A New Water Index for Surface Water Extraction Using OHS Hyperspectral Imagery

Zhixi Qiao (1), Xin Huang (1)*, Song Ma (2)*
[1] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China
[2] Shenzhen Environmental Monitoring Centre, 202# Meiao 7th street, Futian District, Shenzhen 518049, Guangdong Province, China
Email: xhuang@whu.edu.cn; szhecms@126.com

KEY WORDS: Water index, OHS, Water extraction, Hyperspectral

ABSTRACT: Surface water, a vital earth resource, has an essential influence on ecological environment, society, economy, human health and many other aspects, therefore it is significant to extract and monitor water bodies. Due to characteristics of the large-scale instantaneous observation and revisiting, remote sensing data play an important role in exploring the earth natural resources. The strategy of using water indices, which is simple, easy to use and less time-consuming, is one of the mainstream methods of water body extraction based on remote sensing data. Most traditional water indices, for example, Modified Normalized Difference Water Index (MNDWI) and Automated Water Extraction Index (AWEI), are devised for Landsat such medium spatial resolution imagery and have to use short-wave infrared. As the only commercial hyperspectral satellite with satellite constellation in China, the OHS hyperspectral satellite, launched in 2018, has 32 spectral bands with a spatial resolution of 10m. In spite of the high spectral resolution, the upper limit of its band is only to the near infrared band (1000nm), so the existing water indices basically cannot be applied to this data. To solve this plight, we focus on the characteristics of OHS hyperspectral data, analyze the spectral differences between water bodies and other land cover types and propose a new water index (namely, OHS-WI). A set of pure pixels, consisting of typical water and nonwater (e.g. soil, vegetation and shadow), are selected. By analyzing the spectral patterns, we devised an index (OHS-WI) to discriminate water from other land cover types. By selecting Beijing, Guangzhou and Shenzhen such cities as test sites, we compared the proposed OHS-WI with the traditional Normalized Difference Water Index (NDWI) under two threshold selection strategies, i.e., OTSU (Maximum between-class variance) and optimum (Manually selected). We applied the proposed OHS-WI to generate water extraction results in a series of test sites, which reached a better accuracy than the NDWI method for both threshold selection strategies. The results using optimal threshold achieved high accuracy than that using OTSU threshold in three test sites, with the Kappa coefficient of 0.80, 0.73 and 0.86 (OHS-WI) and 0.67, 0.62 and 0.71 (NDWI). According to the quantitative evaluation metrics, the omission error is lower when using the OTSU threshold while the optimal threshold balances the commission and omission of the extraction results. It is worth noting that the results obtained by OHS-WI under the OTSU threshold are still better than the NDWI extraction results when the optimal threshold is used. The proposed OHS-WI outperformed the NDWI in all the study areas, which can be proved by the metrics of producer's accuracy, user's accuracy and Kappa coefficient. The extracted water maps show that the proposed method not only accurately extracts large water bodies, but also effectively retains small water bodies, which reduces the omission error of extraction results. A series of experimental results show that the OHS-WI water extraction method is full of potential for automatically extracting and monitoring water resources using hyperspectral

remote sensing data.

1. INTRODUCTION

Due to the essential role surface water plays in urban environment, public health and regional ecosystem cycle, extracting and monitoring surface water is a topic of strong demand(Tulbure et al. 2016). In view of the abundant observation archive at multiple scales and time series, remote sensing data are commonly used in water detection(Olthof et al. 2015). There are various water extraction methods for optical imagery, which can be roughly categorized into the following types: (a) spectral characteristics enhancement method, i.e., highlighting the radiation differences between water and nonwater bodies in bands, such as single-band threshold method or water index method(Feyisa et al. 2014; Fisher et al. 2016; Guo et al. 2017; Malahlela 2016; McFeeters 2007; Xu 2006), (b) image classification method(Sun et al. 2015) and (c) linear unmixing method(Olthof et al. 2015). In particular, with the benefit of easy to apply and less time-consuming, the water index methods are widely accepted for water body extraction. However, devised for Landsat such medium spatial resolution imagery, most existing water indices, for example, Modified Normalized Difference Water Index (MNDWI)(Xu 2006), Automated Water Extraction Index (AWEI)(Feyisa et al. 2014), Simple Water Index (SWI)(Malahlela 2016), Weighted Normalized Difference Water Index (WNDWI)(Guo et al. 2017) and 2015 Water Index (WI_{2015})(Fisher et al. 2016), are not suitable for images without short-wave infrared band.
Launched in 2018, the OHS (Zhuhai-1) hyperspectral satellite, which is the only commercial hyperspectral satellite with satellite constellation in China, has 32 spectral bands with a spatial resolution of 10m. Nevertheless, subjected to the band range (400nm-1000nm), the mainstream water indices basically cannot be applied to this data, which limits the application in water extraction. Thus, to address the problem, in this study, we focus on the characteristics of OHS hyperspectral data, analyze the spectral differences between water bodies and other land cover types and propose a new water index (namely, OHS-WI).

