A Support Vector Machine-Based Water Detection Analysis in a Heterogeneous Landscape Using Landsat TM Imagery

Saeideh Sharifzadeh
GIS Research Data Analyst, Caltrans, Los Angeles
Sanchayeeta Adhikari
California State University, Northridge

Abstract

Surface water mapping is essential for studying global environmental changes in the quantity and quality of water bodies. This study explores the applicability of machine learning algorithm Support Vector Machine (SVM) in detecting open water surface from land. The study also compares the SVM based water extraction method with computationally simpler water index, modified Normalized Difference Water Index (mNMDI). The St. Croix Watershed Area is used as a test site for its humid environment with wetlands, built area, forest cover, and shadows as the background noise. The study uses Landsat TM data to generate spectral water index of mNDWI. Zero thresholding is used to generate binary images showing water and non-water areas. Two different SVM models, that is, Linear and Radial Basis Function (RBF) are also used to classify the Landsat TM image into water and non-water class. The accuracy of mNDWI and SVM classifiers are tested and compared using error matrices, Kappa coefficient, overall accuracy and McNemar chi-square test. The results show that the mathematically simpler mNDWI performed better than computationally complex SVM in terms of overall accuracy and Kappa coefficient. Furthermore, mNDWI accurately extracted water from narrow streams and wide rivers while SVM extracted water more accurately from locations in close proximity to urban areas such as reservoirs, boat launching ramps and locations with preponderance of wetlands.
Keywords: Support vector machine, Water detection, Water Indices, Landsat TM

Introduction

Water is one of the most vital natural resources responsible for human survival and development. The detection and observation of water has become a trending research topic due to increasing population pressure on

quality and quantity of water (Niemczynowicz 1990; Palmer et al. 2015), climate change (Middelkoop et al. 2001; Bouraoui et al. 2004; Barnett et al. 2004; Piao et al. 2010), and the need for efficient monitoring of natural hazards such as floods and droughts (Alsdorf et al. 2000; Islam and Sado 2000). In addition, water stored in lakes, rivers, reservoirs, floodplains, and wetlands varies in quantity and quality at different spatial scales (Alsdorf et al. 2007). Development of remote sensing technology for measurement of water variables has received widespread recognition in the past decade, in comparison to traditional *in situ* measurements, as it provides water data at varying spatial, spectral, and temporal scales (Alsdorf et al. 2003; Alsdorf and Lettenmaier 2003). Consequently, a multitude of water detection methods has evolved using both active and passive remote sensing data to manage the wide variety of remotely sensed dataset. Some of the commonly used approaches to detect water are the photo interpretative methods of water detection (Mckay and Blain 2013), traditional supervised and unsupervised classification approaches (Frazier and Page 2000; Sheng et al. 2008), digitization-based water-detection methods (Nath and Deb 2009; Li et al. 2014), and single-band/multi-band continuous water indices methods (Li et al. 2013; Feyisa et al. 2014; Zhou et al. 2017). Among the different methodologies, mathematically simpler continuous water indices have been considered as effective as the traditional thematic classification methods for water detection, especially at global and regional scale analysis (Ryu et al. 2002; Davranche et al. 2010; Yang et al. 2015). Yet these continuous indices-based methods to detect water address only one or two background noises in a satellite image.

Background noises are defined as the pixels that are more often confused with the pixels of interest in a particular study. For example, if low-albedo open-water pixels are the feature of interest, the background noise for the water feature detection could be other low-albedo pixels such as wetlands, built-up areas, shadow, forest cover, and paddy fields. Multi-band indices have been the prevalent choice for water extraction because of their mathematical simplicity, ease of use, less ancillary data, rapid reproducible products, and results comparable to the more complex counterparts of classification algorithms (Fisher and Danaher 2013; Du et al. 2014; Jiang et al. 2014). Yet, these indices were developed to address one or two background noises with respect to water extraction. For example, modified Normalized Difference Water Index (mNDWI) was developed to address the built area noise that NDWI was unable to address (Xu 2006), while NDMI was developed to detect water in a drought conditions (Gao 1996). Other indices, such as the wetness feature of the Tasseled Cap Transformation (TCW), were developed

to measure soil and plant moisture (Crist and Cicone 1984). Similarly, an Automated Water Extraction Index (AWEInsh and AWEIsh) was developed to address topographic shadows that are often confused with pure water pixels (Feyisa et al. 2014). Regardless of the popularity of the multi-band indices, they are disadvantaged in one profound way: most of our real-world landscapes do not have one or two background noises, especially at a global or regional scale. Furthermore, multi-band indices have been preferably used for global and regional scale analysis where field data or ancillary data is unavailable to use in a traditional classifier. Yet, regional- and global-scale water mapping would mean heterogeneous landscapes with different types of background noises that multi-band indices are ill-equipped to address.

To address a broad range of background noises, traditional classification algorithms such as a Gaussian Maximum Likelihood classification or the expert classifier have been the other methodologies for water detection when large number of training data, validation data, and ancillary data are available (McCarthy et al. 2003; Sheng et al. 2008; Li et al. 2014; Pekel et al. 2016). Yet, these traditional classifiers and expert systems are time consuming as they require a large number of training and other ancillary data and are not suitable for global- and regional-scale studies (Ouma and Tateishi 2006). Machine Learning (ML) has recently gained much attention from academics and practitioners in remote sensing as an improvement over traditional methods, in terms of accuracy of performance and automation (Yu et al. 2014; Belgiu and Drăguţ 2016; Lary et al. 2016) (Figure 1). One advantage of ML algorithms is that they limit human assistance in task performance. ML algorithms are also known for faster processing time for large study areas (Robinson et al. 2019). Within ML algorithms, Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Network (ANN) are more commonly used algorithms for satellite image classifications (Lary et al. 2016). Amongst the ML algorithms in remote sensing applications, Support Vector Machine (SVM) was developed as a binary classifier (Cortes and Vapnik 1995). SVM is also an ML approach that produces equally robust results with less training data, as compared to the more traditional classification methods such as Gaussian Maximum Likelihood classification, which require large training data (Wang et al. 2005; Foody and Marthur 2004; Foody and Mathur 2006; Mountrakis et al. 2011). Consequently, SVM could address the multiple background noise of a heterogeneous landscape that multi-band indices are unable to address yet require less training data.

