MAPPING LAKES ON THE TIBETAN PLATEAU WITH LANDSAT IMAGERY AND OBJECT-BASED IMAGE ANALYSIS

K. Korzeniowska [a b], O. Korup [b]

[a] BSF Swissphoto GmbH, 12529 Schönefeld, Germany - karolina.korzeniowska@bsf-swissphoto.com
[b] Geohazards Research Group, University of Potsdam, 14476 Potsdam, Germany – korup@geo.uni-potsdam.de

KEY WORDS: Tibetan Plateau, lakes, LANDSAT, SRTM, MNDWI, OBIA, changes detection

ABSTRACT:

The Tibetan Plateau, the world's largest orogenic plateau, hosts thousands of lakes that play a prominent role as water resources, environmental archives, and sources of natural hazards such as glacier lake outburst floods. Previous studies reported that the size of the lakes on the Tibetan Plateau has been changing rapidly in recent years, possibly because of contemporary global warming. Tracking these changes systematically from remote sensing data offers new challenges and opportunities for automated classification methods such as object-based image analysis (OBIA). We present a method for an automatic mapping of lakes combining LANDSAT images, Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM), and OBIA. We tested our method for most of the area of the Tibetan Plateau where lakes formed in tectonic depressions or blocked by glaciers and sediments have different spectral colours, and different physical states of water (frozen and not-frozen). We analysed images obtained in 1995 and 2015. For detecting the lakes we used the Modified Normalized Difference Water Index (MNDWI) combined with OBIA and a slope map derived from the DEM. Our classification of lakes with an area >10 km² derived 323 water bodies with a total area of 31,258 km², which is 2.6% of the analysed area in 2015. The same number of lakes had covered only 24,892 km² in 1995, so that lakes grew by ~26% in the past two decades on average. The classification had an overall, producer's and user's accuracies of 0.98, and a Cohen's kappa of 0.98, and may be useful step towards quantifying regional-scale hydrological budgets.

1. INTRODUCTION

The Tibetan Plateau is the world's largest orogenic plateau with a mean elevation exceeding 4000 m above sea level (a.s.l.), and often named as "the Roof of the World" or "the 3rd Pole of the Earth" (Song et al. 2014), (Wang 2016). The Plateau is surrounded by high mountains ranges such as the Himalayas, the Kunlun Shan, Pamir, and Qilian Shan (Liu and Chen 2000), and hosts glaciers and thousands of lakes that play a prominent role as water resources, environmental archives, and sources of natural hazards such as glacier lake outburst floods.
Previous studies reported that the Tibetan Plateau could be one of the most sensitive places to global warming (Zhang et al. 2011). Symptoms attributed to rising temperatures on the Tibetan Plateau include retreating glaciers (Liu and Chen 2000), (Yang et al. 2008), and the degradation of permafrost (REF). A number of studies reported that the size of the lakes on the Tibetan Plateau has been changing rapidly in recent years, possibly because of a warming climate. Some studies (Yang et al. 2008), (Song et al. 2014) hold that meltwater from the surrounding glaciers significantly changes the size distribution of these lakes. In this context, satellite-based monitoring offers repeatability of images and standardized output of results. Tracking lake changes systematically also offers new challenges and opportunities for automated classification methods such as object-based image analysis (OBIA). Accurate and reliable automatic mapping of landforms or landform elements from remote sensing data decreases subjective operator bias, and allows rapid investigation of large regions. However, the training of such algorithms still requires carefully designed strategies.
We address this issue and present a semi-automatic method for classifying large lakes on the Tibetan Plateau from LANDSAT images and digital elevation models. Our approach is universal to detecting and classifying lakes prone to seasonal ice cover, as is the case on the Tibetan Plateau. We use the modified normalized difference index (MNDWI) to detect water areas, and applied OBIA to extract lake boundaries and distinguish them from rivers and glaciers.