2. STUDY AREA AND DATA

2.1 Study area

Taking into account the diversity of water body types and the complexity of the surface environment, the proposed water index aims to reduce the false alarms as much as possible while at the same time, keep a satisfactory producer's accuracy. In order to test the accuracy and robustness of the OHS-WI, three urban areas (Beijing, Guangzhou and Shenzhen) were selected as the study sites (**Figure 1**), in which there are water bodies with variable shapes, sizes and water quality conditions (e.g., rivers, lakes, reservoirs, ponds).

2.2 OHS hyperspectral imagery

As the only commercial hyperspectral satellite with satellite constellation in China, the OHS (Zhihai-1) hyperspectral satellite, with a spectral resolution of 2.5m and a spatial resolution of 10m, has 32 spectral bands, spanning from 400nm to 1000nm. Owing to its abundant information, the OHS hyperspectral imagery could be used in several applications, such as natural resource monitoring, disaster emergency management, urban planning and so on. Moreover, compared with a 10-day revisit frequency of a single satellite, the OHS satellite constellation, equipped with four satellites, could revisit the same location every two days, making it an ideal data for hyperspectral remote sensing in water extraction.

In this study, the image preprocessing includes radiometric calibration, atmospheric correction, and orthorectification, which converting the original DN value to the surface reflectance and correcting the obvious geometric distortion caused by the terrain and sensor.

