LIMNOLOGY
and
OCEANOGRAPHY: METHODS

ASLO

Limnol. Oceanogr.: Methods 00, 2017, 00–00

doi: 10.1002/lom3.10215

Development of a spectral-pattern-analysis-based method for automated water body extraction using Landsat image data: A case study in central Vietnam and southern Laos

Nguyen Dinh Duong ,[1]* **Le Minh Hang,**[2] **Tran Anh Tuan,**[3] **Zutao Ouyang**[4]

[1]Department of Environmental Information Study and Analysis, Institute of Geography, Vietnam Academy of Science and Technology, Cau Giay, Hanoi, Vietnam

[2]Department of Geodesy and Mapping, Institute of Techniques for Special Engineering, Le Quy Don Technical University, Cau Giay, Hanoi, Vietnam

[3]Department of Remote Sensing Ecology, Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology, Cau Giay, Hanoi, Vietnam

[4]Michigan State University, Center for Global Changes and Earth Observations, East Lansing, MI 48824

Abstract

Water body mapping is an important application of optical remote sensing. Methods for classifying water bodies using multispectral image data have been successfully developed for water resource monitoring and management. However, in most cases, these traditional methods provide only partial automation capabilities. In this study, we propose a new spectral-pattern-analysis-based (SPAB) method for water body extraction using simplified spectral patterns (SSPs). Simplified spectral patterns are a generalized transformation of spectral patterns from analogue to digital format, realized by nonrepetitive pairwise comparison of reflectance values between two bands. Simplified spectral patterns allow for the direct incorporation of spectral patterns into the recognition process. The advantage of this method compared to traditional methods is that the SPAB method allows automation of water body classification using an SSP database. In this study, we selected a mountainous region in central Vietnam and southern Laos as a study site. Four Landsat 8 OLI scenes from 2015 and four Landsat 7 ETM+ scenes from 2001 and 2002 were used. The results of our proposed method were compared with visual interpretation, normalized differential water index (NDWI) thresholding, and the Global Inland Water (GIW) dataset provided by the Global Land Cover Facility of the University of Maryland. It is concluded that the classification results of the SPAB method agree by more than 98.0% with NDWI results for Landsat 8 OLI images, and by more than 95.0% with the GIW dataset for Landsat 7 ETM+ images.

Water bodies are important components of Earth's environment. They reserve water for use in agriculture, power generation, and daily human consumption. Timely and accurate information of the extent of water bodies is important for many applications, including flood prediction, monitoring, and relief, wetland mapping, and the evaluation of water resources (Smith 1997; Tholey et al. 1997; Livingston et al. 2000). Especially, monitoring water body changes over Tibetan Plateau can provide us many important information about climate change since lakes and other water bodies in this area is minimally disturbed by human activities and their information can be used as key indicator of climate change and climate variability (Song et al. 2014; Zhang et al. 2017*b*).

Remote sensing data are considered an important information source for water surface mapping and are used in the extraction of water body extent and the analysis of temporal changes (Papa et al. 2010). With the integration of geospatial data, remote sensing techniques provide sophisticated data for studying issues related to water resource, for example, flood hazards and changes in surface water (Haas et al. 2009; Ji et al. 2009; Gardelle et al. 2010; Proud et al. 2011).

There are four basic types of open water body classification methods (Ji et al. 2009; Feyisa et al. 2014): (1) Thematic classification, (2) Linear unmixing, (3) Single-band thresholding, and (4) Two-band spectral water indices (McFeeters 1996; Xu 2006; Zhanget al. 2017*a*).

Some scientists have proposed combining various methods, including geospatial analysis, to improve water extraction accuracy (Sheng et al. 2008; Verpoorter et al. 2012; Jiang et al. 2014). The use of single-band or two-band methods for water extraction is common because of their ease of application (Ryu et al. 2002). Even though there are a

*Correspondence: duong.nguyen2007@gmail.com

Table 1. Image data used in this study.

Dataset	Scene ID	Date of observation	Land cloud cover (%)
Landsat image data	LC81240492015161LGN00	10/06/2015	1.73
	LC81250492015024LGN00	24/01/2015	0.35
	LC81240502015065LGN00	6/03/2015	2.95
	LC81250502015104LGN00	14/04/2015	0.13
	LE71240492001082SGS00	23/03/2001	1.0
	LE71240502001290EDC00	17/10/2001	6.0
	LE71250492001329SGS00	25/11/2001	6.0
	LE71250502002044SGS00	13/02/2002	0.0
Global Inland Water dataset	p124r049_WC_20010323.TIF	23/03/2001	1.0
	p124r050_WC_20011017.TIF	17/10/2001	6.0
	p125r049_WC_20011125.TIF	25/11/2001	6.0
	p125r050_WC_20020213.TIF	13/02/2002	0.0

number of methods for water body extraction detailed in the literature, the selection of a particular technique depends on many factors such as accuracy requirements, purpose of the application, and availability of resources (remote sensing data, ancillary information, and human resources).

The need for automated water classification over large areas has recently become urgent, but published algorithms are either too complicated or need substantial ancillary data. For example, automated mapping of inland surface water based on atmospherically corrected reflectance, topographic indices, and prior coarse-resolution water layers was applied to the Global Land Survey collection of Landsat images (Feng et al. 2016). This required computation of the normalized differential vegetation index (NDVI), the normalized differential water index (NDWI), and the modified NDWI, as well as application of a predetermined threshold of these indices. Another algorithm, which was recently used to study surface water dynamics in the North American high northern latitudes, was based on a decision tree with multiple Landsat data converted from radiance to surface reflectance (Carroll et al. 2016). Each Landsat scene was processed through a decision tree classification to generate six land cover classes: water, vegetated land, bare ground, snow/ice, cloud, and cloud shadow.

Spectral pattern analysis has been in use for many years in the analysis of multispectral data (Zhang and Sriharan 2005). The basis of spectral pattern analysis is the use of a reference spectral library or reference endmembers selected from the image. Popular spectral analysis methods include Spectral Feature Fitting, Spectral Angle Mapper, or binary encoding classification (Zhang and Sriharan 2005; Canty 2014). A common feature of these methods is the direct use of spectral values for computation, which is typically complex and resource demanding, and cannot adequately exploit the graphical characteristics of the spectral reflectance curve, such as shape or modulation, required for classification.

Modulation of the spectral pattern of water has been proposed for water body classification (Duong 2012). In this method, four spectral bands (green, red, near infrared, and short-wave infrared bands) are used to assess the spectral pattern of open water, and the pattern is combined with the ratios of green to short-wave infrared or red to short-wave infrared to extract the water body. The procedure is not fully automated because threshold values can change both spatially and temporally.

In this paper, we introduce a novel spectral-pattern-analysis-based (SPAB) method for automated water body extraction. In contrast to traditional spectral analysis methods, the proposed method uses geometrical characteristics of the spectral pattern, such as modulation, for water classification so that computation is significantly simplified and overall performance is fundamentally improved. This method relies on a simplified spectral pattern concept, described by Duong in Decomposition of Landsat 8 OLI Images by Simplified Spectral Patterns (unpubl.), applied to the Landsat TM/ETM+ and OLI reflective bands. We conducted an experiment using eight Landsat 7 ETM+ and Landsat 8 OLI scenes (Table 1) that cover the study area. To evaluate the reliability of the proposed method, we compared our results with visual interpretations, the NDWI, and the Global Inland Water (GIW) dataset for the study area. The comparison with the NDWI was performed using Landsat 8 OLI scenes while the comparison with the GIW dataset was performed using Landsat 7 ETM+ images. The GIW dataset (Feng et al. 2016) was downloaded from the Global Land Cover Facility website (http://glcf.umd.edu/data/watercover/). The proposed method allows the processing of multiple time series image data and the classification of water types by different simplified spectral patterns (SSPs). It improves on previous methods in faster processing and simpler classification procedure. This method lays the foundation for rapid water body monitoring over large areas covered by multiple Landsat scenes.