This research explores how well the computationally complex ML algorithm SVM performs, as compared to the traditional, simpler water index meth-

od of mNDWI. Among the many water indices, mNDWI is selected for comparison, as past research has shown an overall better performance of mNDWI in comparison to other water indices in a heterogeneous landscape (Ji et al. 2009; Li et al. 2013; Feisha et al. 2015; Yang et al. 2015; Fisher et al. 2016; Adhikari 2019). The main objectives of this study are thus twofold: (1) extract open-water pixels using continuous water indices of mNDWI- and SVM-based linear and RGB algorithms; (2) statistically compare the robustness of the spectral indices-based method versus the machine learning-based classification algorithms SVM. With this paper we expect to contribute in one key aspect: By comparing the computationally simpler index (mNDWI) and computationally complex (SVM Linear and RGF), we identify the technique best suited to separate water from land in a very large, diverse landscape with wetland and built area as the source of noise.

This research uses St. Croix Watershed Area (SCWA) as a representative test site to test the robustness of accurately extracting water using an mNDWI and SVM classifier. The location has a large number of glaciated lakes and other open-water pixels in the form of reservoirs, rivers, and streams. The test site is ideal, as it includes a large number of wetland area with varying amount of grass, sedge, shrub, and woody cover in the northeastern and central part. The test site also has a substantial urban built area as well as forest cover, cloud shadows, and agricultural fields. Consequently, the test site provides ample background noises for our water detection comparison.

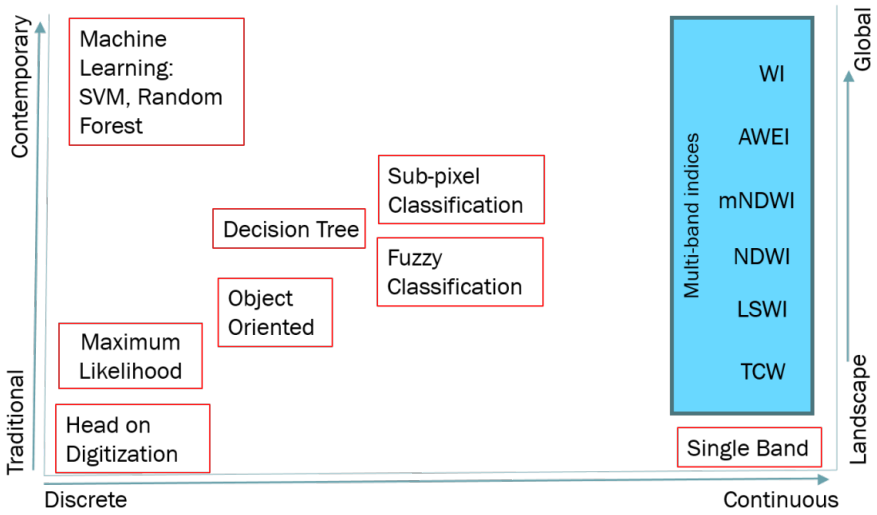

Figure 1.—Development of remote sensing techniques in water extraction (adapted from Southworth and Gibbes 2010).

Site Description

St. Croix watershed area has a total area of approximately 20,000 km^2 and is located in northwest of Minnesota and east-central of Wisconsin (Figure 2). The average rainfall is around 28 to 33 inches annually, with humid continental climatic conditions. The watershed is drained by the St. Croix River for 276 km from its headwaters to its confluence with the Mississippi River. Some of the major tributaries that drain into St. Croix River are the Namekago, Snake, Kettle, Clam, and Yellow Rivers. The topography of the watershed is flat to rolling glaciated terrain, with a large number of lakes. Apart from the lakes and rivers, water in this region is also stored in smaller streams and a large number of seasonally and permanently inundated wetlands (Wegner et al. 2000). The wetlands in this region are diverse, ranging from large alder thicket/swamp hardwood riparian wetlands and shallow marshes in the northeast to open bog wetlands and conifer swamps in the northern part of the watershed and sedge meadows in the south (Bazzell et al. 2002).

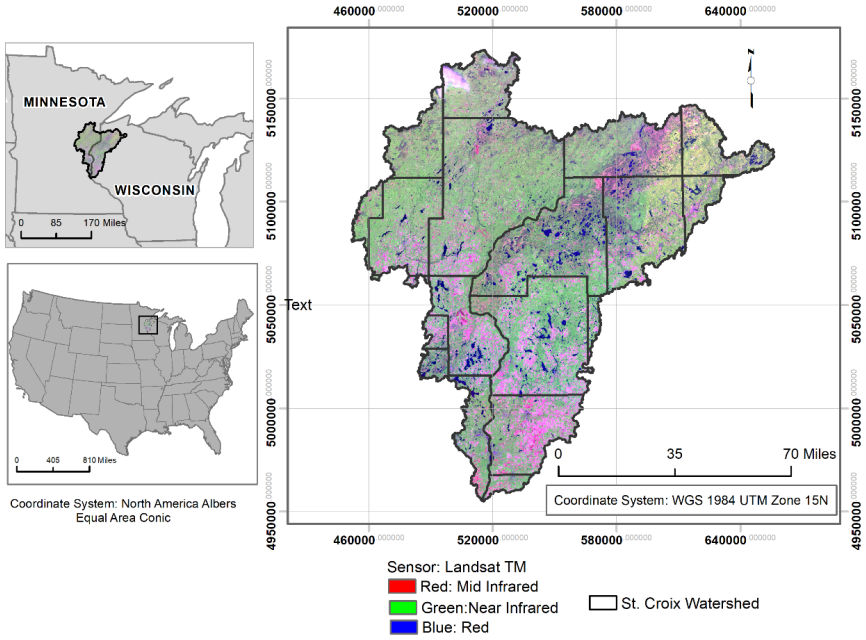

Figure 2.—Location of St. Croix Watershed in Wisconsin and Minnesota.

The watershed area is also divided into the upper and lower St. Croix, due to significant variations in land-cover and land-use types and dominant vegetation species. Upper St. Croix is predominantly forest cover (approximately 60 percent) followed by wetland, grass/pastureland, open water, and

developed land. The Lower St. Croix is dominated by cropland, rangeland, and forest/shrub followed by developed, open water, and wetland. However, cropland has declined since 1940s with increased urban and suburban development of St. Paul-Minneapolis (Edlund et al. 2009). Less than 1 percent of land use has been for residential and commercial development, where the majority of the development occurred in the southern part of the St. Croix region, largely as a result of Minneapolis-St. Paul metropolitan sprawl (Wenger et al. 2000).