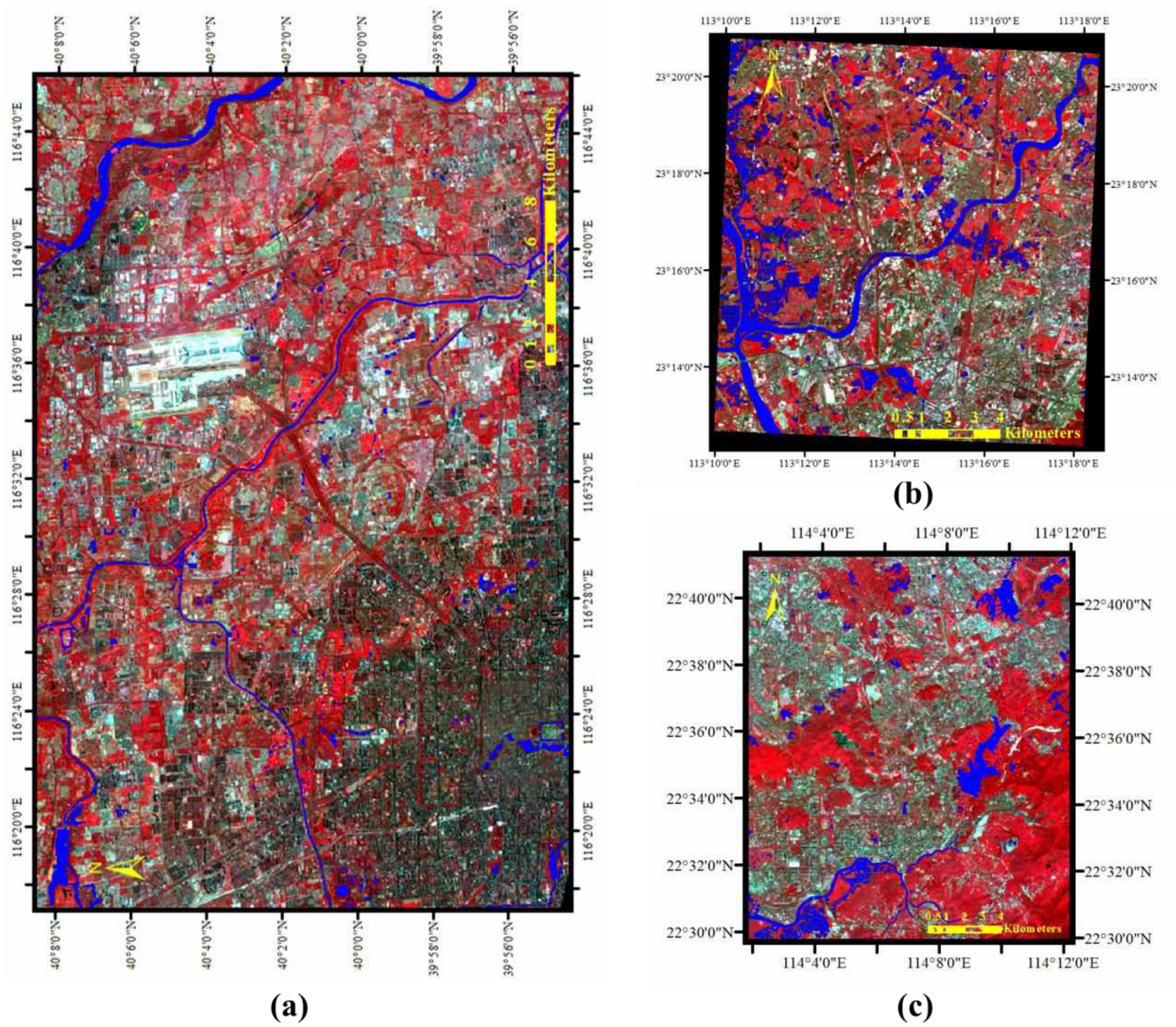


Figure 1. OHS (Zhuhai-1) false color images overlaid with ground truth references of study areas: (a), (b) and (c) are Beijing, Guangzhou and Shenzhen, respectively.

3. METHODOLOGY

The proposed method consists of the following five steps (**Figure 2**): (1) selecting pure pixels; (2) developing a new water index by stepwise linear regression; (3) calculating the water indexes for comparison; (4) selecting the thresholds to extract water; (5) assessing the accuracy. The details of each step are described below.

3.1 Pure-pixel selection

In order to explore the spectral variation between water bodies and other objects, in this study, a set of pure pixels of five classical land cover types, i.e., water, vegetation, soil, building and shadow, were selected in all study sites. Furthermore, because of the diversity among different cities, a special kind of object was added. That is, dark greenhouse in Guangzhou, which might be one of the source of confusion when extracting water bodies. Analyzing the spectral features of these pure pixels, this study aims to design a method to distinguish water from nonwater accurately.
Taking the Google Earth images acquired at the corresponding time as reference, the pure pixels were delineated manually. Meanwhile, the selection of pure pixels obeyed the principles beneath. Water samples were taken from the middle of water bodies, avoiding the mixed pixels on the boundary. For vegetation, try to choose pixels from a large area of homogeneous lawn or forest with high canopy closure. Similarly, the pure pixels of soil were also chosen from agricultural fields in the suburbs. The airport runways, asphalt roads and roofs of large factories were the main sources of building samples. The pure pixels of shadow were distributed as evenly as possible in the urban area.
Finally, in each city, 300 pure pixels were selected for each land cover type. Calculating the separability of each land cover type samples through the Jeffries-Matusita's distance (the value range is between 0 and 2). The J-M values between all pairs of sample types were greater than 1.9, indicating that these samples were separable.

3.2 OHS water index (OHS-WI)

The existing water indexes are usually designed based on subjective judgment of the difference in the spectral curves between water bodies and other land cover types. Normalized Difference Water Index (NDWI) used a ratio method to highlight the difference in reflection between water bodies, bare soil, and vegetation in the green band and near-infrared band so as to distinguish them. By observing the spectral patterns of different land features, the coefficients of the AWEI formulas are obtained by empirical methods. Similarly, analyzing the spectral reflectance of water bodies, shadow as well as built-up areas, the SWI was constructed of blue band and short-wave infrared band.
Nevertheless, the above-mentioned water indexes are mostly used in the Landsat multispectral images. As mentioned before, the band range of Landsat satellite extends to mid-infrared, while the upper limit of the band of OHS hyperspectral date is only to the near infrared band. Therefore, except for NDWI, the existing water indexes cannot be used for water extraction of OHS (Zhuhai-1) hyperspectral data. Moreover, unlike the number of bands of TM/ETM+ sensor and OLI sensor, which is six and eight respectively without the thermal infrared band, the total number of bands of the OHS hyperspectral data is 32, which is 4-5 times the data volume of the former. Hence, the method searching for spectral patterns through subjective observation that used before was time-consuming and not feasible. Introducing the idea of linear regression, this study constructed a new water index – OHS-WI.
In the conventional regression method, to get the regression equation, there are two parts, the real value and the observed value. In this study, the observed value is the spectral reflectance of the pure pixels and the real value is the water index value that needs to be obtained. It should be noted that there is no true index value at that time, because the water index has not been proposed yet.