Materials and procedures

Study area

The study area is located in the central part of Vietnam and southern Laos (Fig. 1). The terrain is rugged, steep, and

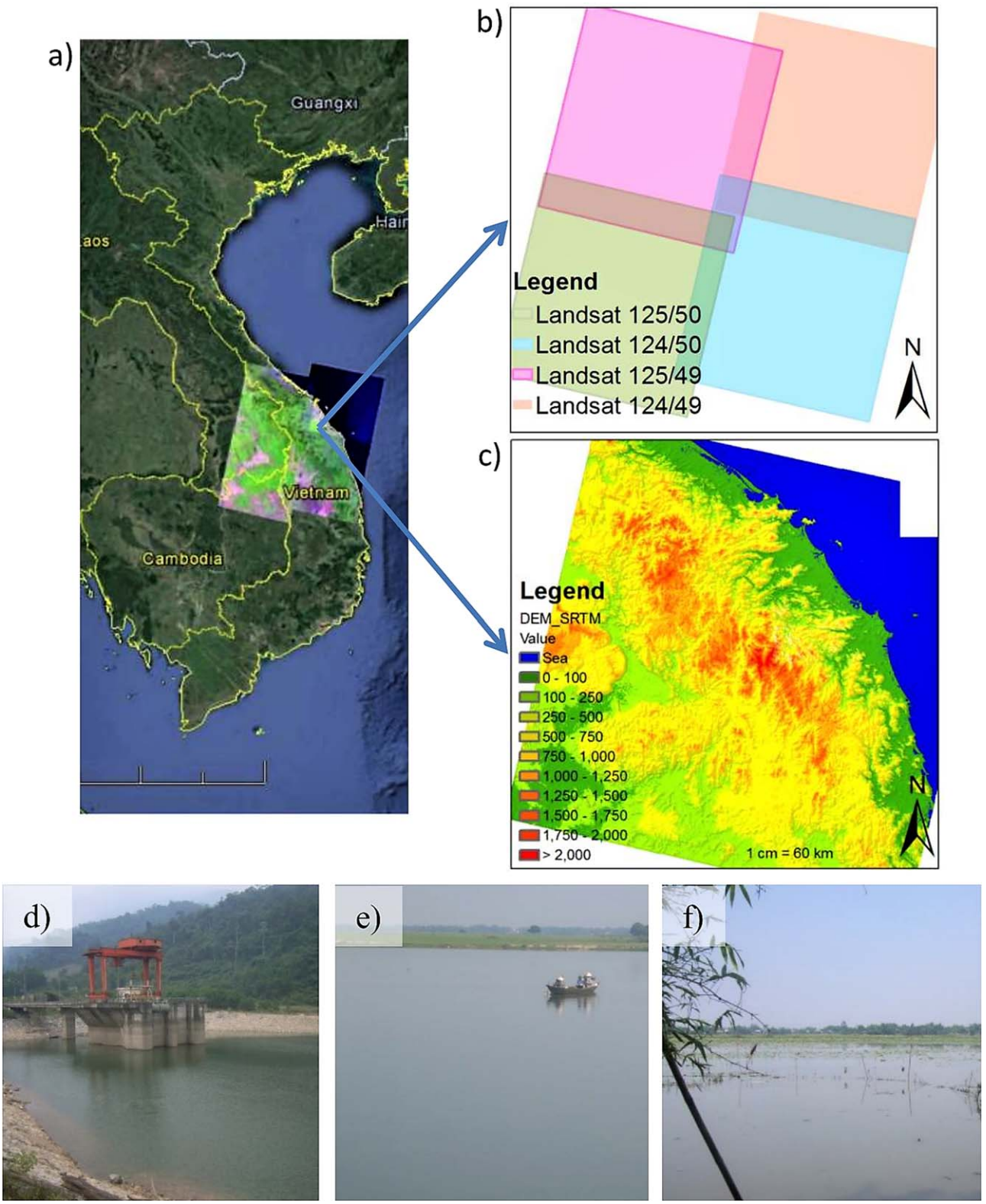

Fig. 1. Location of the study area in Vietnam and major inland water body types. (**a**) Location of the study area in central Vietnam and southern Laos, (**b**) Footprints of Landsat scenes with path/row numbers of 124/49, 124/50,125/49, and 125/50, (**c**) DEM of the study area, (**d**) artificial water reservoir, (**e**) river, and (**f**) natural lake.

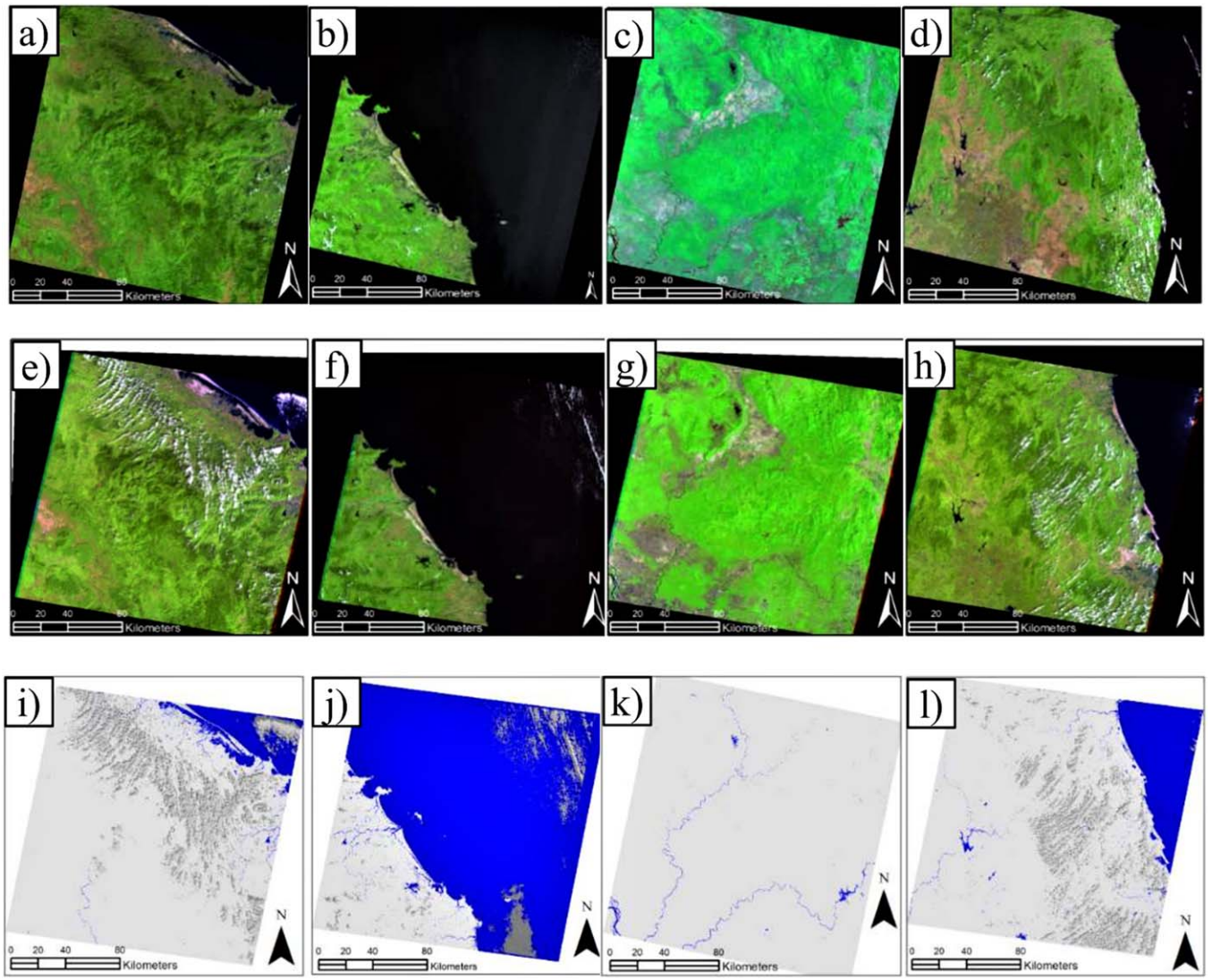

Fig. 2. Pseudo natural color composite of Landsat 8 OLI, Landsat 7 ETM+ scenes, and four Global Inland Water (GIW) dataset scenes: (**a**) LC81250492015024LGN00, (**b**) LC81240492015161LGN00, (**c**) LC81250502015104LGN00, (**d**) LC81240502015065LGN00, (**e**) LE71250492001329SGS00, (**f**) LE71240492001082SGS00, (**g**) LE71250502002044SGS00, (**h**) LE71240502001290EDC00, (**i**) p124r049_WC_20010323.tif, (**j**) p124r050_WC_20011017, (**k**) p125r049_WC_20011125.tif, and (**l**) p125r050_WC_20020213. In the GIW dataset, blue represents surface water, light gray represents land, and dark gray represents cloud.

mountainous with extremely high elevation differences (2000 m) over a relatively small region. In this area, the climate has two distinct seasons: dry and rainy. The rainy season lasts 3 months from September to December and the dry season lasts from January to August. Cultivation of two rice crops forms the major agricultural component in the part of the study area in Vietnam, while only one rice crop is cultivated in the study area in Laos. Water bodies in the study area include artificial reservoirs, natural lakes, streams and rivers, and submerged low land functioning as water reservoirs (Fig. 1). The water level is not stable and varies throughout the year according to the season.

Materials

Four Landsat 8 OLI scenes with different acquisition times from the first half of 2015 were used for the study (Table 1). Cloud coverage in these scenes is low and mostly affects the sea and high mountain areas. Four Landsat 7 scenes with the same path/row number, used to develop the GIW dataset, were selected to achieve results that could be compared to those of the GIW dataset (Feng et al. 2016). Pseudo natural color composites have been used to illustrate our research (Fig. 2), which give clear distinctions between land and water surfaces and provide a suitable means for visual inspection of the subsequent classification.