2. PREVIOUS METHODS FOR DETECTING WATER AREAS

2.1 Water detection algorithms

Methods for extracting water areas from remote sensing data involve analysing objects features such as colour, texture, and shape; thresholding of image bands and band ratio: or spectral water indexes in both supervised and unsupervised classifications (Song et al. 2014).The most common approaches use thresholding of a single band, or a ratio of bands, because they are easy and quick to use (Ryu, Won, and Min 2002), (Song et al. 2014). However, simple thresholding often yields inaccurate classifications that methods based on segmentation or spectral matching hope to improve (Jawak, Kulkarni, and Luis 2015). Frazier et al. (2000) proposed a method applying single-band (ρ) density slicing of LANDSAT images on each individual band separately to determine the most useful ones. They found that the highest accuracy in classifying water areas was based on the ρ short-wave infrared. They also compared the results of band slicing with a maximum likelihood classifier on six bands of LANDSAT for their study area and found that the maximum likelihood approach only slightly outperformed the method of simple ρ density slicing. McFeeters (1996) found that using the green ρ and Near Infrared (NIR) ρ emphasised water areas on top of other objects of the land surface, and proposed the normalized difference water index (NDWI) to delineate water areas: $NDWI = (\rho_{\text{Green}} - \rho_{\text{NIR}})/(\rho_{\text{Green}} + \rho_{\text{NIR}})$, where the water surfaces have positive values, and other objects have negative values. Rogers and Kearney (2004) suggested using

the combination of red ρ and SWIR1 ρ as a ratio $(\rho_{Red} - \rho_{SWIR1})/(\rho_{Red} + \rho_{SWIR1})$ to automatically delineate water boundaries, arguing that only water is more reflective in SWIR1 ρ than red ρ. Xu (Xu 2006) found that applying the NDWI misclassified noise of built-up land in urban areas, revealing that the reflectance pattern of built-up areas on green ρ and NIR ρ is similar to these of water. He also found that the built-up areas on SWIR1 ρ have much greater values than on NIR ρ, and suggested replacing this band by a modified normalized difference water index $MNDWI = (\rho_{Green} - \rho_{SWIR1})/(\rho_{Green} + \rho_{SWIR1})$. (Feyisa et al. 2014) proposed non-normalised multi-band ratios of LANDSAT data to delineate water areas. Their automated water extraction indexes (AWEI) specifically target areas with dark surfaces: $AWEI_{nsh} = 4 * (\rho_{Green} - \rho_{SWIR1}) - (0.25 * \rho_{NIR} + 2.75 * \rho_{SWIR2})$; in mountainous areas they used $AWEI_{sh} = \rho_{Blue} + 2.5 * \rho_{Green} - 1.5 * (\rho_{NIR} + \rho_{SWIR1}) - 0.25 * \rho_{SWIR2}$ to avoid confusion of deep shadows with water bodies. Another non-normalized water index (WI) proposed for large regions by Fisher, Flood, and Danaher (2016) uses the combination of five LANDSAT bands. The index is applied for surface reflectance data and evaluated as $WI = 1.7204 + 171\rho_{Green} + 3\rho_{Red} - 70\rho_{NIR} - 45\rho_{SWIR1} - 71\rho_{SWIR2}$.

2.2 Lake detection methods

Habib et al. (2006) suggested three methods for detecting lakes from SPOT images, i.e. the spectral angle mapper (SAM) classification method, the irregular pyramid, and the watershed with markers methods. They concluded that all of these methods had advantages and disadvantages, and to achieve the best results they suggested combining them. Verpoorter, Kutser, and Tranvik (2012) introduced a method for extracting water bodies from LANDSAT GeoCover^{TM} mosaics (GWEM), delineating lakes >0.0002 km² in size. They used thresholding and supervised classification to map water boundaries, and classified lake outlines by texture analysis. In a refinement step they converted lakes boundaries to vector format and evaluated their area, perimeter, and shape for evaluating additional morphometric parameters. In a last step they removed falsely classified shadows. Jiang et al. (2014) proposed an automated method for extracting rivers and lakes (AMERL) from LANDSAT images. They compared their method with the NDWI, MNWDI, and AWEI, and concluded that their method was superior, because it accounts for not only the spectral characteristics of pixels, but also their neighbourhood. Jawak and Luis (2014) suggested a semi-automatic approach for extracting lakes from World View-2 images based on a modified NDWI algorithm, and suggested to modify the NDWI and applying PAN-sharpening of images from 2 m into 0.5 m to improve classification.