Methods

Data Acquisition and Image Preprocessing

The satellite images used for this study area are level 1 US Geological Survey (USGS) Landsat 5 TM, 2011. The images are cloud free and corrected for sensor and atmospheric calibration errors, to minimize the effects of sun angle and atmospheric condition, using the Center for the Study of Institutions, Populations, and Environmental Change (CIPEC) methods (Green et al. 2005). The manual, image-to-image co-registration method has been used to geometrically correct the images with a Root Mean Square Error (RMSE) of less than 0.5 pixels. The images are projected to zone 15 of WGS 1984 UTM. The study area boundary of the St. Croix watershed was created by using a vector layer collected from the Wisconsin and Minnesota Department of Natural Resources (DNR). Imagery for this study is chosen for June 7 (path/row: 27/28 and 27/29) and July 2 (path/row: 26/28), 2011, which are both within the wet season. These dates were selected as they are right after winter snow season in the St. Croix watershed area, the snow is mostly melted, and rainfall allows for easier detection of all open surface water features full of water.

Water Index-Based Approach—Modified Normalized Difference Water Index (mNDWI)

After careful consideration and literature review of commonly used water indices such as TCW, NDVI, NDMI, NDWI, mNDWI, and AWEI, mNDWI was chosen as the best water index approach for detecting water to compare with a machine learning approach. The mNDWI is considered an improvement over the NDWI (Xu 2006). Xu's mNDWI is calculated as (Green-SWIR1) / (Green+SWIR1), where Green represents the Green spectral band and SWIR1 is the shortwave infrared spectral band. The mNDWI values ranges from -1 to +1, whereas water pixel values range between zero and +1. The mNDWI attempts to maximize the reflectance of water body by Green band and minimizes the low reflectance of water in the SWIR1 band, thus providing better contrast between water and non-water pixels (Xu 2006).

The delineation of water pixels from non-water pixels in spectral indices such as mNDWI is done using image binarization method. Image binarization is a method of using an optimal threshold value to convert a continuous satellite image or a spectral index image into two discrete classes of zero and one (binary) (Jensen 2015). This study applies a zero binarization method. Zero thresholding is the simplest binarization method, which assigns zero to all values less than zero and 1 to all values greater than zero in the index. Therefore, for mNDWI, all values coded 1 represent water features and all values coded 0 indicate non-water features.

Machine Learning Approach—Support Vector Machines (SVM)

A Support Vector Machine (SVM) is a supervised, non-parametric statistical learning technique originally formulated by Vapnik (1979). SVM constructs optimal separating hyperplanes between classes in feature space through the use of support vectors. An optimal hyperplane is a decision surface that can separate classes from each other with maximum margin, and support vectors are the data points nearest to the hyperplane that are most difficult to classify (Figure 3).

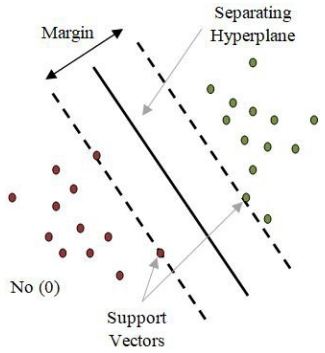

Figure 3.—A demonstration of SVM hyperplane, margin, and support vectors.

The hyperplane can be expressed mathematically as follows:

$$h(x_i) = w.x_i + b \quad \begin{cases} \mathbf{w.xi + b \geq 0} & \mathbf{y = +1} \\ \mathbf{w.xi + b < 0} & \mathbf{y = -1} \end{cases}$$

Where the feature space {xi, yi}, i=1,…,n and yi$\in${-1,+1}. The reflectance value (pixel value) of Landsat TM bands is represented by x, and y is a binary label for water and non-water pixels. The coefficient and constant of

the hyperplane are represented by w and b, respectively. Many hyperplanes could possibly fit the equation (x), but SVM looks for the optimal hyperplane, which maximizes the margin between support vectors. Studies that focused on mathematical formulation of SVMs can be found in previous researches on machine learning (Cortes and Vapnik 1995; Burges 1998).

SVM was originally designed to partition two linearly separable classes (a binary linear classifier) (Cortes and Vapnik 1995). However, SVMs are further extended to deal with non-linear classification by using a non-linear kernel function. In other words, kernel functions offer the option of transforming non-linear spaces into linear ones. Several commonly used kernel functions include linear kernel, polynomial kernel, radial basic function (RBF) kernel, and sigmoid kernel (Ayat et al. 2002; Ayat et al. 2005) (Table 1). It is difficult to determine which kernel would work the best for the study. This study used the radial basic function (RBF) kernel which is commonly used for image classification.

Table 1. SVM Kernel Functions

Kernels	Formula
Linear	$k(x,y) = x.y$
Sigmoid	$k(x,y) = \tanh(ax.y+b)$
Polynomial	$k(x,y) = (1+x.y)^2$
RBF	$k(x,y) = \exp(-a\|\|x-y\|\|^2)$

The SVM algorithm consists of two stages of training and testing. In the training stage, this study used Quantum GIS to collect training samples and labeled them as 1 for water and 0 for non-water pixels. The study used National Land Cover Database 2011 (NLCD, 2011) definition to define water and non-water. Water is defined as any open water pixel with less than 25% cover of vegetation or soil. Non-water is defined as all pixels that are not water pixels, such as urban, agriculture, soil, wetland, bare, forest, and shrubland. These training sites are input data to the SVM algorithm to train the SVM Classifier. The decision boundary generated based on the training sites is used to classify all the image pixels into water and non-water. The study used SVM classifier of the Orfeo Toolbox (OTB) using Monteverdi software, an open source toolbox for remote sensing image processing developed by the French government space agency, National Centre for Space Studies (CNES).

Validation and Comparison Test

For testing the classification accuracy of all the three binary images of SVM Linear, SVM RBF, and mNDWI, reference point samples were collected based on Foody's (2004) binomial probability theory. A total of 200 random samples were generated with equal proportions of water and non-water points, using the NLCD 2011 land cover data as the stratum. The reference points were further verified using Collect Earth, free and open-source software developed by the Food and Agriculture Organization of the United Nations (FAO) to assist in land data collection, management, and analysis (Bey et al. 2016). After careful visual analysis of the random points, 100 non-water and 99 water points were used to run accuracy assessment for the three classified images.