Procedures

We introduce a new method for water classification using the shape of spectral patterns, which has previously received little attention. The mathematic background of our proposed method builds on the application of SSPs. An SSP is a

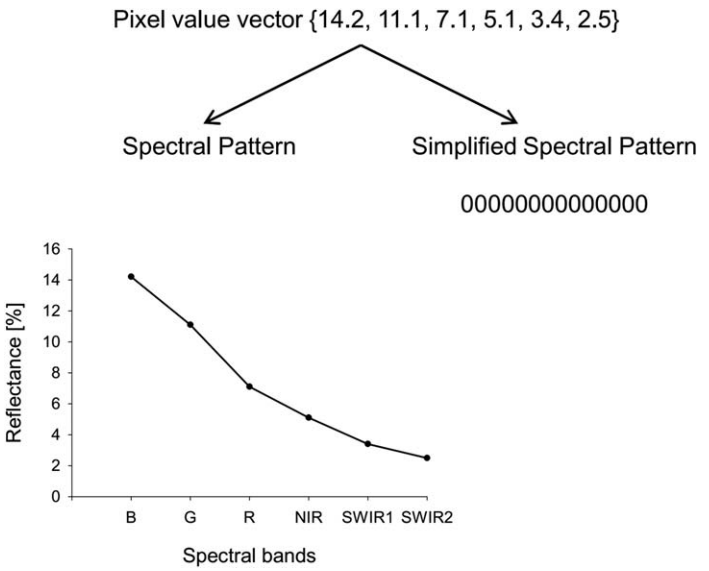

Fig. 3. Example of a transformation of the spectral pattern into a simplified spectral pattern. The curve line on the left site shows full spectral pattern of the given pixel value vector and the fifteen digits on the right site is the simplified spectral pattern.

transformation of the full spectral pattern into a simplified digital form, which allows direct incorporation of the spectral pattern into the classification. The SSP is constructed by nonrepetitive pairwise comparison of reflectance values between two bands of the same image (Charalambides 2002). Given a pixel value vector $B_6 = \{b_1,b_2,b_3,b_4,b_5,b_6\}$, where $b_1, b_2,\ldots,b_6$ denote top of atmosphere (TOA) reflectance of bands 2, 3, 4, 5, 6, and 7 of OLI, and bands 1, 2, 3, 4, 5, and 7 of ETM+ sensor, respectively, the simplified spectral pattern is defined by a new 15 digit vector, shown in Eq. 1.

$$m_{1,2}m_{1,3}m_{1,4}m_{1,5}m_{1,6}m_{2,3}m_{2,4}m_{2,5}m_{2,6}m_{3,4}m_{3,5}m_{3,6}m_{4,5}m_{4,6}m_{5,6} \quad (1)$$

where $m_{i,j}$ is the result of comparison between the reflectance of b_i and b_j and has values of 0 (if $b_j < b_i$.), 1 (if $b_j = b_i$), or 2 ($b_j > b_i$). An example of SSP construction is presented in Fig. 3.

The construction of SSPs for a given image is performed for pixels that have valid digital numbers without saturated values. To perform an automated classification of water bodies using TOA reflectance, we first developed an SSP database for various water body types. Development of the SSP database was realized by interactive visual inspection using a special software tool. A pixel of water area was selected for SSP construction and the remaining pixels in the image with similar SSPs were rendered by a given color. If the SSP being assessed accurately depicts a water area, it is added to the SSP database.

Due to slight differences in spectral characteristics of sensors TM, ETM+, and OLI, we need to develop different databases of SSPs for each sensor and apply them correspondingly during classification.

In some cases, topographic and cloud shadows have spectral characteristics similar to water, complicating water classification (Liao et al. 2014; Carroll et al. 2016; Feng et al. 2016). When using SSPs for experimental water mapping, we found that cloud and topographic shadow affected only some SSPs of water. To separate water bodies from shadows with the same SSP, we used the total reflected radiance index (TRRI) from Duong (1998), using Eq. 2.

$$\mathrm{TRRI} = \frac{\int_1^n I_i \Delta}{\int_1^n I_{\max} \Delta} \quad (2)$$

where I_i is the reflectance of band I, n is the number of spectral bands, $I_{\max}$ is the maximal reflectance for a given quantization level, and Δ is the spectral band difference.

In this research, we computed the TRRI based on red, green, and blue bands using a simplified Eq. 3:

$$\mathrm{TRRI}_{\mathrm{rgb}} = \frac{(b_\mathrm{r} + 2b_\mathrm{g} + b_\mathrm{b})}{2} \quad (3)$$

where b_r, b_g, and b_b stand for the TOA reflectance of the red, green, and blue bands, respectively.

Two methods were used to validate our proposed method: visual interpretation and water extraction using the NDWI. Finally, the results of our method were compared to the GIW dataset provided by the Global Land Cover Facility of the University of Maryland.

Assessment

Simplified spectral patterns of water bodies

In nature, open water is composed of water of highly variable quality. This includes clear water with very low levels of contamination, turbid water with organic or inorganic particles, and water with different levels of nutrient richness (i.e., eutrophic, oligotrophic, or mesotrophic water). Each water type presents a typical spectral pattern associated with a particular SSP. By visual inspection, we identified six SSPs from Landsat 8 OLI and nine SSPs from Landsat 7 ETM+, representative of different water types in the study area. Five common SSPs were observed for both sensors (Fig. 4).

Simple spectral patterns 1 and 7 are associated with water spreading over the sea, reservoirs, lakes, and parts of rivers where water is sufficiently clear and deep that there is no bottom reflectance. The spectral pattern of this water shows a gradual decreasing trend from the blue region to the shortwave infrared region without any local maximum. Simple spectral patterns 2, 5, 8, and 10 have a local maximum in the near infrared band. This maximum value may be affected

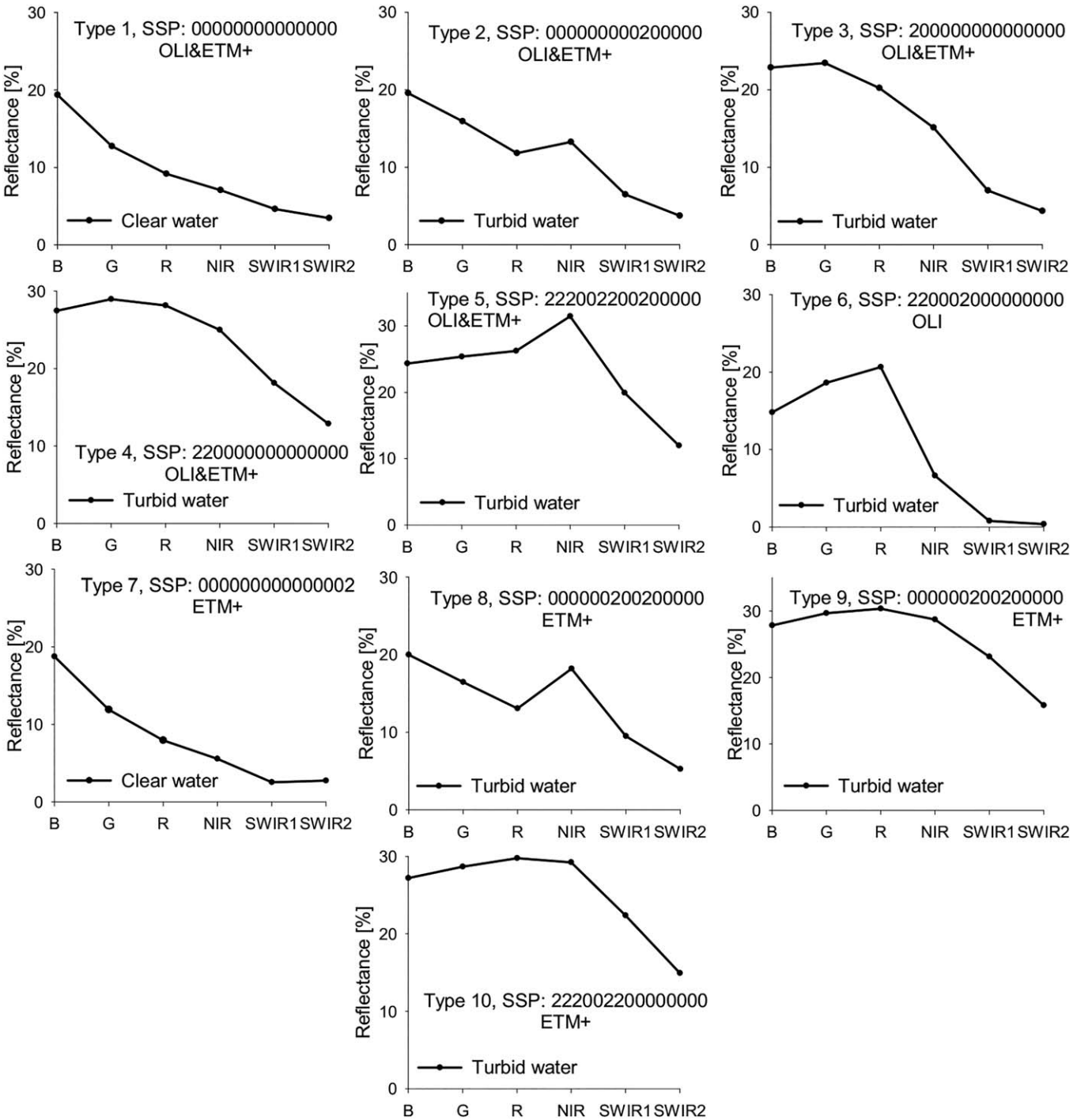

Fig. 4. Simplified spectral patterns from ETM+ and OLI sensors for different water types in the study area.

by the existence of submerged aquatic vegetation along the beach of water bodies or a high concentration of algae. Simple spectral pattern 3 has a local maximum in the green band and a relatively high reflectance in the red band. This characteristic denotes high turbidity of water. Water displaying SSP 3 is distributed in some parts of rivers and natural ponds. Water displaying SSP 4, 6, or 9 has a local maximum in the red band and relatively high reflectance in the blue, green, and red bands. This type of water does not cover large areas and may be associated with a type of turbid water. In some cases, it could be the result of a spectral mixture of water and other ground-based objects.

