contributions: F.P. conceived and designed the study; F.P., X.X., and C.W. processed and analyzed Google Earth and Landsat-8 OLI images. F.P. conducted image classifications, evaluated results, and wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded in part by the Joint Research Fund for Overseas Chinese Scholars and Scholars in Hong Kong and Macao of National Natural Science Foundation of China (No. 41628101), General Program of National Natural Science Foundation of China (No. 41871264), and Guangxi Innovation Driven Development Project (No. 2018AA13005).

Acknowledgments: The authors would like to thank four anonymous reviewers for their constructive comments and suggestions.

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

References

- 1. Allen, G.H.; Pavelsky, T.M. Global extent of rivers and streams. *Science* **2018**, *361*, 585–588. [CrossRef] [PubMed]
- 2. Shiklomanov, I.A.; Sokolov, A.A. Methodological Basis of Water Balance Investigation and Computation. Available online: http://hydrologie.org/redbooks/a148/iahs_148_0077.pdf (accessed on 16 April 2020).
- 3. Dingman, S.L. *Physical Hydrology*, 3rd ed.; Waveland Press, Inc.: Long Grove, IL, USA, 2015; p. 643.
- 4. Hornberger, G.M.; Wiberg, P.L.; Raffensperger, J.P.; D'Odorico, P. *Elements of Physical Hydrology*, 2nd ed.; Johns Hopkins University Press: Baltimore, MD, USA, 2014; p. 378.
- 5. Overton, I.C. Modeling floodplain inundation on a regulated river: Integrating GIS, remote sensing and hydrological models. *River Res. Appl.* **2005**, *21*, 91–101. [CrossRef]
- 6. Matgen, P.; Schumann, G.; Henry, J.P.; Hoffmann, L.; Pfister, L. Integration of SAR derived river inundation areas, high-precision topographic data and a river flow model toward near real-time flood management. *Int. J. Appl. Earth Obs. Geoinf.* **2007**, *9*, 247–263. [CrossRef]
- Khan, S.I.; Hong, Y.; Wang, J.; Yilmaz, K.K.; Gourley, J.J.; Alder, R.F.; Brakenridge, G.R.; Policelli, F.; Habib, S.; Irwin, D. Satellite remote sensing and hydrologic modeling for flood inundation mapping in Lake Victoria Basin: Implications for hydrologic prediction in ungauged basins. *IEEE Trans. Geosci. Remote Sens.* 2011, 49, 85–95. [CrossRef]
- 8. Ban, H.; Kwon, Y.; Shin, H.; Ryu, H.; Hong, S. Flood monitoring using satellite based RGB composite imagery and refractive index retrieval visible and near-infrared bands. *Remote Sens.* **2017**, *9*, 313. [CrossRef]
- 9. Pelletier, J.D.; Mayer, L.; Pearthree, P.A.; House, P.K.; Demsey, K.A.; Klawon, J.E.; Vincnet, K.R. An integrated approach to flood hazard assessment on alluvial fans using numerical modeling, field mapping, and remote sensing. *Geol. Soc. Am. Bull.* **2005**, *117*, 1167–1180. [CrossRef]
- 10. Sanyal, J.; Lu, X. Remote sensing and GIS-based flood vulnerability assessment of human settlements: A case study of Gangetic West Bengal, India. *Hydrol. Process.* **2005**, *19*, 3699–3716. [CrossRef]