Considering the final water index, which generally follows the rule that the index value of water bodies is positive while the index value of nonwater is negative, at first, the "index value" of water samples is set to 1 and that of the nonwater samples is set to -1. After stepwise linear regression, further artificial adjustments were made according to the spectral curves of various ground features, and finally the OHS-WI proposed in this study was obtained:

$$\mathrm{OHS-WI} = 0.001(-\rho_4 + \rho_7 + \rho_9 - \rho_{10} + \rho_{12} - \rho_{14} - \rho_{19} + \rho_{23} - \rho_{28}) - 0.43 \qquad (1)$$

where ρ_x represents the reflectance of the *x-th* band.

3.3 Optimal NDWI for comparison

Since OHS hyperspectral data has 32 bands, a series of NDWI can be calculated (Function 2). In order to select the optimal NDWI for comparative experiment, this study carried out 91 combinations of seven possible green bands and thirteen possible near-infrared bands, according to the band range. Based on the pure pixels, the J-M distance was used to quantify the separability of water and nonwater. Eventually, the optimal NDWI combination of each of the three study areas was obtained: Beijing (580nm, 760nm), Guangzhou (610nm, 790nm) and Shenzhen (550nm, 880nm).

$$\mathrm{NDWI} = (\rho_{Green} - \rho_{NIR})/(\rho_{Green} + \rho_{NIR}) \qquad (2)$$

3.4 Threshold selection

The key step of the water index extraction method is the selection of the threshold. In this study, we compared the extraction accuracy under two threshold selection methods (automatic Otsu-based method, manual optimal threshold selection).

The manual threshold is selected based on the F1-score that is an index used in statistics to measure the accuracy of a binary classification model. It takes into account the accuracy and recall of the classification model, which can be regarded as a harmonic mean of model accuracy and recall, ranging from 0 to 1. According to the F1-score, the corresponding optimal NDWI and OHS-WI thresholds of Beijing, Guangzhou, and Shenzhen were obtained: -0.103, 0.084, -0.048 and -1.24, -1.42, -1.25, respectively.

3.5 Estimation of water extraction results accuracy

Compared with the results of NDWI, the extraction results of the OHS-WI were evaluated in three study areas. The producer's accuracy (PA), user's accuracy (UA), and Kappa coefficient calculated from the confusion matrix were used as indicators to evaluate the results of water extraction.

Since the smallest water body of interest in this study is 500 square meters (that is, 5 pixels), all extraction results are constrained in area, which means the extracted water bodies less than 500 square meters were removed.