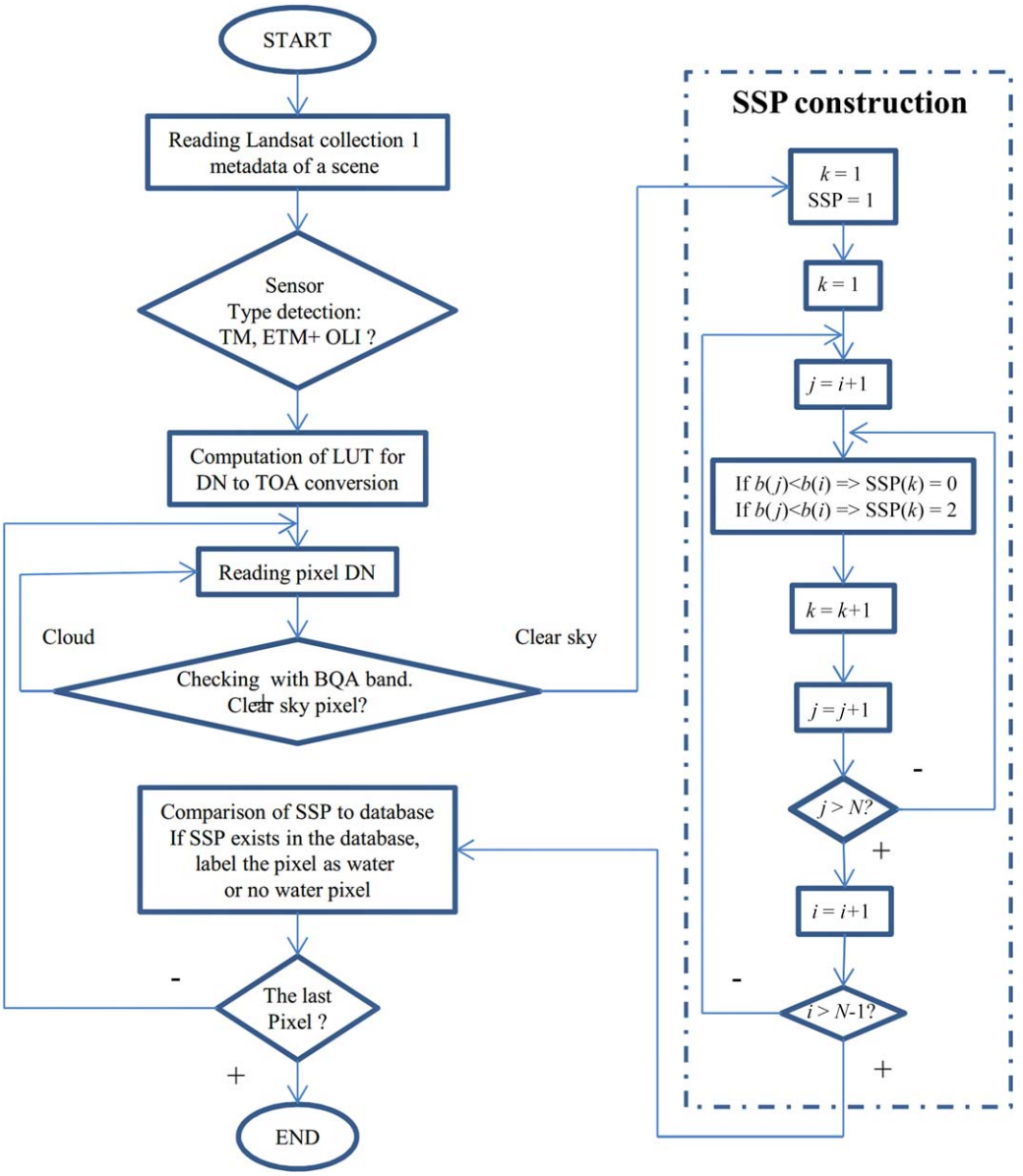

Fig. 5. Schematic diagram of the algorithm for automated water body extraction using SSPs. DN, digital number; SSP, simplified spectral pattern; TOA, top of the atmosphere; TRRI, total reflected radiance index; LUT, look up table for conversion of DN to TOA; *k*, index of SSP members, $k = 1–15$; *i,j*, spectral band number; *b*, pixel value vector; *N*, number of spectral bands, $N = 6$.

Algorithm for automated water body extraction

We propose an algorithm for automated water body extraction, which is schematically presented in Fig. 5. The algorithm was implemented in a software module using the C++ programming language. This module runs in command mode with three basic input parameters: name of the Landsat metadata file, classification rule file, and output file names for classified images. The results were stored in GeoTIFF format. Areas with no associated data were indicated with no data values. This data structure allows fast mosaicking of multiple scenes when mapping a large area, as well as fast per-scene processing. It takes approximately 20 s for the classification of one OLI or ETM+ scene using a Windows-based laptop with Intel Core i7 CPU 2.20 GHz and 8 GB of RAM. The classification rule is organized as a text file composed of classes to be classified. Each class is arranged into a structure that provides information on class name, SSP code, color used for class display, and additional parameters used for classification. The source code of a demonstration module is given in the Supplementary material.

Algorithm testing and validation

We carried out automated water body extraction with the predefined SSPs and proposed algorithm (Fig. 5) for the eight Landsat 8 OLI and Landsat 7 ETM+ scenes (Table 1).