The three classified images were validated using total error, overall accuracy, Kappa coefficients, and error matrices. As classification accuracies are based on sample points used and only provide an estimate (Foody 2004), McNemar's test was calculated. To calculate statistically significant differences between pairwise classifications, McNemar's chi-square statistics with continuity correction were computed as shown in the equation below (Foody 2004) for mNDWI, SVM Linear, and SVM RBF.

$$X^2 = \frac{([f_{12} - f_{21}] - 1)^2}{f_{12} + f_{21}}$$

where, f_{12} and f_{21} denote the number of samples that are correctly classified by one classification method but wrongly classified by the other. McNemar's chi-square test was computed for pairwise comparisons for mNDWI, SVM Linear, and SVM RBF. A ≥ 3.84 would show a significant difference between two classified maps at one degree of freedom at a significance level $\alpha = 0.05$.

Results

This research applied mNDWI and SVM methods of water extraction to Landsat TM image of St. Croix watershed area to separate water and non-water features. The research dealt with a heterogeneous landscape with background noise such as forest cover, built areas, wetlands, agricultural area, cloud cover, and shadow. A zero-thresholding method was used for mNDWI to separate water and non-water. For SVM models, training data for water and non-water was used to create the hyperplane to separate water from non-water. Results were analyzed visually and quantitatively. The visual overviews of classified images are presented in the qualitative results section,

and a comparison of statistical accuracy assessment results is discussed in the quantitative results section.

Qualitative Results

The visual inspection of the three water maps (Figure 4) shows that the water surfaces delineated using mNDWI, SVM Linear, and SVM RBF were similar to the actual water distribution in the NLCD 2011 data. To verify the accuracy of the water extraction, the maps were compared visually for different water types such as narrow streams, man-made reservoirs, and lakes. Reservoirs and other water bodies in close proximity to urban features are better extracted by SVM models.

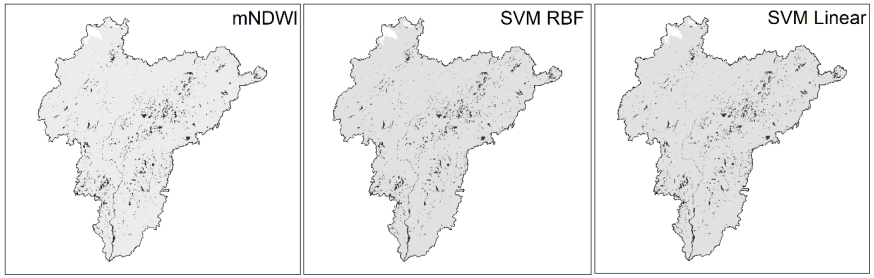

Figure 4.—Classified map of St. Croix watershed showing water and non-water classes using mNDWI, SVM RBF, and SVM linear methods of image segmentation.

Figure 5 shows a bridge and boat ramp over the river, which are misclassified as water by mNDWI water index approach. Comparatively, both SVM linear and SVM RBF extracted the bridge and boat ramp more accurately than mNDWI as a non-water feature. Similarly, mNDWI detected all three reservoirs as one water feature, while SVM classifiers, particularly SVM RBF, accurately separated each reservoir with non-water pixels (Figure 6). However, visual inspection shows that the narrow water features such as streams are more accurately extracted by mNDWI method (Figure 7).

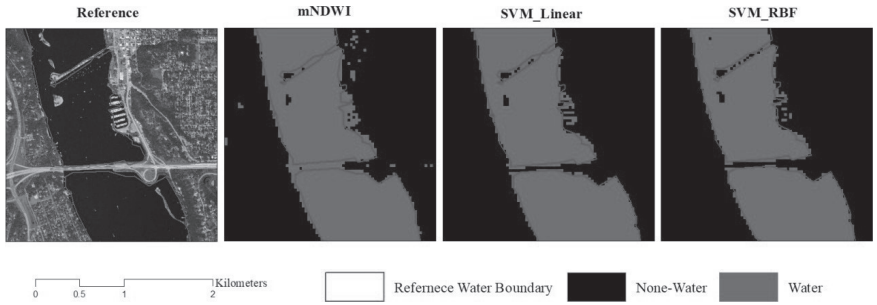

Figure 5.—Comparison of three classifiers to detect bridge over the river.

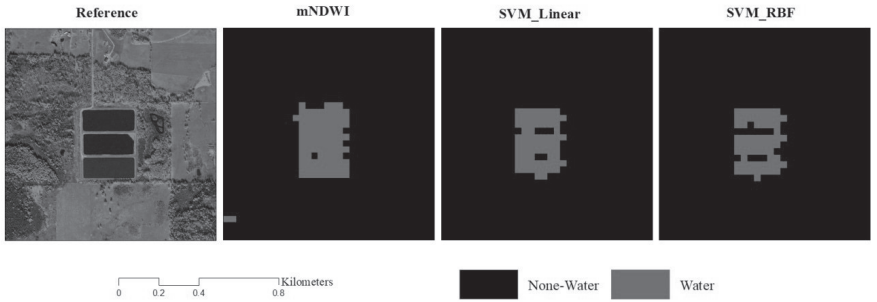

Figure 6.—Comparison of three classifiers to detect three adjacent reservoirs as separate features.

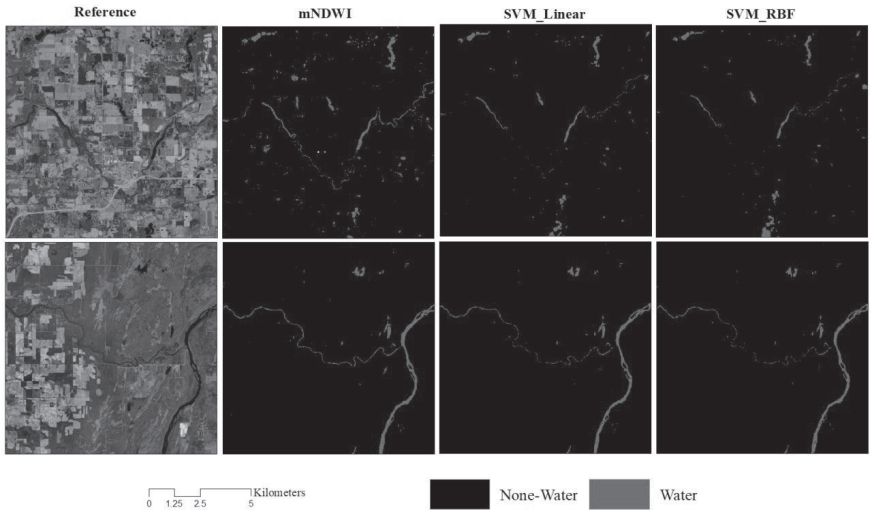

Figure 7.—Comparison of three classifiers to detect narrow streams of water body in the St. Croix watershed area.