3. TEST AREA AND DATA

3.1 Test area

We selected the greater part of the Tibetan Plateau as our test area (Fig. 1). We excluded from the analysis the eastern part of the Tibetan Plateau mainly because of the few high-quality cloud free LANDSAT images. We analysed all lakes on the selected satellite images, also these beyond the Tibetan Plateau margin. We tested our method for semi-automatically extracting lake outlines for most of the Tibetan Plateau where lakes formed in tectonic depressions or behind glaciers and sediments, and have different spectral colours, and different physical states of water (frozen and non-frozen).
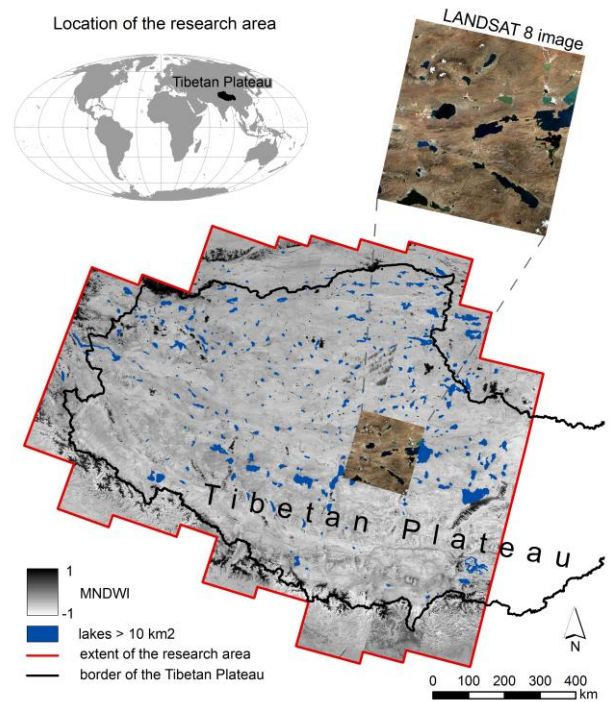

Figure 1. Study area on the Tibetan Plateau and modified normalized difference water index (MNDWI) for major lakes.

3.2 Data

We analysed 47 LANDSAT-5 images for 1995, and 47 LANDSAT-8 images for 2015. The size of the study area and the different weather conditions captured on these images required that we analysed different days of the year, i.e. from 25 April to 17 December 1995, and from 8 June to 22 November 2015. Due to the lack of high-quality images for the study area in 1995, we also included 22 images from 1994 and four images from 1996. We selected only images with <10% cloud cover, and atmospheric and sun-angle correction provided by the U.S. Geological Survey (http://espa.cr.usgs.gov/). We used Top-Of-Atmosphere (TOA) reflectance bands because the cosine effect of different solar zenith angles linked to different time in data acquisition is already removed from these data (Chander et al. 2009). To delineate the lake boundaries we used the SRTM DEM version 4 (Reuter, Nelson, and Jarvis 2007) (http://srtm.csi.cgiar.org/) as a supporting layer in the OBIA analysis. We used the SRTM DEM to generate a local slope map of the Tibetan Plateau. The slope map records the rate of elevation change between neighbouring pixels.

4. METHODS

Our workflow uses satellite images taken in all seasons, thus involving seasonally changing water colours and ice cover (Fig. 2). Consequently the RGB values of satellite images that are used for classifying water boundaries are not a proper input to delineate the lake shapes. Therefore, we applied an MNDWI algorithm for distinguishing water areas irrespective of the physical state of the water bodies and their visual appearance.
In this regard, we tested the modified normalized difference water index *MNDWI* (Xu 2006). In this algorithm only those pixels representing the water areas should have positive values. However, objects like glaciers, and mountains shadows have also values above the zero, making the classification of water areas in mountainous and ice-covered areas problematic. To

