Remote Sensing of Environment 140 (2014) 23–35

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery

Gudina L. Feyisa [a,*], Henrik Meilby [a], Rasmus Fensholt [b], Simon R. Proud [b]

[a] Department of Food and Resource Economics, University of Copenhagen, Rolighedsvej 23, DK-1958 Frederiksberg C, Denmark
[b] Department of Geosciences and Natural Resource Management, University of Copenhagen, Oester Voldgade 10, DK-1350 Copenhagen K, Denmark

ARTICLE INFO

Article history:
Received 16 April 2013
Received in revised form 3 August 2013
Accepted 21 August 2013
Available online 17 September 2013

Keywords:
Classification accuracy
Threshold stability
Subpixel
Mixed pixel

ABSTRACT

Classifying surface cover types and analyzing changes are among the most common applications of remote sensing. One of the most basic classification tasks is to distinguish water bodies from dry land surfaces. Landsat imagery is among the most widely used sources of data in remote sensing of water resources; and although several techniques of surface water extraction using Landsat data are described in the literature, their application is constrained by low accuracy in various situations. Besides, with the use of techniques such as single band thresholding and two-band indices, identifying an appropriate threshold yielding the highest possible accuracy is a challenging and time consuming task, as threshold values vary with location and time of image acquisition. The purpose of this study was therefore to devise an index that consistently improves water extraction accuracy in the presence of various sorts of environmental noise and at the same time offers a stable threshold value. Thus we introduced a new Automated Water Extraction Index (AWEI) improving classification accuracy in areas that include shadow and dark surfaces that other classification methods often fail to classify correctly. We tested the accuracy and robustness of the new method using Landsat 5 TM images of several water bodies in Denmark, Switzerland, Ethiopia, South Africa and New Zealand. Kappa coefficient, omission and commission errors were calculated to evaluate accuracies. The performance of the classifier was compared with that of the Modified Normalized Difference Water Index (MNDWI) and Maximum Likelihood (ML) classifiers. In four out of five test sites, classification accuracy of AWEI was significantly higher than that of MNDWI and ML (*P-value* < 0.01). AWEI improved accuracy by lessening commission and omission errors by 50% compared to those resulting from MNDWI and about 25% compared to ML classifiers. Besides, the new method was shown to have a fairly stable optimal threshold value. Therefore, AWEI can be used for extracting water with high accuracy, especially in mountainous areas where deep shadow caused by the terrain is an important source of classification error.

1. Introduction

Environmental changes and their impacts on natural systems and human societies are topics of research in a wide range of scientific fields. Surface water is among the most vital earth resources undergoing changes in time and space as a consequence of land use/cover (LULC) changes, climate change and other forms of environmental changes in many parts of the world. The ecological, social, health and economic effects of surface water changes have been the subject of academic study for many years (Alderman, Turner, & Tong, 2012; Bond, Lake, & Arthington, 2008; Charron et al., 2004; Kondo et al., 2002; Lake, 2003; Li, Wu, Dai, & Xu, 2012); Sun, Sun, Chen, and Gong (2012). Changes in surface water may result in disasters such as flooding, outbreaks of waterborne disease and water shortage in dry tropical areas, which may involve loss of lives. Timely monitoring and delivery of data on the dynamics of surface water are, therefore, essential for policy and decision making processes (Giardino, Bresciani, Villa, & Martinelli, 2010; Morss, Wilhelmi, Downton, & Gruntfest, 2005).

Remote sensing has become an important source of information in analyzing and delivering data on changes in different earth resources, and surface water in particular. Examples of studies applying remote sensing and GIS techniques for various applications in relation to water resources include flood hazard/damage assessment and management (Dewan, Islam, Kumamoto, & Nishigaki, 2007; Ji, Zhang, & Wylie, 2009; Proud, Fensholt, Rasmussen, & Sandholt, 2011), change in surface water resources (Gardelle, Hiernaux, Kergoat, & Grippa, 2009; Haas, Bartholomé, & Combal, 2009; Prigent et al., 2012), water quality assessment and monitoring (Guttler, Niculescu, & Gohin, 2013; He et al., 2012; Novoa et al., 2012), and water-related disease epidemiology (Charoenpanyanet & Chen, 2008; Dambach et al., 2012; Lacaux, Tourre, Vignolles, Ndione, & Lafaye, 2007).