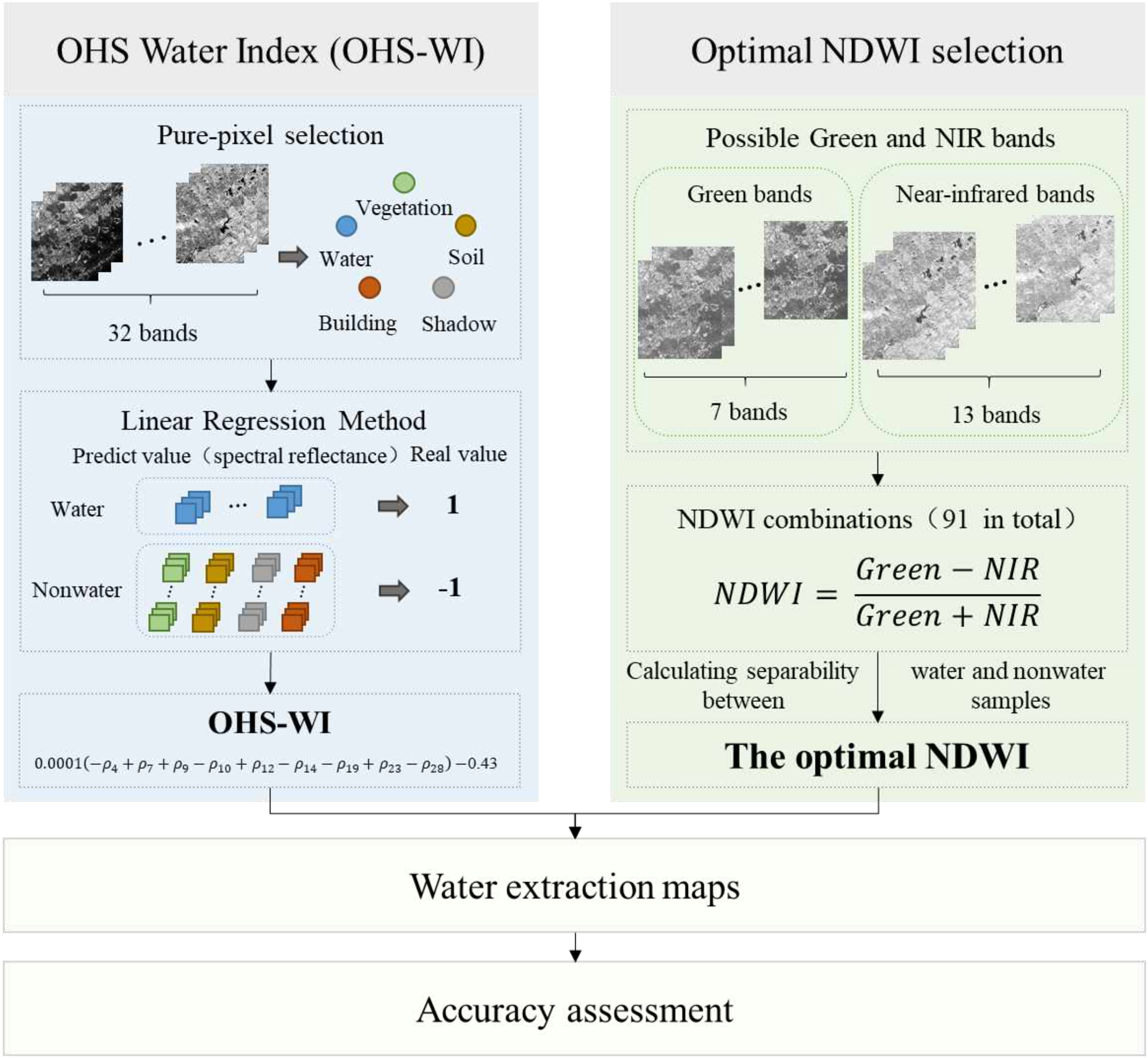


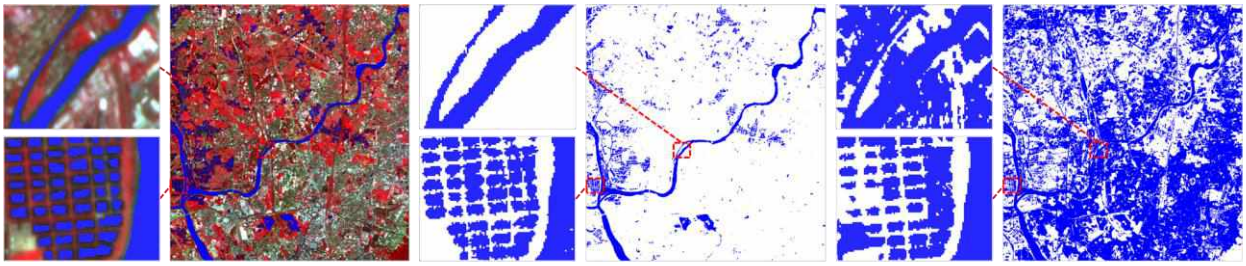
Figure 2. Flowchart of the proposed water extraction method

4. RESULTS

4.1 Water extraction maps using automatic Otsu-based threshold method

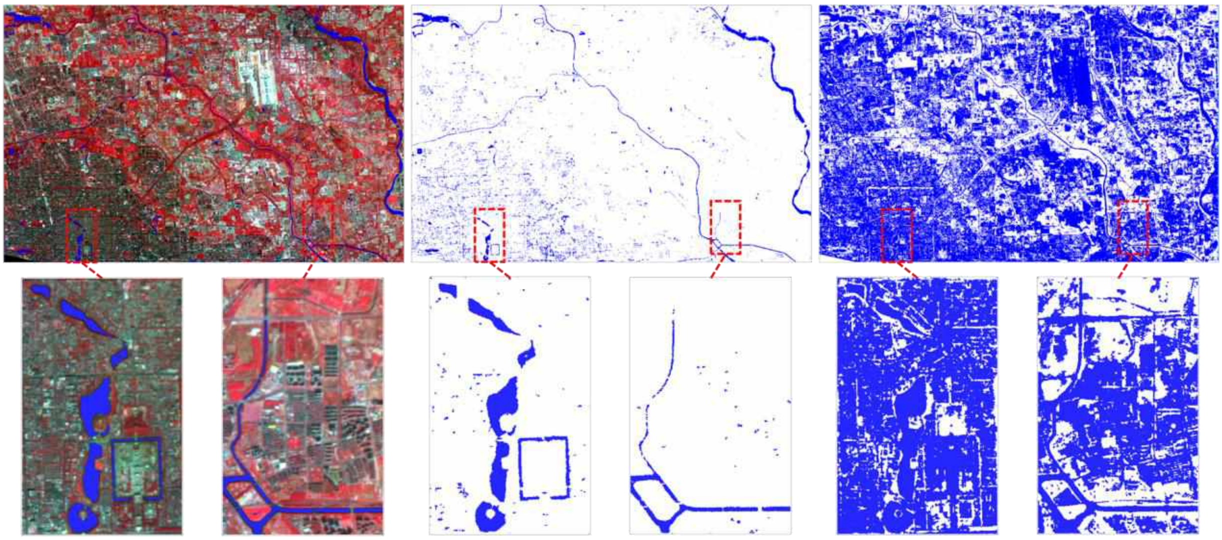
Taking the Otsu-based threshold method, two kinds of water index were used to extract water bodies in all study sites. The final water mapping results are shown in Figure 3-5. Through visual inspection, it is easy to find that the results by NDWI have led to a large number of misclassifications, which means in Guangzhou and Beijing almost 90% of the regions are divided into water bodies. On the other hand, the extraction results using OHS-WI are obviously closer to the real situation and the boundaries of water are complete and clear, although there may be some false positives caused by shadow in dense built-up areas.