meet this challenge, we combined the MNDWI and its thresholds with OBIA and SRTM data for a more integrative approach for automatically detecting lakes.
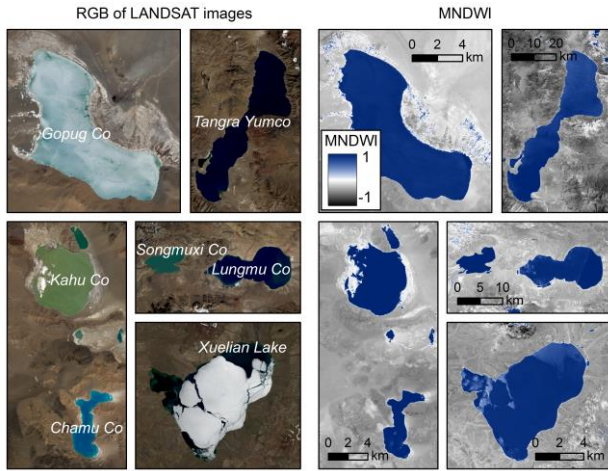

Figure 2. The RGB values of sample lakes on the Tibetan Plateau and their modified normalized difference water index (MNDWI).

In the first step we generated mosaics using standard method *Mosaic to New Raster* available in ArcGIS software for bands from all images for 1995 and 2015. Then we evaluated the MNDWI for the whole study area. We also evaluated the slope map [°] in a 3 x 3 search window from the SRTM DEM data using the *Slope* function with default options available in ArcGIS software. We then applied a multi-resolution segmentation algorithm (Baatz and Schäpe 2000) to segment the MNDWI values and classify the segments as either 'water' or 'other'. We applied the multi-resolution segmentation only to the MNWDI without the slope map, because the SRTM data date back to March 2000, and might thus influence the segmentation process. Using a manual trial and error approach we selected 100 as the scale of segmentation (Fig. 3). For this scale the segment size was smaller than lakes >10km² in area, and large enough to maintain feasible computing time. We arbitrarily selected 0.1 as the shape, and 0.7 as the compactness parameters, after performing tests with different scales.
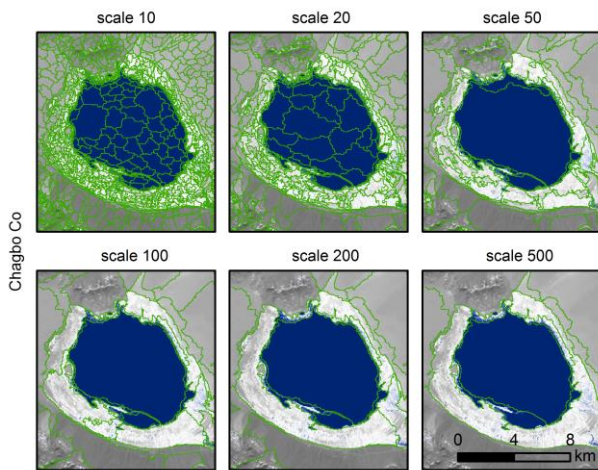

Figure 3. Multiresolution segmentation of Chagbo Co with different scales.

In the next step we used the threshold value of the MNDWI to classify water segments. To differentiate lakes from rivers we used the asymmetry of segments and their relation to the neighbouring segments and also their classification. The asymmetry we computed with the *Asymmetry* function in the eCognition software, where it is defined as a relative length of the segment to a regular polygon drawn around the segment. We find that many glaciers on the Tibetan Plateau have MNDWI values similar to those of lakes, and that a pure OBIA-driven classification based on MNDWI produced many errors, so that we used the slope map from the SRTM DEM as a supporting layer. We assumed that the glaciers are located on areas with slope >2° to distinguish them from lakes. To omit resolution errors, we used only those lakes >10 km² and exported them as vector polygons for further analysis. We used the exported lakes polygons to generate a reference data set, and verified visually the correctness of lakes boundaries using natural colour mosaic (RGB) generated from LANDSAT data. In some cases, especially in areas where the lakes are very shallow and visual interpretation of RGB image was difficult, we supported the delineation of lakes boundaries by analysing the MNDWI map. We edited the edges of all lakes in areas where the boundary was incorrectly delineated, and delated all falsely classified polygons not representing lakes.

5. RESULTS

Our classification of Tibetan lakes found that in 2015 the total area of lakes >10 km² was 31,258 km², or 2.6% of the total study area. Twenty years earlier, these same lakes covered only 24,892 km². The total area of large lakes thus grew by ~26% over two decades on average.