Satellite sensors of varying spatial, temporal and spectral resolution have been used to extract and analyze information regarding surface water. Landsat satellites are among the most widely used optical

* Corresponding author. Tel.: +45 91414185; fax: +45 353 31508.
E-mail addresses: fgudina@gmail.com, fgudina@ifro.ku.dk (G.L. Feyisa), heme@ifro.ku.dk (H. Meilby), rf@geo.ku.dk (R. Fensholt), srp@geo.ku.dk (S.R. Proud).

0034-4257/$ – see front matter
http://dx.doi.org/10.1016/j.rse.2013.08.029

sensors in surface water and other environmental research. The use of these remotely sensed data commonly starts with classification of land use/cover types. Common water classification methods for optical imagery could be categorized into four basic types (Ji et al., 2009): (a) thematic classification (Lira, 2006), (b) linear unmixing (Sethre, Rundquist, & Todhunter, 2005), (c) single-band thresholding (Jain, Singh, Jain, & Lohani, 2005) and (d) two-band spectral water indices (Jain, Saraf, Goswami, & Ahmad, 2006; McFeeters, 1996; Rogers & Kearney, 2004; Xu, 2006). Combinations of various methods are also proposed to improve water extraction accuracies. Examples are, Jiang, Qi, Su, Zhang, and Wu (2012), Sheng, Shah, and Smith (2008), Sun et al. (2012) and Verpoorter, Kutser, and Tranvik (2012). Single band thresholding and two-band indices are commonly used water extraction methods because of ease of use and the fact that these methods are computationally less time-consuming than alternative approaches (Ryu, Won, & Min, 2002).

McFeeters (1996) introduced the Normalized Difference Water Index (NDWI) to delineate open water features using the *green* (*band* 2) and *near-infrared* (*band* 4) of Landsat TM. Rogers and Kearney (2004) used another NDWI for water extraction where they applied *bands* 3 and 5 of Landsat TM. McFeeters (1996) proposed a threshold of 0 for extracting surface water using the raw digital number of Landsat, where all positive NDWI values would be classified as water and negative values as nonwater. However, Xu (2006) found that the NDWI cannot efficiently suppress the signal from built-up surfaces and using an NDWI threshold of 0 does not accurately enable discriminating built-up surfaces from water pixels. Xu (2006) therefore proposed another index, called Modified Normalized Difference Water Index (MNDWI), where McFeeters (1996) NDWI was modified by replacing *band* 4 by *band* 5 of Landsat 5 TM. The MNDWI of Xu (2006) is one of the most widely used water indices for various applications, including surface water mapping, land use/cover change analyses and ecological research (Davranche, Lefebvre, & Poulin, 2010; Duan & Bastiaanssen, 2013; Hui, Xu, Huang, Yu, & Gong, 2008; Poulin, Davranche, & Lefebvre, 2010).

Even though a number of water extraction techniques are described in the literature, the choice between them is constrained by accuracy problems. Environmental monitoring and change detection techniques such as post-classification comparison are likely to be less reliable when classifiers of low accuracy are used (Congalton & Green, 2009; Mucher, Steinnocher, Kressler, & Heunks, 2000). For instance, in a study focusing on water dynamics monitoring, Ji et al. (2009) faced two major problems in appropriately using water indices: first, the results obtained using different indices were inconsistent and unreliable; second, the threshold values applied to distinguish water from nonwater were unstable, varying with scene and locations. These authors compared four different water indices using simulated datasets of four satellite sensors: Landsat ETM+, Système Pour l'Observation de la Terre (SPOT), the Advanced Space-borne Thermal Emission and Reflection radiometer (ASTER), and the Moderate Resolution Imaging Spectroradiometer (MODIS), aiming to identify the best method for delineating water features. Among the four alternatives, they found that the MNDWI performed best in delineating water, and featured the most stable threshold.