According to the consequence of the experiments, under the automatic threshold condition, the OHS-WI is a more reliable choice, because it not only has excellent extraction capability for large water bodies, but also for the smaller ones it could extract the complete shapes, which conforms to the requirements of water body monitoring.



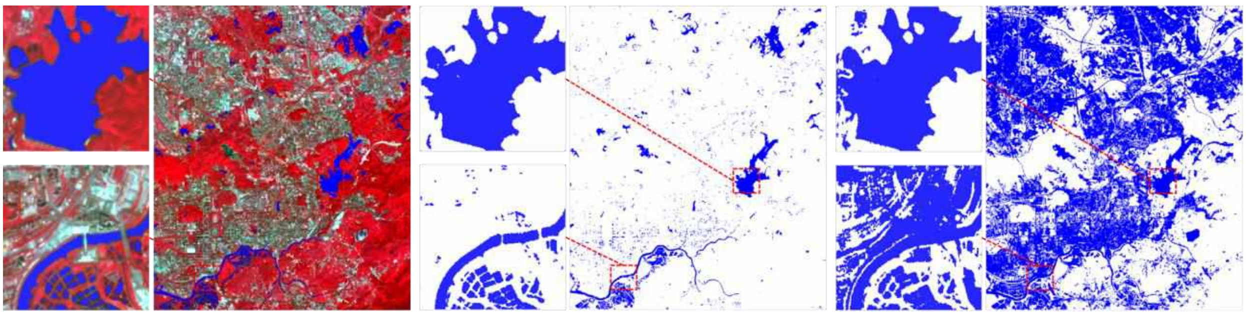
(a) False color image with ground truth (b) OHS-WI (Otsu-based threshold) (c) NDWI (Otsu-based threshold)

Figure 3. Comparison of water extraction results using Otsu-based threshold in Guangzhou



(a) False color image with ground truth (b) OHS-WI (Otsu-based threshold) (c) NDWI (Otsu-based threshold)

Figure 4. Comparison of water extraction results using Otsu-based threshold in Beijing



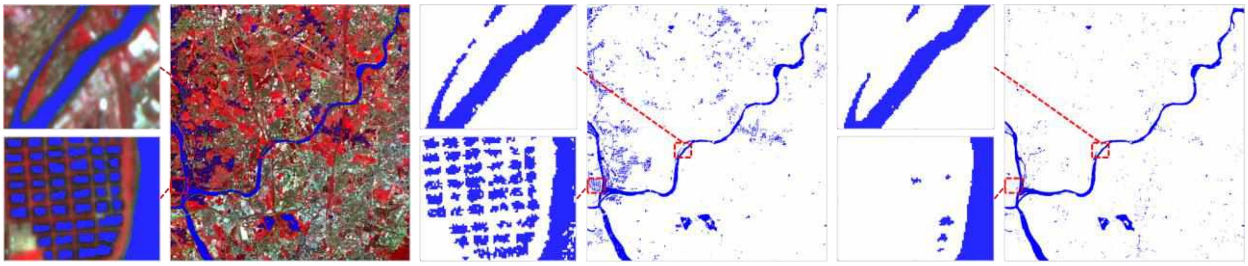
(a) False color image with ground truth (b) OHS-WI (Otsu-based threshold) (c) NDWI (Otsu-based threshold)

Figure 5. Comparison of water extraction results using Otsu-based threshold in Shenzhen