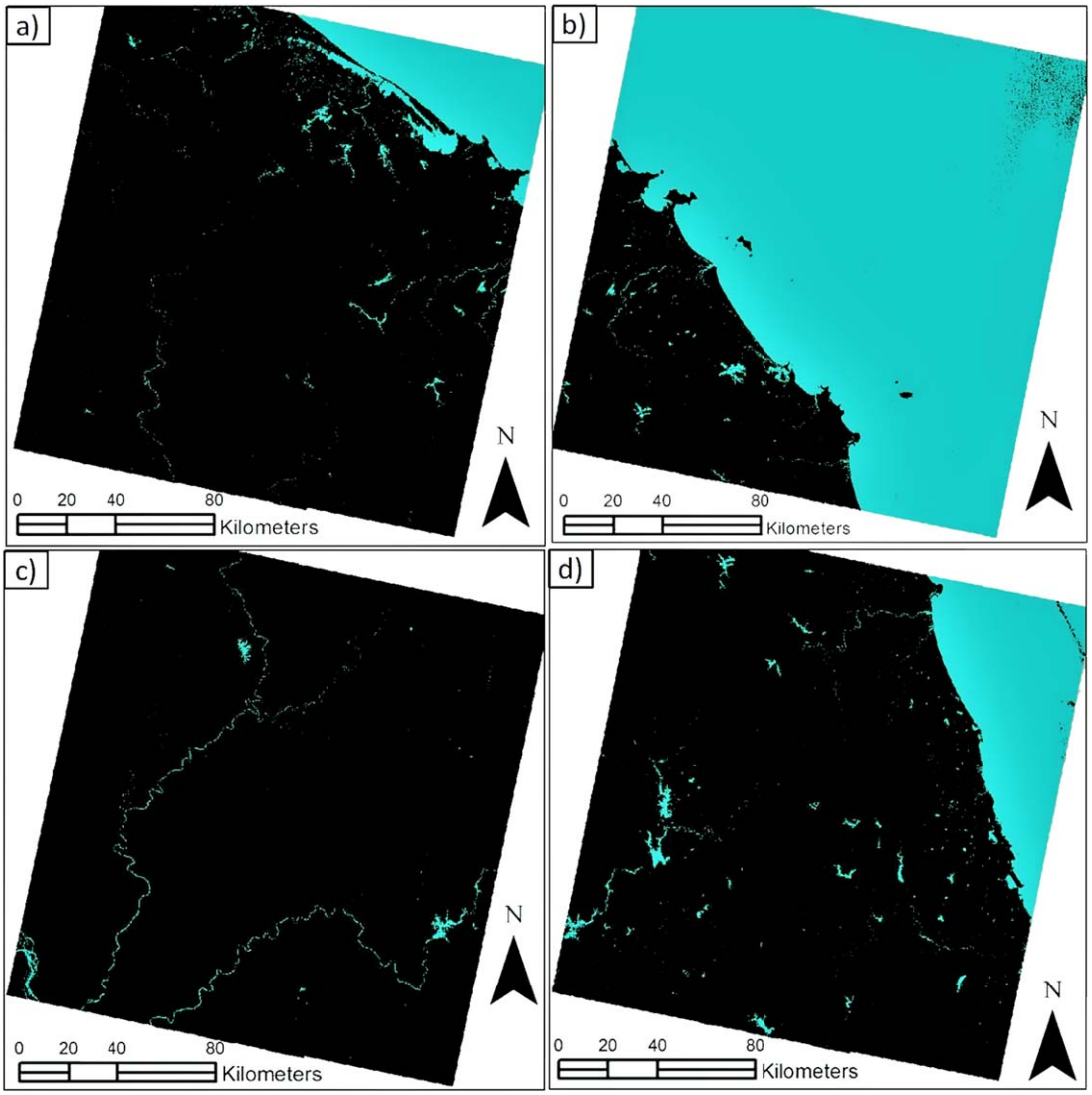

Fig. 6. Results of the automated water body extraction for the Landsat 8 OLI scenes: (**a**) LC81250492015024LGN00, (**b**) LC81240492015161LGN00, (**c**) LC81250502015104LGN00, and (**d**) LC81240502015065LGN00. Water is represented as cyan. Land is shown in black. The black patterns over the sea in (**b**) and (**d**) indicate the presence of cloud.

Although these eight scenes were acquired at different times, the proposed algorithm successfully classified the water bodies (Figs. 6, 7), and the results can be compared to those produced by other methods.

To determine their accuracy, we compared our results with pseudo natural color composites and overlays for two selected windows in scene LC81240492015161LGN00 (Fig. 8). The boundaries of open water bodies extracted by our method followed the waterlines of hydrographical features in the study window, such as reservoirs, ponds, rivers, streams, and sea. Based on a visual inspection, there was good agreement between the automated extraction results and the visual interpretation results. There was agreement over the sea, inland water, and water of different qualities. However, water bodies that were too narrow and water that was too shallow were not correctly extracted, although they were detected by visual interpretation.

One advantage of our method, in comparison to traditional methods, is its capability to differentiate dissimilar water types. In most cases, traditional water classification methods (e.g., thematic classification or spectral indices) only distinguish water extent from land areas. Theoretically, it is possible to use unsupervised or supervised classification techniques to recognize water types, but it is difficult to collect training samples and, thus, it is rarely applied in practice. In our method, water is classified by different SSPs according to water type. When dissimilar SSPs are visualized with different colors, the water body can be mapped by both water extent and water type (Fig. 9). The spatial distribution of different water types can provide useful information about

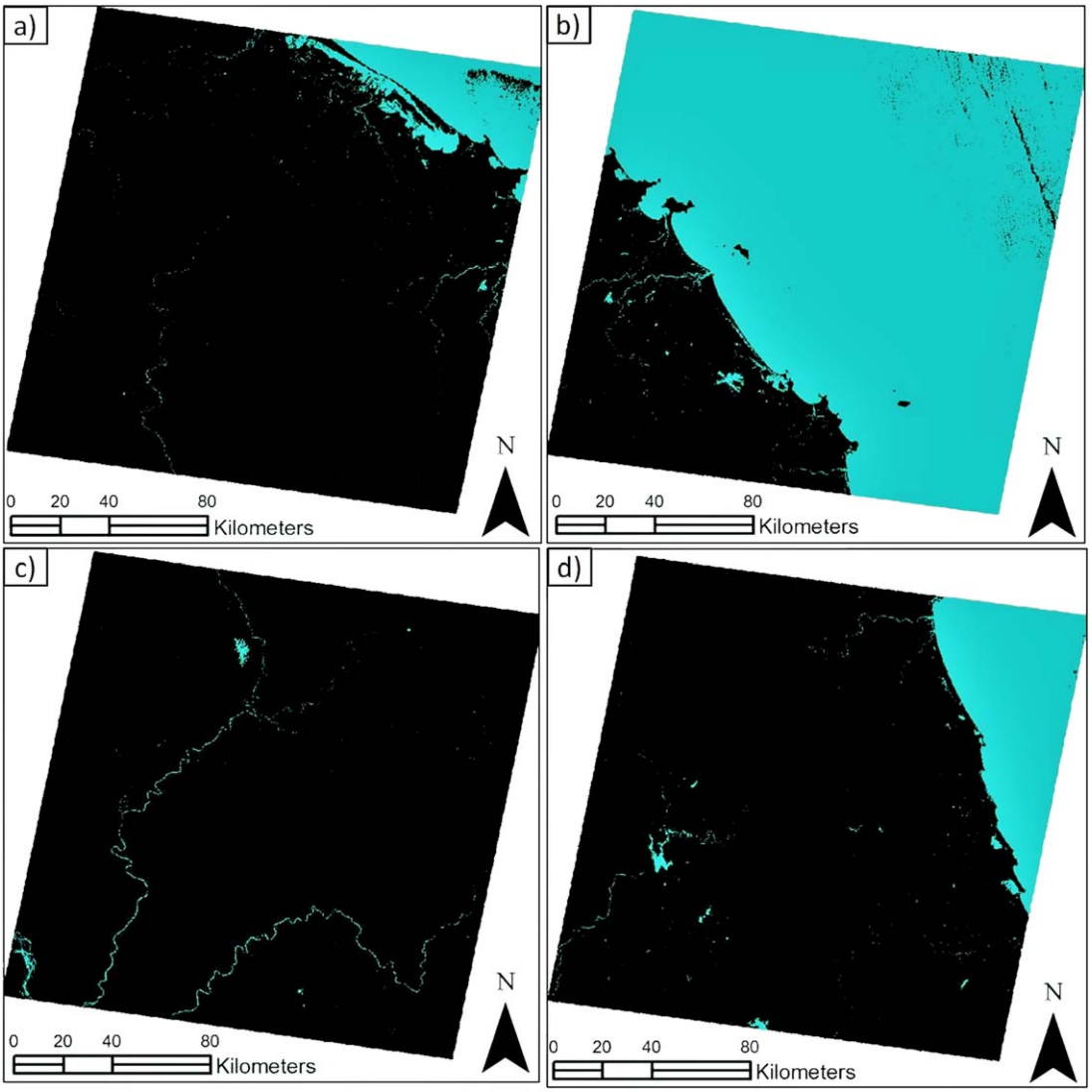

Fig. 7. Results of the automated water body extraction for the Landsat 7 ETM+ scenes: (**a**) LE71250492001329SGS00, (**b**) LE71240492001082SGS00, (**c**) LE71250502002044SGS00, and (**d**) LE71250492001329SGS00. Water is represented as cyan. Land is shown in black. The black patterns over the sea in (**b**) and (**d**) indicate the presence of cloud.

the water environment, and the environment of an area through which a stream passes.

To assess the reliability of our proposed method, we compared our results with other methods both qualitatively, by visual examination, and quantitatively, using the performance index (PI) (Bagli et al. 2004). The PI is computed from Eq. 4:

$$\mathrm{PI}=100-\left(\sum_{i}^{n}\frac{A_i}{N_{\mathrm{pix}}}\right) \tag{4}$$

where A_i is the area of discrepancy and N_{pix} is the area of the reference water body. P_i refers to the ability of a method to detect water relative to other methods.

Assessment with Landsat 8 OLI data

We applied our method to Landsat 8 OLI data and compared the results with those achieved by application of the NDWI. The NDWI is a widely accepted method for water body mapping (Du et al. 2016) and is computed using Eq. 5 (McFeeters 1996):

$$\mathrm{NDWI}=\frac{\rho_{\mathrm{Green}}-\rho_{\mathrm{NIR}}}{\rho_{\mathrm{Green}}+\rho_{\mathrm{NIR}}} \tag{5}$$

where ρ is the TOA reflectance of the respective bands.