- 11. Sakamoto, T.; Van Nguyen, N.; Kotear, A.; Ohno, H.; Ishitsuka, N.; Yokozawa, M. Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. *Remote Sens. Environ.* **2007**, *109*, 295–313. [CrossRef]
- 12. Taubenbock, H.; Wurm, M.; Netzband, M.; Zwenzner, H.; Roth, A.; Rahman, A.; Dech, S. Flood risks in urbanized areas-multi-sensoral approaches using remotely sensed data for risk assessment. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 431–444. [CrossRef]
- 13. Skakun, S.; Kussul, N.; Shelestov, A.; Kussui, O. Flood hazard and flood risk assessment using a time series of satellite images: A case study in Namibia. *Risk Anal.* **2014**, *34*, 1521–1537. [CrossRef]
- 14. Rahman, M.S.; Di, L. The state of the art of spaceborne remote sensing in flood management. *Nat. Hazards* **2017**, *85*, 1223–1248. [CrossRef]
- 15. Rosser, J.F.; Leibovici, D.G.; Jackson, M.J. Rapid flood inundation mapping using social media, remote sensing an topographic data. *Nat. Hazards* **2017**, *87*, 103–120. [CrossRef]
- 16. Huang, X.; Wang, C.; Li, Z. Reconstructing flood inundation probability by enhancing near real-time imagery with real-time gauges and tweets. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4691–4701. [CrossRef]
- Psomiadis, E.E.; Soulis, K.X.; Zoka, M.; Decas, N. Synergistic approach of remote sensing and GIS techniques for flash-flood monitoring and damage assessment in Thessaly Plain Area, Greece. *Water* 2019, *11*, 448. [CrossRef]
- Frappart, F.; Papa, F.; Famiglietti, J.S.; Prigent, C.; Rossow, W.B.; Seyler, F. Interannual variations of river water storage from a multiple satellite approach: A case study for the Rio Negro River basin. *J. Geophys. Res.-Atmos.* 2008, 113, D21104. [CrossRef]
- Cai, X.; Gan, W.; Ji, W.; Zhao, Z.; Wang, X.; Chen, X. Optimizing remote sensing-based level-area modeling of large lake wetlands: Case study of Poyang Lake. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2015, *8*, 471–479. [CrossRef]
- 20. Normandin, C.; Frappart, F.; Lubac, B.; Belanger, S.; Marieu, V.; Blarel, F.; Robinet, A.; Guiastrennec-Faugas, L. Quantification of surface water volume changes in the Mackenzie Delta using satellite multi-mission data. *Hydrol. Earth Syst. Sci.* **2018**, 22, 1543–1561. [CrossRef]
- 21. Schwerdtfeger, J.; da Silveira, S.W.; Zeilhofer, P.; Weiler, M. Coupled ground- and space- based assessment of regional inundation dynamics to asses impact of local and upstream changes on evaporation in tropical wetlands. *Remote Sens.* **2015**, *7*, 9769–9795. [CrossRef]
- Pan, F.; Nichols, J. Remote sensing of river stage using the cross sectional inundation area-river stage relationship (IARSR) constructed from digital elevation model data. *Hydrol. Process.* 2012, 27, 3596–3606. [CrossRef]
- 23. Pan, F. Remote sensing of river stage and discharge. Spie Newsroom 2013. [CrossRef]
- 24. Pan, F.; Liao, J.; Li, X.; Guo, H. Application of the inundation area-lake level rating curves constructed from the SRTM DEM to retrieving lake levels from satellite measured inundation areas. *Comput. Geosci.* **2013**, *52*, 168–176. [CrossRef]
- 25. Pan, F.; Wang, C.; Xi, X. Constructing river stage-discharge rating curves using remotely sensed river cross-sectional inundation areas and river bathymetry. *J. Hydrol.* **2016**, *540*, 670–687. [CrossRef]
- 26. Ridolfi, E.; Di Francesco, S.; Pandolfo, C.; Berni, N.; Biscarini, C.; Manciola, P. Coping with extreme events: Effect of different reservoir operation strategies on flood inundation maps. *Water* **2019**, *11*, 982. [CrossRef]
- 27. Papa, F.; Prigent, C.; Durand, F.; Rossow, W.B. Wetland dynamics using a suite of satellite observations: A case study of application and evaluation for the Indian Subcontinent. *Geophys. Res. Lett.* **2006**, *33*, L08401.
- Dadson, S.J.; Ashpole, I.; Harris, P.; Davies, H.N.; Clark, D.B.; Blyth, E.; Taylor, C.M. Wetland inundation dyanmics in a model of land surface climate: Evaluation in the Niger inland delta region. *J. Geophys. Res. Atmos.* 2010, 115, D23114. [CrossRef]
- 29. Prigent, C.; Papa, F.; Aires, F.; Rossow, W.B.; Matthews, E.E. Global inundation dynamics inferred from multiple satellite observations. *J. Geophys. Res. Atmos.* **2007**, *112*, D12107. [CrossRef]
- 30. Sheffield, J.; Ferguson, C.R.; Troy, T.J.; Wood, E.F.; McCabe, M.F. Closing the terrestrial water budget from satellite remote sensing. *Geophys. Res. Lett.* **2009**, *36*, L07403. [CrossRef]
- 31. Shi, L.; Ling, F.; Foody, G.M.; Chen, C.; Fang, S.; Li, X.; Zhang, Y.; Du, Y. Permanent disappearance and seasonal fluctuation of urban lake area in Wuhan, China monitored with long time series remotely sensed images from 1987 to 2016. *Int. J. Remote Sens.* **2019**, *40*, 8484–8505. [CrossRef]