Quantitative Results

The classification accuracy is evaluated by error matrix, total error (error of omission and error of commission), overall accuracy, Kappa coefficient, and McNemar's chi-square test. Tables 2 and 3 show overall accuracy and Kappa coefficient results from the mNDWI method and SVM Linear and SVM RBF approaches. The quantitative results show that water index and SVM methods have a similar overall accuracy of nearly 98% (Table 2). mNDWI performed marginally better (0.51%) than the SVM classifiers in overall accuracy.

Table 2. Summary of accuracy assessment results of mNDWI and SVM Models.

	mNDWI	SVM Linear	SVM RBF
Overall Accuracy (%)	98.49	97.98	97.98
Kappa Coefficient	0.96	0.96	0.96

Table 3. Confusion matrices of mNDWI, SVM linear and RBF.

Method		Water	Non-Water	User's Accuracy
mNDWI	Water	96	0	100%
	Non-Water	3	100	97%
	Producer's Accuracy	97%	100%	99%
SVM Linear	Water	95	0	100%
	Non-Water	4	100	96%
	Producer's Accuracy	96%	100%	98%
SVM RBF	Water	95	0	100%
	Non-Water	4	100	96%
	Producer's Accuracy	96%	100%	98%

Comparing omission error (water pixels classified as non-water pixels) between the three classifiers show that all three produced an omission error of 0% (Table 3), which means all the reference non-water pixels are classified correctly. With reference to commission error (non-water pixels classified as water pixels), the results show that the mNDWI method classified 3% of reference water pixels as non-water, while SVM approaches (both linear and RBF kernels) classified 4% of water pixels from reference points incorrectly as non-water pixels.

To assess the statistically significant differences between the three classified maps, the study also calculated McNemar's chi-square test. This test is instrumental in identifying the best model among the mNDWI, SVM Linear, and SVM RBF, as the differences in overall accuracy is merely 0.51%. The McNemar's test result shows that the classification performance was not equal. In fact, mNDWI performed significantly better than both SVM linear (= 60.0161, $p < 0.05$) and SVM RBF (= 72.3049, $p < 0.05$).

Comparison of Background Noise

In order to compare the classification accuracy of water, different background noises are compared (Figures 8 and 9). A closer visual inspection of the background noises shows that SVM models extracted water from urban and wetland background noise much better than the mNDWI model. Two examples of built area in St. Croix watershed area are shown in Figure 8. In both examples, mNDWI misclassified low albedo built-up areas as water ,while SVM classifiers extracted water comparatively better in this background noise. Figure 9 shows water pixels in wetland background. The results from three classifiers indicate that mNDWI misclassified wetlands as water pixels, while SVM classifiers extracted water from a wetland more accurately.

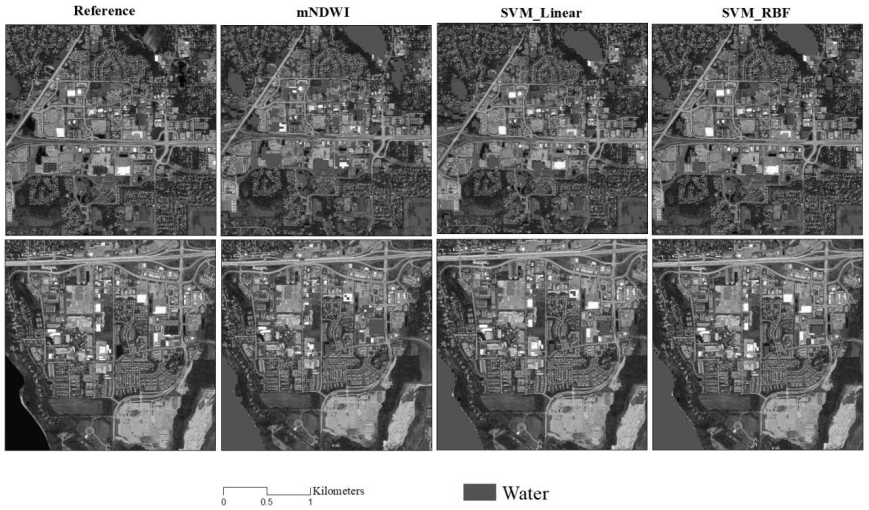

Figure 8.—Comparison of three classifiers to detect water in urban/developed area.

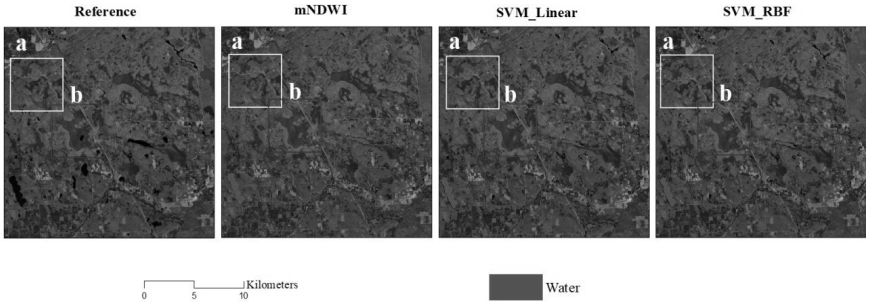

Figure 9.—Comparison of three classifiers misclassifying wetlands as water. Two wetlands are marked as a and b.