5.1 Classification accuracy

To verify the correctness of our classification we compared the results with manually generated reference data, and computed several standard performance statistics (Sithole and Vosselman 2004), (Congalton 1991), (Cohen 1960). Our OBIA method for extracting lakes with the area >10km² had an overall accuracy of 0.99 (Tab. 1); the producer's and user's accuracy as well as Cohen's kappa for both analysed time slices was >0.98.

Performance statistics	1995	2015
Type I error	0.0134	0.0169
Type II error	0.0003	0.0005
Total error	0.0006	0.0010
Overall accuracy	0.9994	0.9990
Producer's accuracy	0.9866	0.9831
Users accuracy	0.9850	0.9808
Cohen's kappa	0.9855	0.9815

Table 1. Performance statistics for automatic lakes classification

By cross-checking these classification results visually we see that the most of misclassified areas occur on the border of lakes (Fig. 4), especially in areas where the lakes has very complex coastlines or where rivers flows into the lake; it is there that our object-based approach tends to create only single segments. Clouds also favoured misclassification on some images, especially where located directly above the lake or its coasts. The physical state of the lake surface did not influence our automatic OBIA approach (Fig. 4 – Xuelian Lake), and both frozen and not-frozen parts of lake were correctly classified. We also observed that glaciers were distinguished from lakes with high accuracy, and only few glaciers were wrongly classified as lakes. The biggest confusion potential occurs where glacier segments are sharing a border with lake segments. In some

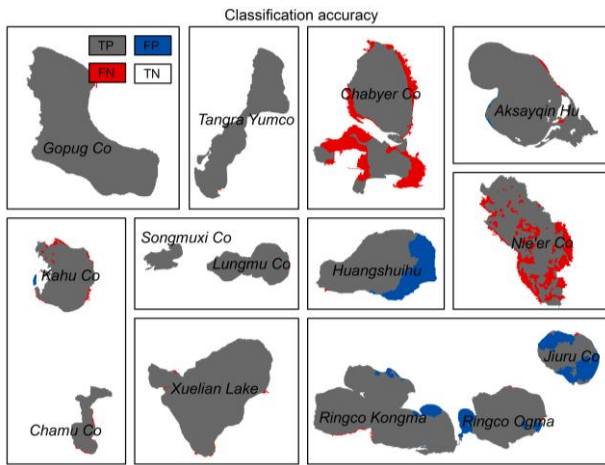cases small island lakes were falsely misclassified by adding them into the lake class.

Figure 4. OBIA classification accuracy examples for 2015 data based on confusion matrix; TP is true positive, FP is false positive, FN is false negative, and TN is true negative.

A good example of a rather low classification accuracy is Chabyer Co (Fig. 4); there, many pixels were falsely classified as 'lake', because this lake is changing its size rapidly, so that misclassified pixels represent marshy areas where shallow water occurred in the past. Nie'er Co is another example of a lake detected with low accuracy. Although this lake has been correctly detected by the thresholding of the MNDWI, the assumptions which we used in the OBIA approach included small lake islands and surrounding areas as lakes.

5.2 Lakes changes

Our analysis shows that in general the total area of lakes on the Tibetan Plateau increased about 0.5% or some 6,366 km². Changes in the lakes area are not evenly spread out on the whole area of the Tibetan Plateau (Fig. 5).
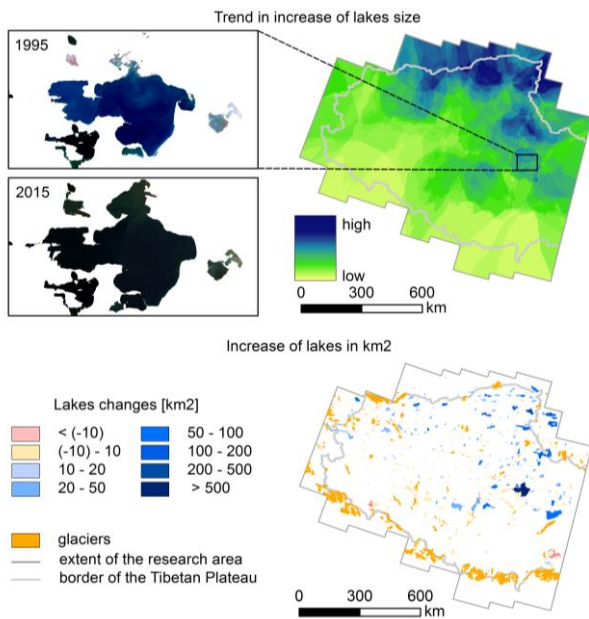

Figure 5 Changes in lake areas on the Tibetan Plateau from 1995 to 2015, compared to glacier area.