4.2 Water extraction maps using manual optimal threshold method

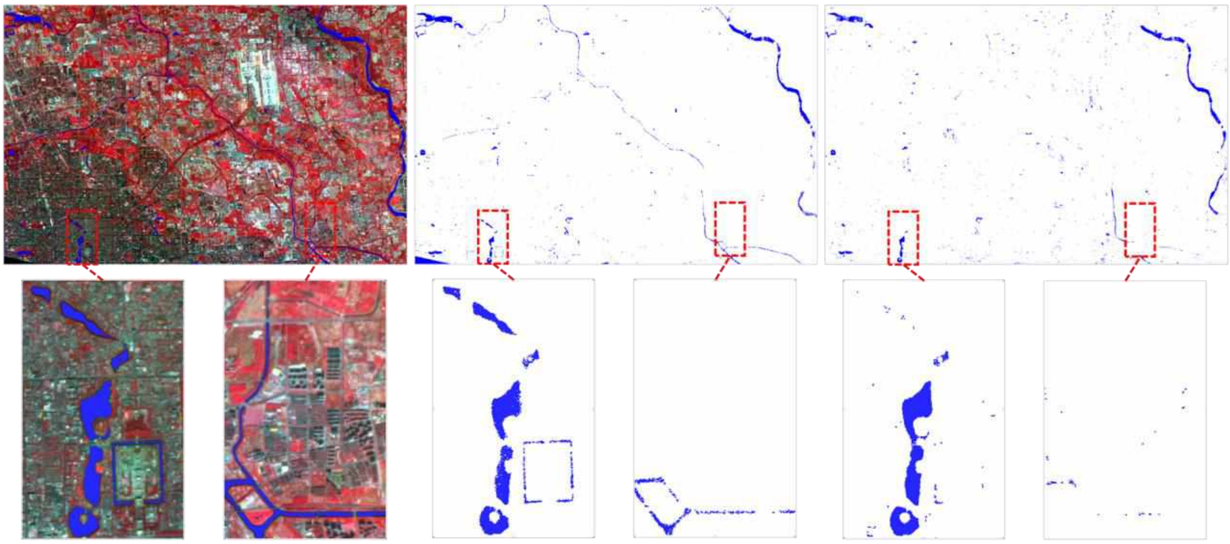
Figure 6-8 show the water extraction results using the two water indexes based on the manual optimal threshold method. Although the lakes such large water bodies could be completely extracted by both two water indexes, the capabilities of the indexes for extracting the small size water bodies are poles apart. In the three study areas, the NDWI results still have many omissions when there are many false alarms, which is mainly due to the lack of river channels in Beijing and the absence of ponds in Guangzhou and Shenzhen. In contrast, compared to the use of automatic threshold, when adopting the manual optical ones, the OHS-WI can retain a relatively complete

water extraction result while eliminating a large number of misclassification errors.
From the results' perspective, the proposed OHS-WI can balance the completeness and accuracy of water extraction, while NDWI cannot achieve the same ideal effect.



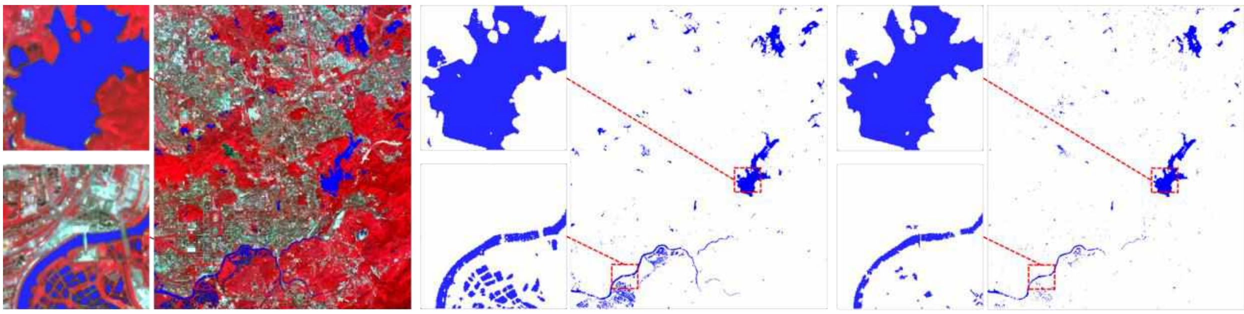
(a) False color image with ground truth (b) OHS-WI (manual optimal threshold) (c) NDWI (manual optimal threshold)

Figure 6. Comparison of water extraction results using manual optimal threshold in Guangzhou



(a) False color image with ground truth (b) OHS-WI (manual optimal threshold) (c) NDWI (manual optimal threshold)

Figure 7. Comparison of water extraction results using manual optimal threshold in Beijing



(a) False color image with ground truth (b) OHS-WI (manual optimal threshold) (c) NDWI (manual optimal threshold)

Figure 8. Comparison of water extraction results using manual optimal threshold in Shenzhen