Discussion

The St. Croix Watershed Area, in Wisconsin and Minnesota, USA, was selected as a test site for this research. The study area is an ideal test site for comparing different water extraction methods, as there are massive amounts of open surface water area varying in depth of water (deep or shallow water), shape of water boundaries (well-defined lake boundaries, wide rivers, narrow streams), and type of water (pure water or containing vegetation or sediments). Additionally, the study site is a very heterogeneous landscape presenting different land uses and land-cover types. The heterogeneous landscape also means that the region provides a large number of low-albedo land surfaces that are often confused in water extraction such as wetlands with varying amount of grass, shrub and woody cover, built areas, cloud shadows, and forested areas. These background noises and the heterogeneous landscape have made this watershed a very challenging case study for water extraction. Yet, this same challenge of the study site provides an ideal test site to compare between the more traditional and mathematically simpler index-based water extraction method of mNDWI with a more contemporary and computationally complex statistical machine-learning-based algorithm such as the SVM.

Accuracy of water extraction depends on the range of background noises (non-water surface) the research is trying to address; as such, much of the research on water extraction has focused on one or two background noises and developing algorithms that best address those one or two background noises. The water index-based methods are more commonly used because of its simplicity, automated approach, rapidity in application, and less data dependency compared to more complex classification-based water extraction methods. Generally, the threshold values used in these methods are fixed, but it can be challenging in the case of environmental noises such as shadow, forest, clouds, and build-up/urban areas. Among the water indices, mNDWI was developed to address low albedo built area confusion with water bodies (Xu 1996). Studies have also shown that mNDWI performs much better than other index-based methods such as NDWI, AWEI, and NDMI in extracting water from a heterogeneous landscape with multiple background noises (Ji et al. 2009; Li et al. 2013; Feisha et al. 2014; Yang et al. 2015; Fisher et al. 2016; Adhikari 2019). Thus, it is not surprising that mNDWI performed well in extracting water in the heterogeneous landscape of the St. Croix watershed area with an overall accuracy of 98.49% and a kappa coefficient of 0.95. The statistical assessment of classification accuracy performance showed that SVM did not offer an improvement over mNDWI in terms of overall accuracy and Kappa coefficient. The slightly higher overall accuracy of mNDWI in comparison to SVM's overall accuracy (97.98%) is also on par with other similar studies that compared water indices and ML algorithms for water extraction with built areas as background noise (Xu 2007; Sarp and Ozcelik 2017).

Nevertheless, visual inspection of water maps demonstrated that SVM models extracted water from certain sites better than mNDWI in this study site. The sites where SVM Linear and RBF extracted water better than mNDWI are located mostly in the central and northeastern part of the St. Croix watershed area, where there is a dominance of wetland land cover. Wetlands are difficult to separate, as the spectral responses of wetlands can be very similar to open water surfaces, depending on the percentage and type of vegetative cover in the wetland. Thus, the deliberate training data from wetland areas as non-water pixels in an SVM classifier might have allowed for more accurate water extraction in the northeastern part of the study area. The visual inspections also show that SVM models extracted water more accurately than mNDWI in locations that are in close proximity to built-up areas such as bridges, boat launching ramps, man-made reservoirs, and buildings in the southern part of St. Croix watershed area, even though mNDWI was devised to address the built area noise in water extraction mapping. Sanchez et al. (2018) found that SVM performed better in extracting reservoirs, as compared to mNDWI. The present research results resonate with other previous researches that found SVM to maintain spatial characteristics of isolated features such as reservoirs, bridges, and roads better than other algorithms (Huang and Zhang 2009; Mountrakis et al. 2010; Poursanidis et al. 2015). Thus, SVM might be a preferable model over mNDWI in areas where fragmented and isolated built-up areas are the background noise.

Overall, this research concludes that SVM models would be great for areas that are very heterogeneous, with a preponderance of built and wetland land-covers with isolated and fragmented water spaces such as reservoirs and bridges. Yet, the slightly higher accuracy of the quantitative results of mNDWI show that a simple, water-index-based method can produce rapid and reproducible water maps for regional- and global-scale mapping. Additionally, this research shows that the performance of ML algorithms such as SVM are best suited for locations where precise water mapping might be required, and where training data collection and expert knowledge of the study site to validate the training data are available. The contradictory results of the quantitative analysis with the visual interpretations could also be the result of a pixel-based accuracy assessment, a popular approach for classification validation (Ye et al. 2018). Pontius and Millones (2011) have

argued against an accuracy assessment based on the Kappa coefficient that are derived from a confusion matrix. Polygon-based sampling for accuracy assessment, a less common approach, might address the contradictory quantitative and qualitative results better (Ye et al. 2018).

Conclusion

The present study focused on open surface water detection using Landsat TM satellite data, with the St. Croix watershed area as the test site. The St. Croix watershed area provided an ideal site for testing the robustness of a mathematically simpler water index mNDWI with a computationally complex machine-learning algorithm SVM Linear and RBF. The study site of the St. Croix watershed area is an ideal test site, as it provided a large, heterogeneous landscape with many low-albedo background noises. Three water maps were generated from mNDWI, linear SVM, and RBF SVM. To statistically assess the results, error matrices were calculated for each of these water maps, providing overall accuracy and kappa coefficient. McNemar's chi-square test was calculated to check whether the differences in accuracy of the three maps were statistically significant. The quantitative results show that the index method (mNDWI) and classifiers (SVM) produced similar overall accuracies, with slightly better performance for mNDWI. McNemar's chi-square test confirmed that MNDWI was a better method than SVM classifiers. Visual inspection of the images show contradictory results where the SVM classifiers extracted water more accurately than mNDWI in close proximity to built-up areas and wetlands. The quantitative and qualitative results show that a SVM classifier may not provide better performance accuracy of mNDWI, but it can be an improvement over mNDWI to detect open surface water in a landscape with wetland and built area as the dominant background noise. We can conclude that the mNDWI might be best used when rapid and reproducible maps are needed for global- and regional-scale water mapping, as it is algorithmically simple, computationally easy, requires less ancillary data, and does not need expert knowledge about the landscape. In contrast, SVM might be a better approach when precision mapping is needed for a very heterogeneous landscape with urban and wetland as the background noise.

Acknowledgments

We would like to acknowledge and thank the Department of Geography and Environmental Studies, California State University, Northridge, California, USA, for providing work space and research equipment for this study. I also thank Fernando Romero Galvan, an undergraduate student at Department of Geography and Environmental Studies, California State University, Northridge, for assisting with software-specific concerns.

References

Adhikari, Sanchayeeta. 2019. A Comparison of Water Indices and Binary Thresholding Techniques for Water Surface Delineation for St. Croix Watershed Area. *Yearbook of the Association of Pacific Coast Geographers* 81, no. 1: 182–204.