A threshold value of 0.07 was selected for the NDWI for the delineation of open water bodies. Figure 10 shows a visual comparison of water classification using NDWI and our proposed method. The comparison of our results

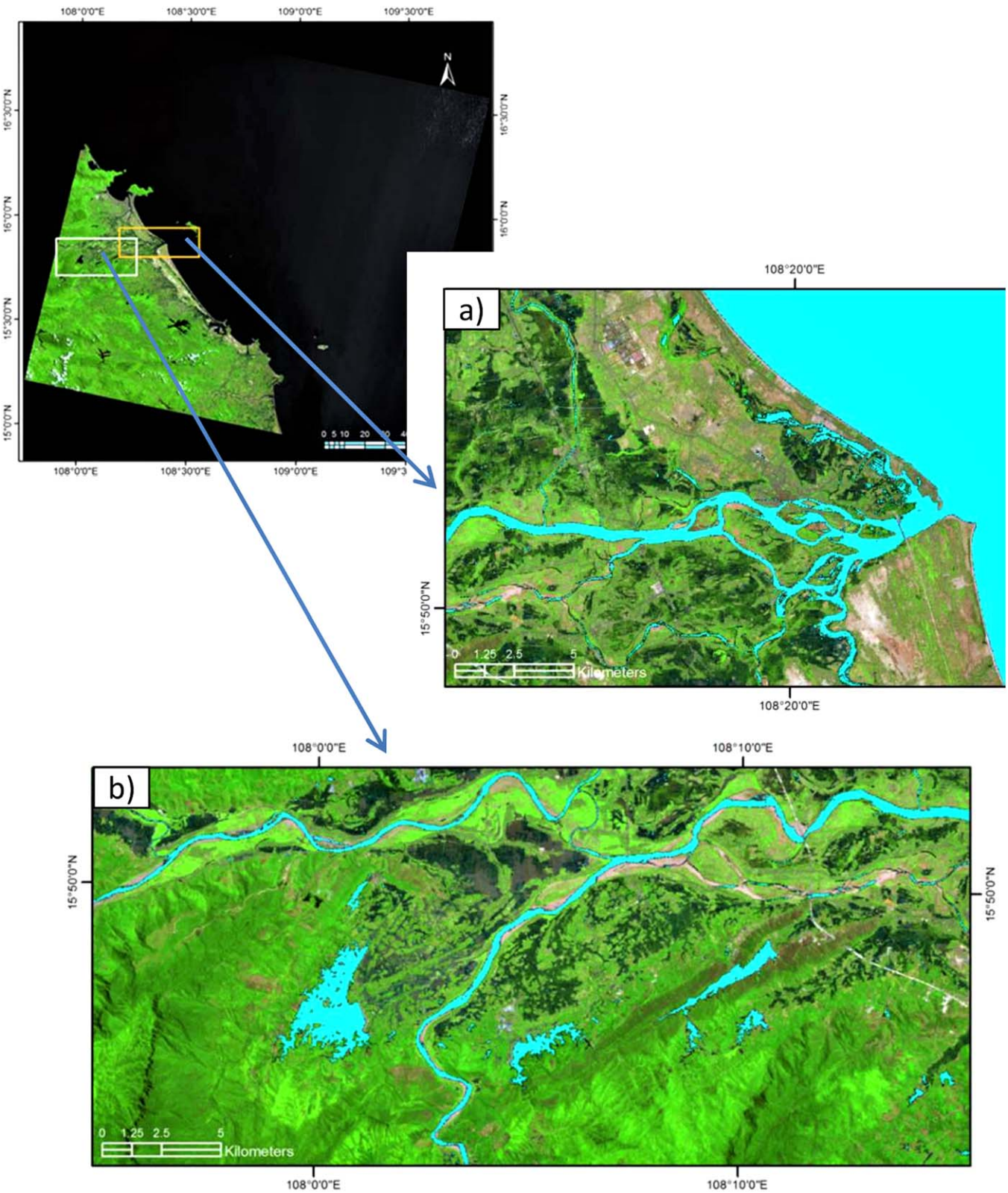

Fig. 8. Layout of the two windows used for detailed study. Window (**a**) covers part of Thu Bon River and Cua Dai estuary. Window (**b**) covers inland water bodies including the Vu Gia River (upper), Thu Bon River (lower), and Khe Tan reservoir (the largest one), as well as several small lakes and ponds. This figure shows the overlay of our classification results over a pseudo natural color composite. Water bodies are colored cyan.

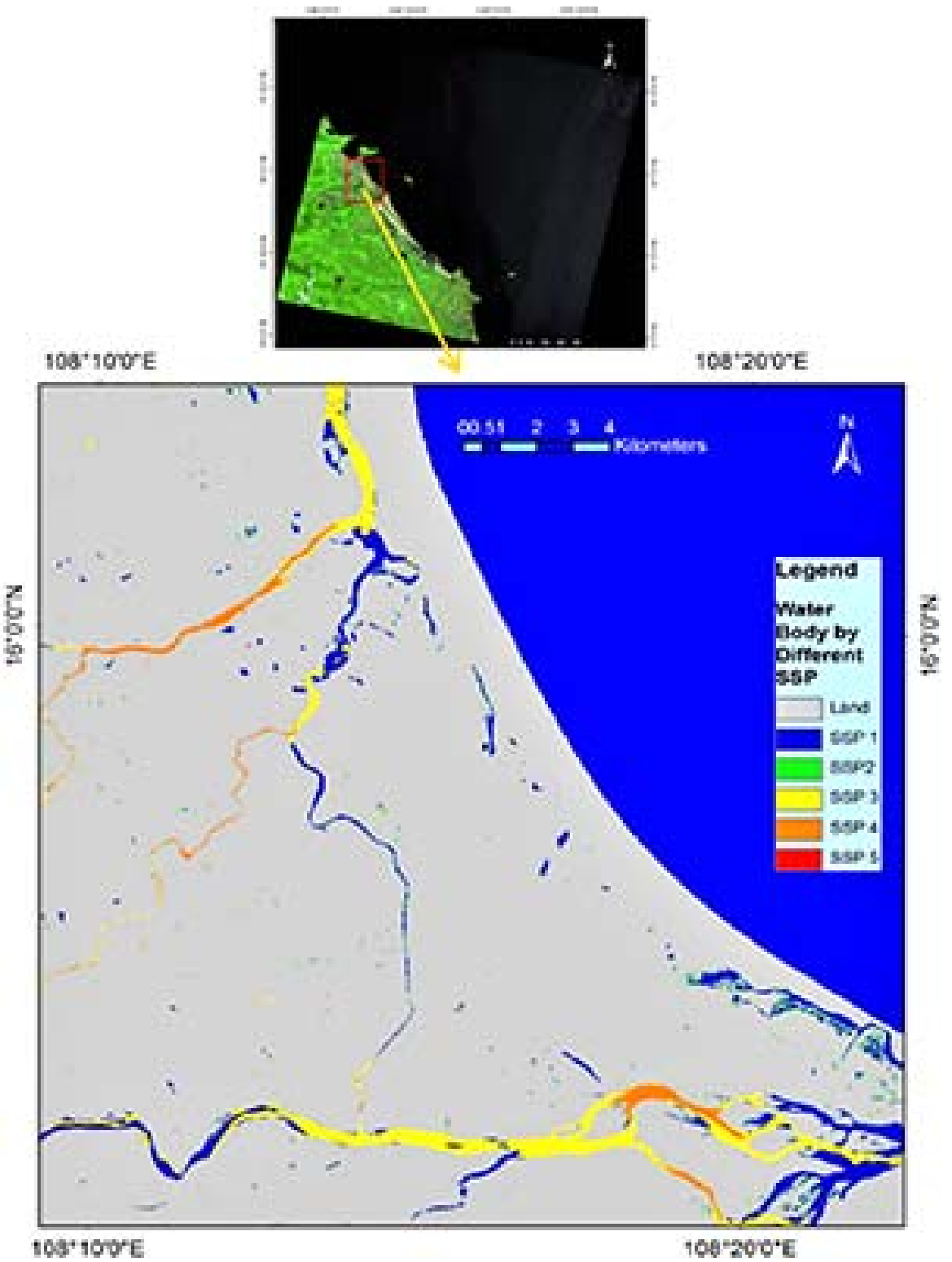

Fig. 9. Example of the spatial distribution of water bodies of different SSPs represented by different colors.

with the NDWI for Landsat 8 OLI images showed high agreement between the two methods. In this case, we did not apply any cloud or shadow mask. Both methods provide good classification of large water surfaces but for smaller ponds, the SSP method yielded more detailed results.

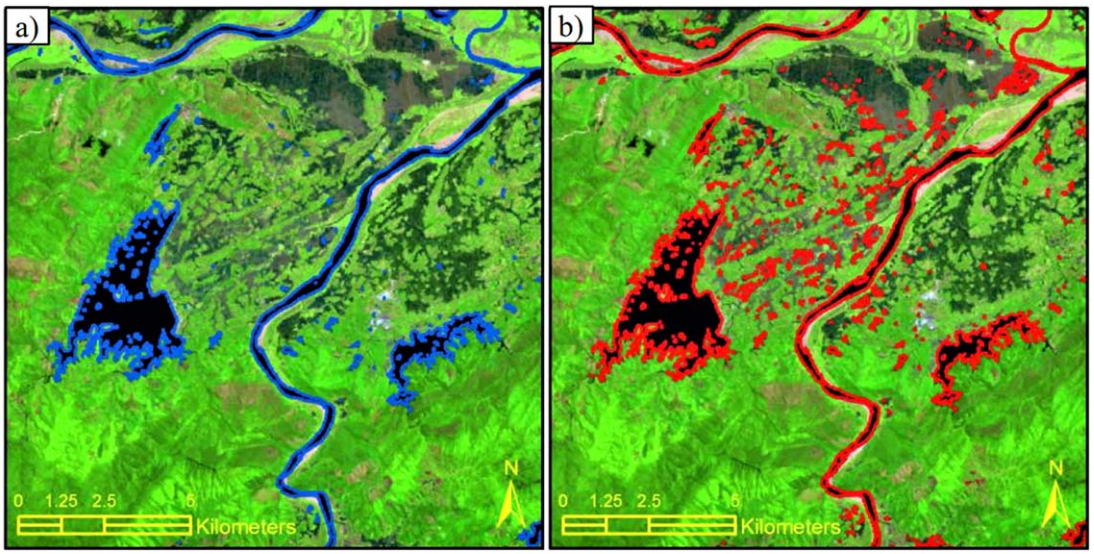

Fig. 10. Comparison of water classification by (**a**) NDWI with water body boundaries highlighted in red color and (**b**) the proposed SSP method with boundaries of water bodies highlighted in blue color using scene LC81240492015161LGN00.