4.3 Accuracy assessment

In order to better quantitatively compare the water extraction results of the two water indexes, the user's accuracy (UA), producer's accuracy (PA) and Kappa coefficient calculated by confusion matrix were used. It can be seen from Table 1 that when using the manually optimal threshold the Kappa coefficients are all better than that of the automatic threshold, which indicates the former

one takes into account the completeness and accuracy of the extraction results. In all study sites, the Kappa of OHS-WI are better than that of NDWI under the same threshold selection method, which proves that the proposed OHS-WI has a better effect.
The higher producer's accuracy of automatic threshold correlates with the fewer water omissions. Comparing OHS-WI based on automatic threshold and NDWI based on optimal threshold, in all experimental areas, the producer's accuracy of OHS-WI extraction results is much higher than that of NDWI. It is pronounced that OHS-WI makes fewer omissions, which is significant for small water bodies monitoring. The Kappa of OHS-WI using automatic threshold is higher than that of NDWI using optimal threshold in all areas, except Beijing, proving that OHS-WI is a robust and effective water index. The accuracy in Beijing is subjected to the misclassification of building shadow, which can be solved by post-processing.

Table 1. Accuracy assessment of the three study sites

Method	Threshold	Guangzhou			Beijing			Shenzhen		
		Kappa	PA %	UA%	Kappa	PA %	UA%	Kappa	PA %	UA%
OHS-WI	Automatic	0.7764	95.16	67.76	0.5787	97.76	42.00	0.7311	98.53	59.44
	Optimal	0.8028	82.89	80.05	0.7341	70.55	77.36	0.8639	86.92	86.70
NDWI	Automatic	0.082	95.76	9.87	0.0237	99.92	2.69	0.0744	99.23	6.83
	Optimal	0.6728	56.45	87.96	0.6245	55.77	72.22	0.7119	60.64	88.33

5. CONCLUSION

In this study, we proposed a new water index which is designed for the OHS (Zhuhai-1) hyperspectral data and constructed by linear regression, based on reflectance of a set of different land cover types, i.e. water, vegetation, soil, building and shadow. According to the experiments in three study sites, no matter using automatic or manually optimal threshold, the proposed OHS-WI could get relatively accurate extraction results, especially for those small water bodies. Moreover, compared with NDWI, under the condition of automatic threshold, the commission error of OHS-WI is obviously less, and in the case of optimal threshold, OHS-WI reaches fewer omission error. When comparing the water extraction results of OHS-WI using automatic threshold and that of NDWI taking optimal threshold, the former show higher extraction accuracy, which proves the OHS-WI is a simple, robust and potential for application method.
Despite there may be some misclassifications due to the shadow of buildings, when taking the OHS-WI in urban area with dense structures, such commission errors can be reduced by further post-processing based on texture information. This index is useful for the effective extraction of small water bodies and provides a favorable reference for water resources monitoring, investigating and evaluating. In the future, post-processing method and analyzing water quality conditions of specific water body categories will be the focus.

REFERENCE

Feyisa, G.L., Meilby, H., Fensholt, R., & Proud, S.R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment, 140*, pp. 23-35.
Fisher, A., Flood, N., & Danaher, T. (2016). Comparing Landsat water index methods for

automated water classification in eastern Australia. *Remote Sensing of Environment, 175*, pp. 167-182.
Guo, Q., Pu, R., Li, J., & Cheng, J. (2017). A weighted normalized difference water index for water extraction using Landsat imagery. *International Journal of Remote Sensing, 38(19)*, pp. 5430-5445.
Malahlela, O.E. (2016). Inland waterbody mapping: towards improving discrimination and extraction of inland surface water features. *International Journal of Remote Sensing, 37(19)*, pp. 4574-4589.
McFeeters, S.K. (2007). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing, 17(7)*, pp. 1425-1432.
Olthof, I., Fraser, R.H., & Schmitt, C. (2015). Landsat-based mapping of thermokarst lake dynamics on the Tuktoyaktuk Coastal Plain, Northwest Territories, Canada since 1985. *Remote Sensing of Environment, 168*, pp. 194-204.
Sun, X.X., Li, L.W., Zhang, B., Chen, D.M., & Gao, L.R. (2015). Soft urban water cover extraction using mixed training samples and Support Vector Machines. *International Journal of Remote Sensing, 36(13)*, pp. 3331-3344.
Tulbure, M.G., Broich, M., Stehman, S.V., & Kommareddy, A. (2016). Surface water extent dynamics from three decades of seasonally continuous Landsat time series at subcontinental scale in a semi-arid region. *Remote Sensing of Environment, 178*, pp. 142-157.
Xu, H.Q. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing, 27(14)*, pp. 3025-3033.