- 32. Li, W.; Du, Z.; Ling, F.; Zhou, D.; Wang, H.; Gui, Y.; Sun, B.; Zhang, X. A comparison of land surface water mapping using the normalized difference water index from TM, ETM+ and ALI. *Remote Sens.* **2013**, *5*, 5530–5549. [CrossRef]
- 33. Ronki, K.; Ahmad, A.; Selamat, A.; Hazini, S. Water feature extraction and change detection using multitemporal Landsat imagery. *Remote Sens.* **2014**, *6*, 4173–4189. [CrossRef]
- Yang, Y.; Liu, Y.; Zhou, M.; Zhang, S.; Zhan, W.; Sun, C. Landsat 8 OLI image based terrestrial water extraction from heterogeneous backgrounds using a reflectance homogenization approach. *Remote Sens. Environ.* 2015, 171, 14–32. [CrossRef]
- 35. Malahlela, O.E. Inland waterbody mapping: Towards improving discrimination and extraction of inland surface water features. *Int. J. Remote Sens.* **2016**, *37*, 4574–4589. [CrossRef]
- 36. Yang, X.; Chen, L. Evaluation of automated urban surface water extraction from Sentineel-2A imagery using different water indices. *J. Appl. Remote Sens.* **2017**, *11*, 026016. [CrossRef]
- 37. Zhou, Y.; Dong, J.; Xiao, X.; Xiao, T.; Yang, Z.; Zhao, G.; Zou, Z.; Qin, Y. Open surface water mapping algorithms A comparison of water-related spectral indices and sensors. *Water* **2017**, *9*, 256. [CrossRef]
- 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]
- 39. Bangira, T.; Alfieri, S.M.; Menenti, M.; van Niekerk, A. Comparing thresholding with machine learning classifiers for mapping complex water. *Remote Sens.* **2019**, *11*, 1351. [CrossRef]
- 40. Schwatke, C.; Scherer, D.; Dettmering, D. Automated extraction of consistent time-variable water surfaces of lakes and reservoirs based Landsat and Sentinel-2. *Remote Sens.* **2019**, *11*, 1010. [CrossRef]
- 41. Bartolucci, L.A.; Robinson, B.F.; Silva, L.F. Field measurements of the spectral response of natural waters. *Photogramm. Eng. Remote Sens.* **1977**, *43*, 595–598.
- 42. Meaden, G.J.; Kapetsky, J.M. *Geographical Information System and Remote Sensing in Inland Fisheries and Aquaculture*; FAO Fisheries Technical Paper No.318; FAO: Rome, Italy, 1991; p. 261.
- 43. Jensen, J.R. Introductory Digital Image Processing: A Remote Sensing Perspective, 2nd ed.; Prentice-Hall: Upper Saddle River, NJ, USA, 1995.
- 44. McFeeters, S.K. The use of the normalized difference water index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [CrossRef]
- 45. Xu, H. Modification of normalized difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033. [CrossRef]
- 46. Feyisa, G.L.; Meilby, H.; Fensholt, R.; Proud, S.R. Automated water extraction index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* **2014**, *140*, 23–35. [CrossRef]
- 47. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Mancybernetics* **1979**, *9*, 62–69. [CrossRef]
- 48. Richards, J.A. *Remote Sensing Digital Image Analysis: An Introduction*, 5th ed.; Springer: Berlin/Heidelberg, Germany, 2013; p. 494.



© 2020 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 (http://creativecommons.org/licenses/by/4.0/).