Table 2. Comparison of water area in individual Landsat 8 OLI scenes derived from the simplified spectral pattern (SPP) and normalized difference water index (NDWI) methods and the performance index (PI).

Scene ID	SSP (m^2)	NDWI (m^2)	PI SSP
LC81240492015161LGN00	27,728,235,952	28,206,571,424	98.30
LC81250492015024LGN00	5,547,227,360	5,556,296,224	99.84
LC81240502015065LGN00	3,619,963,804	3,566,228,084	98.49
LC81250502015104LGN00	247,693,500	249,301,800	99.35

Table 3. Comparison of water area in individual Landsat 7 ETM+ scenes derived from the simplified spectral pattern (SSP) method and Global Inland Water (GIW) dataset and performance index (PI).

Scene ID	SSP (m^2)	GIW (m^2)	PI SSP
LE71240492001082SGS00	23,777,297,096	23,186,757,488	97.45
LE71240502001290EDC00	4,315,610,052	4,162,490,528	96.32
LE71250492001329SGS00	2,038,573,432	1,935,378,700	94.67
LE71250502002044SGS00	190,215,000	194,435,100	97.83

To objectively evaluate the performance of our method, we used the water area classification results from the NDWI method as a reference. Because this method has been accepted and validated, we were able to compute the PI of our method for each Landsat 8 OLI scene (Table 2).

Assessment with Landsat 7 ETM+

Furthermore, we compared our results with the GIW dataset published by the University of Maryland. Although all four Landsat 7 ETM+ scenes used for the comparison have high cloud cover, two have particularly high cloud coverage (LE71250492001329SGS00 and LE71240502001290EDC00); therefore, we used a cloud and shadow mask to reduce their impact on the water classification. We kept the cloud and shadow mask from the GIW to maintain consistency between the results of our method and the GIW. The PI was computed for the four scenes using the GIW as reference data (Table 3). Figure 11 presents a visual comparison of the two methods. We observed high agreement between the results for both large and small water bodies.

Accuracy assessment

The classification accuracy was assessed by using 100 validation points generated by a random sampling scheme for each OLI image and compared with visual interpretation of Google Earth high spatial resolution images in 2015 (Table 4.).

Discussion

Methods for mapping water have shown gradual developments in recent years. However, there are still many problems that must be resolved concerning the automation of water mapping, simultaneous detection of both water extent and water type, and analysis using multitemporal image series. Traditional water classification methods require extensive floating-point computation, including surface reflectance estimation, decision tree classification or threshold value computation, cloud and terrain mask generation, etc. (Verpoorter et al. 2012; Feyisa et al. 2014; Jiang et al. 2014; Carroll et al. 2016; Feng et al. 2016). This computation slows the overall performance of water classification. In our method, the computation is based on integers and is kept as simple as possible. The computation in our proposed algorithm is essentially a comparison of the SSP of every pixel to those pre-collected in

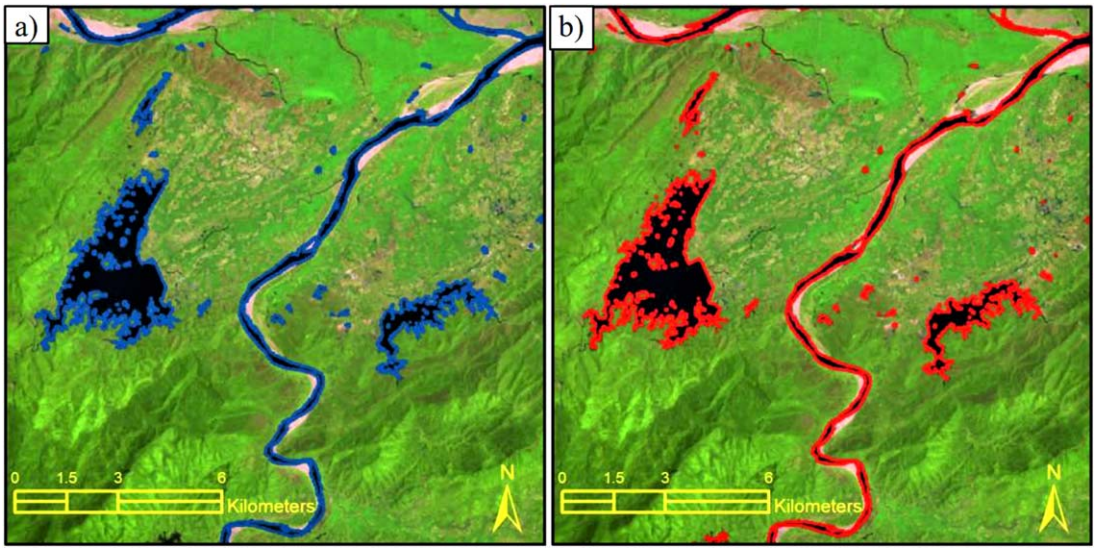

Fig. 11. Comparison of water classification using the (**a**) GIW dataset with water body boundaries highlighted in blue color and (**b**) proposed SSP method with water body boundaries in red color using scene LE71240492001082SGS00. The methods provide almost identical results regarding the detection of small and large water bodies for the area in the selected windows.

Table 4. Accuracy assessment using 100 reference data points for each scene. The reference data are interpreted from high spatial resolution images on the Google Earth.

Image ID	User's accuracy	Producer's accuracy	Kappa coefficient
LC81240492015161LGN00	97.3	97.3	0.89
LC81250492015024LGN00	100.0	90.5	0.92
LC81240502015065LGN00	97.1	91.7	0.91
LC81250502015104LGN00	100.0	100.0	1.0

the SSP database. Thus, the overall performance of the analysis is improved. Due to the simplicity of the algorithm, which does not require special computer resources, and the fact that input image data is available online, it has great potential for large-scale, or even global, water classification using a cloud computing facility such as Google cloud computing.

Our study shows that SSPs enable mapping of both the extent and type of water bodies. This is a unique feature of our proposed method that is not available in other methods. We analyzed Landsat images from different time periods without atmospheric correction, and achieved results comparable to those that have undergone atmospheric correction (i.e., the GIW dataset). This implies that SSPs are stable against variable atmospheric conditions and can enable the automation of water body detection. Since the construction of SSPs is simple, with no complex computation, the classification process is quick and can potentially process a large number of images in a short time. The novelty of our method offers the possibility of fast water body dynamics monitoring in large areas that cannot currently be accessed with existing methods.

The requirements of the input optical image data for water classification vary with the algorithm used. For example, in the GIW dataset, the input data consists of atmospherically corrected Landsat 7 scenes from circa 2000 (Feng et al. 2016). The Automated Water Extraction Index (Feyisa et al. 2014) employs Landsat 5 data after atmospheric correction using the Fast Line-of-site Atmospheric Analysis of Hypercubes method. The Global Wildland Fire Emission Model dataset (Verpoorter et al. 2012) uses GeoCover data circa 2000. Our method can accept any Landsat TM, ETM+, and OLI data for analysis without atmospheric correction. This advantage enables the rapid study of changes in surface water using Landsat archives from the 1980s to the present for any location in the world.

Cloud and cloud shadow are well-known error sources in water mapping, and many algorithms for cloud and shadow masking have been published. Therefore, we do not intend to resolve this problem here. Instead, when working with Landsat data, we use cloud and shadow masks provided by the United States Geological Survey, together with image scenes from The Landsat Collection 1 or pre-collection dataset (https://landsat.usgs.gov/landsat-collections). In this dataset, the C Function of Mask (CFMask) algorithm identifies fill, cloud, cloud confidence, cloud shadow, and snow/ice in Landsat 4–5, Landsat 7, and Landsat 8 scenes for representation in the Quality Assessment band. By accessing the Quality Assessment band, we can easily exclude cloud and shadow from the water classification results.

The success of our method depends on the representativeness of SSPs in the database used for the classification. Theoretically, water bodies across the world have similar spectral reflectance characteristics; however, due to different levels of

pollution by either inorganic or organic contaminants, the number of SSPs may vary between geographic regions. The study area of this research includes various water types characteristic of tropical regions such as seawater, lakes, and ponds, swamps, and rivers and streams. The SSPs collected for the ETM+ and OLI sensors work well for these water bodies. However, there could be other water types present in alpine and high latitude regions. To expand this method for global water mapping, research covering different regions should be conducted, and a common SSP database should be gradually established by adding new SSPs. In this way, a global SSP database representing all water types can be produced to support global water mapping.