Alsdorf, Douglas E., and Dennis P. Lettenmaier. 2003. Tracking fresh water from space. *Science* 301, no. 5639: 1491–1494.

Alsdorf, Douglas E., Ernesto Rodríguez, and Dennis P. Lettenmaier. 2007. Measuring surface water from space. *Reviews of Geophysics* 45, no. 2.

Alsdorf, Douglas E., John M. Melack, Thomas Dunne, Leal AK Mertes, Laura L. Hess, and Laurence C. Smith. 2000. Interferometric radar measurements of water level changes on the Amazon flood plain. *Nature* 404, no. 6774: 174.

Ayat, Nedjem-Eddine, Mohamed Cheriet, and Ching Y. Suen. 2005. Automatic model selection for the optimization of SVM kernels. *Pattern Recognition* 38, no. 10: 1733–1745.

———. 2002. Optimization of the SVM kernels using an empirical error minimization scheme. In *International Workshop on Support Vector Machines*, pp. 354–369. Springer, Berlin, Heidelberg.

Barnett, Tim, Robert Malone, William Pennell, Detlet Stammer, Bert Semtner, and Warren Washington. 2004. The effects of climate change on water resources in the west: introduction and overview. *Climatic Change* 62, no. 1-3: 1–11.

Bazzell, D., F. Fennessy, and B. Zellmer. 2002. The state of the St. Croix Basin. Wisconsin Department of Natural Resources.

Belgiu, Mariana, and Lucian Drăguț. 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing* 114:24–31.

Bey, Adia, Alfonso Sánchez-Paus Díaz, Danae Maniatis, Giulio Marchi, Danilo Mollicone, Stefano Ricci, Jean-François Bastin, et al. 2016. Collect earth: Land use and land cover assessment through augmented visual interpretation. *Remote Sensing* 8, no. 10: 807.

Bouraoui, F., B. Grizzetti, K. Granlund, S. Rekolainen, and G. Bidoglio. 2004. Impact of climate change on the water cycle and nutrient losses in a Finnish catchment. *Climatic Change* 66, no. 1-2: 109–126.

Burges, Christopher JC. 1998. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery* 2, no. 2: 121–167.

Cortes, Corinna, and Vladimir Vapnik. 1995. Support-vector networks. Machine learning 20, no. 3: 273–297.
———. Support-vector networks." 1995. *Machine learning* 20, no. 3: 273–297.
Crist, Eric P., and Richard C. Cicone. 1984. Application of the tasseled cap concept to simulated thematic mapper data. *Photogrammetric Engineering and Remote Sensing* 50, no. 3: 343–352.
Cristianini, Nello, and John Shawe-Taylor. 2000. An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press.
Davranche, Aurélie, Gaëtan Lefebvre, and Brigitte Poulin. 2010. Wetland monitoring using classification trees and SPOT-5 seasonal time series. *Remote Sensing of Environment* 114, no. 3: 552–562.
Du, Zhiqiang, Wenbo Li, Dongbo Zhou, Liqiao Tian, Feng Ling, Hailei Wang, Yuanmiao Gui, and Bingyu Sun. 2014. Analysis of Landsat-8 OLI imagery for land surface water mapping. *Remote Sensing Letters* 5, no. 7: 672–681.
Edlund, Mark B., Daniel R. Engstrom, Laura D. Triplett, Brenda Moraska Lafrancois, and Peter R. Leavitt. 2009. Twentieth century eutrophication of the St. Croix River (Minnesota–Wisconsin, USA) reconstructed from the sediments of its natural impoundment. *Journal of Paleolimnology* 41, no. 4: 641–657.
Feyisa, Gudina L., Henrik Meilby, Rasmus Fensholt, and Simon R. Proud. 2014. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment* 140:23–35.
Fisher, Adrian, and Tim Danaher. 2013. A water index for SPOT5 HRG satellite imagery, New South Wales, Australia, determined by linear discriminant analysis. *Remote Sensing* 5, no. 11: 5907–5925.
Foody, Giles M., and Ajay Mathur. 2006. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment* 103, no. 2: 179–189.
———. 2004. Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment* 93, no. 1-2: 107–117.
Frazier, Paul S., and Kenneth J. Page. 2000. Water body detection and delineation with Landsat TM data. *Photogrammetric Engineering and Remote Sensing* 66, no. 12: 1461–1468.

Gao, Bo-Cai. 1996. NDWI–A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment* 58, no. 3: 257–266.
Green, Glen M., Charles M. Schweik, and J. C. Randolph. 2005. Retrieving Land-Cover Change Information from Landsat Satellite Images by Minimizing Other Sources of Reflectance Variability. *Seeing the forest and the trees: Human-environment interactions in forest ecosystems*: 131.
Huang, Xin, and Liangpei Zhang. 2009. Road centreline extraction from high-resolution imagery based on multiscale structural features and support vector machines. *International Journal of Remote Sensing* 30, no. 8: 1977–1987.
Islam, Md Monirul, and Kimiteru Sado. 2000. Flood hazard assessment in Bangladesh using NOAA AVHRR data with geographical information system. *Hydrological Processes* 14, no. 3: 605–620.
Jensen, John R. 2015. Introductory digital image processing: a remote sensing perspective. No. Ed. 4. Prentice-Hall Inc.
Ji, Luyan, Xiurui Geng, Kang Sun, Yongchao Zhao, and Peng Gong. 2015. Target detection method for water mapping using Landsat 8 OLI/TIRS imagery. *Water* 7, no. 2: 794–817.
Jiang, Hao, Min Feng, Yunqiang Zhu, Ning Lu, Jianxi Huang, and Tong Xiao. 2014. An automated method for extracting rivers and lakes from Landsat imagery. *Remote Sensing* 6, no. 6: 5067–5089.
Lary, David J., Amir H. Alavi, Amir H. Gandomi, and Annette L. Walker. 2016. Machine learning in geosciences and remote sensing. *Geoscience Frontiers* 7, no. 1: 3–10.
Li, Bangyu, Hui Zhang, and Fanjiang Xu. 2014. Water extraction in high resolution remote sensing image based on hierarchical spectrum and shape features. In *IOP Conference Series: Earth and Environmental Science*, vol. 17, no. 1, p. 012123. IOP Publishing.
Li, Wenbo, Zhiqiang Du, Feng Ling, Dongbo Zhou, Hailei Wang, Yuanmiao Gui, Bingyu Sun, and Xiaoming Zhang. 2013. A comparison of land surface water mapping using the normalized difference water index from TM, ETM+ and ALI. *Remote Sensing* 5, no. 11: 5530–5549.
McCarthy, Jenny M., Thomas Gumbricht, Terence McCarthy, Philip Frost, Konrad Wessels, and Frank Seidel. 2003. Flooding patterns of the Okavango wetland in Botswana between 1972 and 2000. *Ambio: A Journal of the Human Environment* 32, no. 7: 453–458.
McKay, Paul, and Cheryl Ann Blain. 2014. An automated approach to extracting river bank locations from aerial imagery using image texture. *River Research and Applications* 30, no. 8: 1048–1055.