Water classification accuracy problems may be especially pronounced in areas where the background land cover includes low albedo surfaces such as asphalt roads in urban areas, and shadows from mountains, buildings and clouds. The presence of shadows may cause misclassification due to the similarity in reflectance patterns, and this may lessen the accuracy of surface water mapping and change analysis (Frey, Huggel, Paul, & Haeberli, 2010; Verpoorter et al., 2012; Xu, 2006). In environments where nonwater dark surfaces are found, simple classification methods such as two-band water indices and single-band thresholding may not sufficiently and accurately distinguish between water pixels and nonwater dark surfaces, particularly shadows (Verpoorter et al., 2012). In a study of land cover dynamics using Landsat TM data, we noted accuracy problems due to failure of existing water extraction methods in accurately distinguishing water from shadows and low albedo urban surfaces. Particularly, no existing water index was able to automatically separate water and shadowed surfaces. In this paper, therefore, we introduce a multiple-band index called Automated Water Extraction Index (AWEI), with the objectives to: (a) improve accuracy of surface water mapping by automatically suppressing classification noise from shadow and other nonwater dark surfaces, and (b) test the robustness of the new method under different environmental conditions and evaluate its relative accuracy in comparison with existing classification techniques.

2. Study areas and data sources

2.1. Test sites

The accuracy and robustness of the Automated Water Extraction Index (AWEI) were tested considering several lakes and other water bodies in different environmental conditions ranging from humid temperate through sub-tropical to tropical dry regions. The test water bodies were obtained from five different countries: Denmark, Switzerland, Ethiopia, South Africa and New Zealand. The water bodies that include small freshwater reservoirs, large lakes, harbors and the sea differ with regard to depth, turbidity, chemical composition and surface appearance. A summary of the basic characteristics of the test sites is shown in Table 1.

The test sites were deliberately selected so that the sub-scenes consist of complex surface features, such as hill shade, built-up areas and other dark surfaces as background to the water bodies. The test sites in Switzerland, Ethiopia and South Africa are characterized by the presence of built-up surfaces and shadows of mountains. The site in Denmark also consists primarily of urban background but with no major shadow problems since the terrain is predominantly flat and tall buildings in the urban area are rare. The test site in New Zealand consists of mountain slopes with deep shadows, but no major urban surfaces are included.

In addition to the five test sites for which detailed accuracy analyses and comparisons were carried out, further validation of the robustness of the new index was undertaken considering shadow-dominated water bodies in Norway, rivers with urban surfaces and shadows from tall buildings in Shanghai, China, and several crater lakes with built-up background surfaces in Bishoftu, Ethiopia. However, these additional test sites were not analyzed in detail and classification output from these sites is not included in the Results section; instead, the classification maps are included in Appendix A for visual inspection of classification accuracy.

2.2. Landsat images

Landsat 5 TM images were acquired from USGS GLOVIS portal (United States Geological Survey (USGS), 2012). All Landsat images used are of product type L1T and with a scene quality score of 9, which means perfect scenes with no errors detected. The images were also georeferenced with precision better than 0.4 pixels (NASA, 2012). The sub-scenes were all free of clouds. Descriptions of the Landsat images are presented in Table 2.

2.3. Reference data

Reference data used in accuracy assessment are described in Table 2. For the test site in Denmark, colored Digital Orthophoto Quadrangles (DOQs) from year 2010 were used as reference. These aerial photos have a spatial resolution of 12.5 cm and location accuracy better than 0.5 m (COWI, 2010). For the four other test sites, high spatial resolution images provided by Google Earth™ were used for reference. The acquisition dates of the reference data and the Landsat 5 TM images were