Lakes were growing most strikingly in the north-eastern parts, which also hosts the largest number of lakes with individual net increases of >100 km². Most lakes with size changes of <10 km² in the past two decades are in the southwestern part of the Tibetan Plateau, mainly along the Himalayas and adjacent mountain belts.

6. DISCUSSION

Our automatic classification results have a strikingly high accuracy. Most lakes have been correctly detected with only few misclassifications along lake boundaries, especially complex coastline geometries. We found the MNDWI an accurate indicator of all water areas, regardless of whether they sustained ice cover or not. The segmentation algorithm in our workflow helped to reduce the salt-and-pepper noise and delineate more reliably the boundaries of lakes, whereas OBIA allowed finding the segments which were incorrectly classified by the MNDWI thresholding. Spuriously low values for thresholding arose from clouds above the lakes, too low MNDWI values, or mountain shadows. Integrating topography via a slope map enabled us to exclude from the lakes class falsely classified glaciers. We infer that using more accurate DEMs might help to achieve better results, especially if the topographic data were captured closer to the timeframe of interest. We used DEM obtained in 2000 to analyse lakes in 1995 and 2015, meaning that we had 10 to 15 years difference between our elevation data and the satellite images. Any OBIA workflow should be corrected for such possible mismatches, wherever possible.

Our change detection analysis reveals that changes in lake size are not uniform across the Tibetan Plateau, thus supporting similar results from previous works (Song, Huang, and Ke 2013), (Song et al. 2014), and (Fang et al. 2016). Comparing our findings to a regional glacier map, we see that the most of glaciers are located in the Himalayan Mountains Range where the growth of large lakes is smallest, hence any excess glacial meltwater might mostly feed lakes <10 km² in that area. It is equally possible that any growth of lakes in the Himalayas remains shrouded from automated detection because of the locally rough terrain characteristics. Changes to lakes in flatter areas involve commensurate larger increases in area than in mountainous areas. Such topographic boundary conditions may partly explain the large net growth of lake areas in the north-eastern Tibetan Plateau, although detailed analyses are clearly desirable to unravel the underlying causes.

7. CONCLUSIONS

In this contribution we presented an approach for automatic detection of lakes >10 km², and their changes through time, drawing on examples from the Tibetan Plateau. Our method is designed for detecting lakes with different state of the water for large areas, especially those where many glaciers occur. Our approach combined a satellite-derived water index, OBIA, and DEM-derived slope data to automatically extract lakes. Our classification derived 323 lakes with a total area of 31,258 km², or 2.6% of the study area in 2015. The same number of lakes had covered only 24,892 km² in 1995, so that lakes grew by ~26% in the past two decades on average. We conclude that our approach is general and flexible enough to provide easy transferability and its application in areas other than the Tibetan Plateau, which we selected as our testing research area.

REFERENCES

Baatz, Martin, and Arno Schäpe. 2000. "Multiresolution Segmentation: An Optimization Approach for High Quality Multi-Scale Image Segmentation." *Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium Salzburg 2000, Karlsruhe, Herbert Wichmann Verlag*: 12–23.

Chander, Gyanesh, Brian L Markham, and Dennis L Helder. 2009. "Remote Sensing of Environment Summary of Current Radiometric Calibration Coef Fi Cients for Landsat MSS , TM , ETM + , and EO-1 ALI Sensors." *Remote Sensing of Environment* 113(5): 893–903.

Cohen, Jacob. 1960. "A Coefficient of Agreement for Nominal Scales.", *Educational and Psychological Measurement* 20(1): 37–46.

Congalton, Russell G. 1991. "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." *Remote Sensing of Environment* 37(1): 35–46.