Middelkoop, Hans, Karlheinz Daamen, Daniel Gellens, Wolfgang Grabs, Jaap CJ Kwadijk, Herbert Lang, Bart WAH Parmet, Bruno Schädler, Jörg Schulla, and Klaus Wilke. 2001. Impact of climate change on hydrological regimes and water resources management in the Rhine basin. *Climatic change* 49, no. 1-2: 105–128.

Mountrakis, Giorgos, Jungho Im, and Caesar Ogole. 2011. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 66, no. 3: 247–259.

Nath, R. K., and S. K. Deb. 2009.Water-body area extraction from high resolution satellite images-an introduction, review, and comparison. *International Journal of Image Processing* **3**(6): 353–372.

Niemczynowicz, Janusz. 1999. Urban hydrology and water management–present and future challenges. *Urban Water* 1, no. 1: 1–14.

Otukei, John Richard, and Thomas Blaschke. 2010. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation* 12:S27–S31.

Ouma, Yashon O., and R. Tateishi. 2006. A water index for rapid mapping of shoreline changes of five East African Rift Valley lakes: an empirical analysis using Landsat TM and ETM+ data. *International Journal of Remote Sensing* 27, no. 15: 3153–3181.

Palmer, Stephanie CJ, Tiit Kutser, and Peter D. Hunter. 2015. Remote sensing of inland waters: Challenges, progress and future directions. 1–8.

Pekel, Jean-François, Andrew Cottam, Noel Gorelick, and Alan S. Belward. 2016. High-resolution mapping of global surface water and its long-term changes. *Nature* 540, no. 7633: 418.

Piao, Shilong, Philippe Ciais, Yao Huang, Zehao Shen, Shushi Peng, Junsheng Li, Liping Zhou et al. 2010. The impacts of climate change on water resources and agriculture in China. *Nature* 467, no. 7311: 43.

Pontius Jr, Robert Gilmore, and Marco Millones. 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing* 32, no. 15: 4407–4429.

Poursanidis, Dimitris, Nektarios Chrysoulakis, and Zina Mitraka. 2015. Landsat 8 vs. Landsat 5: A comparison based on urban and peri-urban land cover mapping. *International Journal of Applied Earth Observation and Geoinformation* 35:259–269.

Robinson, Caleb, Anthony Ortiz, Kolya Malkin, Blake Elias, Andi Peng, Dan Morris, Bistra Dilkina, and Nebojsa Jojic. 2019. Human-Machine Collaboration for Fast Land Cover Mapping. arXiv preprint arXiv:1906.04176.

Ryu, Joo-Hyung, Joong-Sun Won, and Kyung Duck Min. 2002. Waterline extraction from Landsat TM data in a tidal flat: A case study in Gomso Bay, Korea. *Remote Sensing of Environment* 83, no. 3: 442–456.

Sánchez, Gabriela Calvario, Oscar Dalmau, Teresa E. Alarcón, Basilio Sierra, and Carmen Hernández. 2018. Selection and Fusion of Spectral Indices to Improve Water Body Discrimination. *IEEE Access* 6: 72952–72961.

Sheng, Yongwei, Chintan A. Shah, and Laurence C. Smith. 2008. Automated image registration for hydrologic change detection in the lake-rich Arctic. *IEEE Geoscience and Remote Sensing Letters* 5, no. 3: 414–418.

Southworth, J., and C. Gibbes. 2010. Digital Remote Sensing within the Field of Land Change Science: Past, Present and Future Directions. *Geography Compass*: 2/6 1–17.

Vapnik, Vladimir. 2006. Estimation of dependences based on empirical data. Springer Science & Business Media.

Wang, Jigang, Predrag Neskovic, and Leon N. Cooper. 2005. Training data selection for support vector machines. In *International Conference on Natural Computation*, pp. 554–564. Springer, Berlin, Heidelberg.

Wenger, Robert B., Hallett J. Harris, and David S. Devault. 2000. An assessment of ecosystem risks in the St. Croix National Scenic Riverway. *Environmental Management* 25, no. 6: 599–611.

Xu, Hanqiu. 2007. Extraction of urban built-up land features from Landsat imagery using a thematicoriented index combination technique. *Photogrammetric Engineering & Remote Sensing* 73, no. 12: 1381–1391.

———. 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing* 27, no. 14: 3025–3033.

Yang, Yuhao, Yongxue Liu, Minxi Zhou, Siyu Zhang, Wenfeng Zhan, Chao Sun, and Yuewei Duan. 2015. Landsat 8 OLI image based terrestrial water extraction from heterogeneous backgrounds using a reflectance homogenization approach. *Remote Sensing of Environment* 171: 14–32.

Ye, Su, Robert Gilmore Pontius Jr, and Rahul Rakshit. 2018. A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches. *ISPRS Journal of Photogrammetry and Remote Sensing* 141: 137–147.

Yu, Le, Lu Liang, Jie Wang, Yuanyuan Zhao, Qu Cheng, Luanyun Hu, Shuang Liu et al. 2014. Meta-discoveries from a synthesis of sat-

ellite-based land-cover mapping research. *International Journal of Remote Sensing* 35, no. 13: 4573–4588.

Zhou, Yan, Jinwei Dong, Xiangming Xiao, Tong Xiao, Zhiqi Yang, Guosong Zhao, Zhenhua Zou, and Yuanwei Qin. 2017. Open Surface Water Mapping Algorithms: A Comparison of Water-Related Spectral Indices and Sensors. *Water* 9, no. 4: 256.