Although a number of water extraction methods are used in practice, we suggest that our SPAB method using SSPs is the simplest, because it has minimal requirements for the input image data and uses a straightforward algorithm providing fast computation and comparable accuracy to other standard water extraction methods. Therefore, our method makes a significant contribution to the development of tools for monitoring spatiotemporal changes in water bodies, which are of vital importance for limnologic studies, especially under the current changing global climate.

Conclusion

In this paper, we proposed a new spectral-pattern-analysis-based (SPAB) method for automated water body extraction using simplified spectral patterns (SSPs). This new method allows mapping spatial extent as well as different quality type of surface water. This is a unique capability that traditional water classification methods can rarely afford. The use of simplified spectral patterns instead of image digital numbers for computation results in high performance of automated process, which is very demanding for computer resources when mapping in a large region with massive image data.

Our new method was successfully tested in the central part of Vietnam and southern Laos, an area that encompasses different water body types, including seawater and inland water, with various levels of pollution. Water body were extracted for four Landsat 8 OLI scenes from 2015 and four Landsat 7 ETM+ scenes from 2001 and 2002 to assess the performance on water classification of the new method by comparing with visual interpretation, normalized differential water index (NDWI) thresholding, and the Global Inland Water (GIW) dataset provided by the Global Land Cover Facility of the University of Maryland. It was found that the classification results of our new method agree by more than 98.0% with NDWI results by more than 95.0% with the GIW dataset. When assessed by reference data from high spatial resolution images on Google Earth, we also achieved average user's accuracy of 98%, producer's accuracy of 94.8%, and average Kappa coefficients of 0.93.

In the next phase of this research, study will be expanded to other geographical areas to develop a SSPs database, which will represent global water body types. This goal requires a range of experiments to be conducted in different geographical regions including subarctic belt and Tibetan Plateau.

References

Bagli, S., P. Soille, and E. Fermi. 2004. Automatic delineation of shoreline and lake boundaries from Landsat satellite images Proceedings of initial ECO-IMAGINE GI and GIS for Integrated Coastal Management, Seville.

Canty, M. J. 2014. Image analysis, classification and change detection in remote sensing: With algorithms for ENVI/IDL and Python. CRC Press.

Carroll, M., M. Wooten, C. DiMiceli, R. Sohlberg, and M. Kelly. 2016. Quantifying surface water dynamics at 30 meter spatial resolution in the North American high northern latitudes 1991–2011. Remote Sens. **8**: 622. doi: 10.3390/rs8080622

Charalambides, C. A. 2002. Permutation and combinations, p. 40–42. In C. A. Charalambides [ed.], Enumerative combinatorics, Chapman&Hall/CRC.

Du, Y., Y. Zhang, F. Ling, Q. Wang, W. Li, and X. Li. 2016. Water bodies' mapping from sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. Remote Sens. **8**: 354. doi:10.3390/rs8040354

Duong, N. D. 1998. Total reflected radiance index - An index to support land cover classification. p. 16–20. Proceedings of the 19th Asian Conference on Remote Sensing. Available from http://a-a-r-s.org/aars/proceeding/ACRS1998/Papers/LU98-5.htm

Duong, N. D. 2012. Water body extraction from multi spectral image by spectral pattern analysis, p. 181–186. *In* International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives.

Feng, M., J. O. Sexton, S. Channan, and J. R. Townshend. 2016. A global, high-resolution (30-m) inland water body dataset for 2000: First results of a topographic–spectral classification algorithm. Int. J. Digit. Earth **9**: 113–133. doi:10.1080/17538947.2015.1026420

Feyisa, G. L., H. Meilby, R. Fensholt, and S. R. Proud. 2014. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. Remote Sens. Environ. **140**: 23–35. doi:10.1016/j.rse.2013.08.029

Gardelle, J., P. Hiernaux, L. Kergoat, and M. Grippa. 2010. Less rain, more water in ponds: A remote sensing study of the dynamics of surface water from 1950 to present in pastoral Sahel (Gourma region, Mali). Hydrol. Earth Syst. Sci. **14**: 309–324. doi:10.5194/hess-14-309-2010

Haas, E. M., E. Bartholomé, and B. Combal. 2009. Time series analysis of optical remote sensing data for the mapping of temporary surface water bodies in sub-Saharan

western Africa. J. Hydrol. **370**: 52–63. doi:10.1016/j.jhydrol.2009.02.052

Ji, L., L. Zhang, and B. Wylie. 2009. Analysis of dynamic thresholds for the Normalized Difference Water Index. Photogramm. Eng. Remote Sens. **75**: 1307–1317. doi:10.14358/PERS.75.11.1307

Jiang, H., M. Feng, Y. Zhu, N. Lu, J. Huang, and T. Xiao. 2014. An automated method for extracting rivers and lakes from Landsat imagery. Remote Sens. **6**: 5067–5089. doi:10.3390/rs6065067

Liao, A., and others. 2014. High-resolution remote sensing mapping of global land water. Sci. China Earth Sci. **57**: 2305–2316. doi:10.1007/s11430-014-4918-0

Livingston, B., P. Frazier, and J. Louis. 2000. Remote sensing of riverine water bodies, p. 1119–1123. *In* Proceedings of the 10th Australasian Remote Sensing and Photogrammetry Conference, Adelaide, SA, Australia. Citeseer. Available from http://csusap.csu.edu.au/~jlouis/research/papers/arspc2000/10arspc_paper102_00.pdf

McFeeters, S. K. 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int. J. Remote Sen. **17**: 1425–1432. doi:10.1080/01431169608948714

Papa, F., C. Prigent, F. Aires, C. Jimenez, W. B. Rossow, and E. Matthews. 2010. Interannual variability of surface water extent at the global scale, 1993–2004. J. Geophys. Res. Atmos. **115**: D12111. doi:10.1029/2009JD012674

Proud, S. R., R. Fensholt, L. V. RasmussenInge, and I. Sandholt. 2011. Rapid response flood detection using the MSG geostationary satellite. Int. J. Appl. Earth Obs. Geoinf. **13**: 536–544. doi:10.1016/j.jag.2011.02.002

Ryu, J.-H. H., J.-S. S. Won, and K. D. Min. 2002. Waterline extraction from Landsat TM data in a tidal flat: A case study in Gomso Bay, Korea. Remote Sens. Environ. **83**: 442–456. doi:10.1016/S0034-4257(02)00059-7

Sheng, Y., C. A. Shah, and L. C. Smith. 2008. Automated image registration for hydrologic change detection in the lake-rich arctic. IEEE Geosci. Remote Sens. Lett. **5**: 414–418. doi:10.1109/LGRS.2008.916646

Smith, L. C. 1997. Satellite remote sensing of river inundation area, stage, and discharge: A review. Hydrol. Process. **11**: 1427–1439. doi:10.1002/(SICI)1099-1085(199708)11:10<1427::AID-HYP473>3.0.CO;2-S

Song, C., B. Huang, L. Ke, and K. S. Richards. 2014. Remote sensing of alpine lake water environment changes on the Tibetan Plateau and surroundings: A review. ISPRS J. Photogramm. Remote Sens. **92**: 26–37. doi:10.1016/j.isprsjprs.2014.03.001

Tholey, N., S. Clandillon, and P. De Fraipont. 1997. The contribution of spaceborne SAR and optical data in monitoring flood events: Examples in northern and southern France. Hydrol. Process. **11**: 1409–1413. doi:10.1002/(SICI)1099-1085(199708)11:10<1409::AID-HYP531>3.0.CO;2-V

Verpoorter, C., T. Kutser, and L. J. Tranvik. 2012. Automated mapping of water bodies using Landsat multispectral data. Limnol. Oceanogr.: Methods **10**: 1037–1050. doi:10.4319/lom.2012.10.1037

Xu, H. 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. Int. J. Remote Sens.**27**: 3025–3033. doi:10.1080/01431160600589179

Zhang, G., G. Zheng, Y. Gao, and J. Li. 2017*a*. Automated water classification in the Tibetan Plateau using Chinese GF-1 WFV data Automated Water Classification in the Tibetan Plateau Using Chinese GF-1 WFV Data.. Photogramm. Eng. Remote Sens. **83**: 33–43. doi:10.14358/PERS.83.7.415

Zhang, G., and others. 2017*b*. Extensive and drastically different alpine lake changes on Asia's high plateaus during the past four decades. Geophys. Res. Lett. **44**: 252–260. doi:10.1002/2016GL072033

Zhang, W. W., and S. Sriharan. 2005. Using hyperspectral remote sensing for land cover classification, p. 261–270. *In* Fourth International Asia-Pacific Environmental Remote Sensing Symposium 2004: Remote Sensing of the Atmosphere, Ocean, Environment, and Space. International Society for Optics and Photonics.

Acknowledgments

The authors would like to thank Mr. Nguyen Ba Thi for coding the water classification module and the anonymous reviewers whose comments have greatly improved the manuscript. This research was funded by the Vietnam Academy of Science and Technology (grant number VAST05.04/15-16) and Vietnam National Foundation for Science and Technology Development (grant number 105.08-2014.14).

Conflict of Interest

None declared.

Submitted 19 March 2017
Revised 11 September 2017
Accepted 13 September 2017

Associate editor: Xiao Hua Wang