Fang, Yue, Weiming Cheng, Yichi Zhang, Nan Wang, Shangmin Zhao, Chenghu Zhuo, Xi Chen, and Anming Bao. 2016. "Changes in Inland Lakes on the Tibetan Plateau over the Past 40 Years." *Journal of Geographical Sciences* 26(4): 415–38.

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, Neil Flood, and Tim Danaher. 2016. "Comparing Landsat Water Index Methods for Automated Water Classification in Eastern Australia." *Remote Sensing of Environment* 175: 167–82.

Frazier, Paul Shane, and Kenneth John Page. 2000. "Water Body Detection and Delineation with Landsat TM Data." *Photogrammetric Engineering & Remote Sensing* 66(12): 1461–67.

Habib, Tarek, Michel Gay, Jocelyn Chanussot, and Pascal Bertolino. 2006. "Segmentation of High Resolution Satellite Images SPOT Applied to Lake Detection." *International Geoscience and Remote Sensing Symposium (IGARSS)* (April): 3680–83.

Jawak, Shridhar D, Kamana Kulkarni, and Alvarinho J Luis. 2015. "A Review on Extraction of Lakes from Remotely Sensed Optical Satellite Data with a Special Focus on Cryospheric Lakes." *J* 4(4): 196–213.

Jawak, Shridhar D, and Alvarinho J Luis. 2014. "A Semiautomatic Extraction of Antarctic Lake Features Using Worldview-2 Imagery." (October): 33–46.

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(6): 5067–89.

Liu, Xiaodong, and Baode Chen. 2000. "Climatic Warming in the Tibetan Plateau During Recent Decades." *International Journal of Climatology* 20: 1729–42.

McFeeters, S. K. 1996. "The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features." *International Journal of Remote Sensing* 17(7): 1425–32.

Reuter, H. I., A. Nelson, and A. Jarvis. 2007. 21 International Journal of Geographical Information Science *An Evaluation of Void-Filling Interpolation Methods for SRTM Data*.

Rogers, A. S., and M. S. Kearney. 2004. "Reducing Signature Variability in Unmixing Coastal Marsh Thematic Mapper Scenes Using Spectral Indices." *International Journal of Remote Sensing* 25(12): 2317–35.

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(3): 442–56.

Sithole, George, and George Vosselman. 2004. "Experimental Comparison of Filter Algorithms for Bare-Earth Extraction from Airborne Laser Scanning Point Clouds." *ISPRS Journal of Photogrammetry and Remote Sensing* 59(1-2): 85–101.

Song, Chunqiao, Bo Huang, and Linghong Ke. 2013. "Modeling and Analysis of Lake Water Storage Changes on the Tibetan Plateau Using Multi-Mission Satellite Data." *Remote Sensing of Environment* 135: 25–35.

Song, Chunqiao, Bo Huang, Linghong Ke, and Keith S. Richards. 2014. "Remote Sensing of Alpine Lake Water Environment Changes on the Tibetan Plateau and Surroundings: A Review." *ISPRS Journal of Photogrammetry and Remote Sensing* 92: 26–37.

Verpoorter, Charles, Tiit Kutser, and Lars J. Tranvik. 2012. "Automated Mapping of Water Bodies Using Landsat Multispectral Data." *Limnology and Oceanography: Methods* 10: 1037–50.

Wang, Cuizhen. 2016. "A Remote Sensing Perspective of Alpine Grasslands on the Tibetan Plateau: Better or Worse under 'Tibet Warming'?" *Remote Sensing Applications: Society and Environment* 3: 36–44.

Xu, Hanqiu. 2006. "Modification of Normalised Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery." *International Journal of Remote Sensing* 27(14): 3025–33.

Yang, Wei, Tan Dong Yao, Bai Qing Xu, Guang Jian Wu, Ling Long Ma, and Xiao Dong Xin. 2008. "Quick Ice Mass Loss and Abrupt Retreat of the Maritime Glaciers in the Kangri Karpo Mountains, Southeast Tibetan Plateau." *Chinese Science Bulletin* 53(16): 2547–51.

Zhang, Guoqing et al. 2011. "Monitoring Lake Level Changes on the Tibetan Plateau Using ICESat Altimetry Data (2003-2009)." *Remote Sensing of Environment* 115(7): 1733–42.