Verification and analysis of surface water in China based on Landsat8 OLI images

Wei Jiang1,2,3,4, Guojin He4,5, Zhiguo Pang1,2,3, Tengfei Long4,5, JUNE Fu1,2,3, Hongxiang Guo4,6 and Yuan Ni7

1State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, Beijing 100038, China
2China Institute of Water Resources and Hydropower Research, Beijing, China;
3Research Center of Flood and Drought Disaster Reduction of the Ministry of Water Resources, Beijing, China;
4Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China;
5Key Laboratory of Earth Observation Hainan Province, Hainan, China;
6University of Chinese Academy of Sciences, Beijing, China;
7Sichuan Highway Planning, survey, Design and Research Institute Ltd, Sichuan, China;

E-mail: jiangwei@iwhr.com

Abstract. Surface water map is important to understand climate change and the hydrologic cycle and manage water resource. This study introduces multilayer perceptron (MLP) neural network to produce a surface water map of China in 2015 (SWMC-2015), Then, a new framework is created to generate a set of random water and not water validation points to assess the accuracy of the SWMC-2015. The overall accuracy and kappa coefficients of SWMC-2015 are 90% and 0.78, respectively. Finally, with the SWMC-2015, the spatial characteristics of surface water are comprehensively analysed at national, basin and province scales. The surface water area in China is 154,811.1 km², accounting for 1.63% of the total land area. There are considerable differences in surface water abundance at both the basin and province scales. The water body areas are largest in the continental basin and Tibet, and water bodies cover the greatest percentage of total surface area in the Huaihe River Basin and Jiangsu.

1. Introduction

Water plays an important role in human development and natural ecosystems. Lakes, rivers and reservoirs and excluding oceans and wetlands, are defined as surface water [1]. Remote sensing technology provides an effective way to monitor surface water [2, 3]. Before 2013, Moderate Resolution Imaging Spectroradiometer (MODIS) was the main data source used in surface water mapping at the global scale [4, 5]. With the open sharing of Landsat series data, which offer the
longest continuous moderate-resolution data since 1972[6], the timeliness and accuracy of surface water mapping have been greatly improved[3, 7].

In recent years, eight global surface water products using Landsat series satellite data have been released. Landsat Thematic Mapper (TM)/ Enhanced Thematic Mapper (ETM+) are the main data sources used to conduct surface water mapping. Since the launch of Landsat 8 in 2013, the Landsat8 OLI has been used to produce long time series surface water products[8]. A decision tree with supervised classification is the primary classification method. Threshold and object-based methods are also employed to generate surface water products[9].

Until now, most products utilized decision tree methods [3, 10]. Recently, a new and effective surface water extraction method, the multilayer perceptron (MLP) neural network, was proposed to identify surface water in Landsat 8 OLI images[8]. The paper introduces MLP neural network to produce a surface water map of China (SWMC-2015), and then a frameworks of water and not water validation points was proposed to assess the accuracy. Finally, the spatial pattern of the surface water based on SWMC-2015 was analysed at different scales.

2. Methods

2.1. Data and preprocessing

A total of 516 Landsat 8 images were acquired from the United States Geological Survey (USGS). The image quality and seasonal surface water dynamics were considered. Cloud-free images from May to October in 2015 were preferred, and the remaining images were supplemented with images from the closest month. To ensure consistent experimental images, the composite bands 1-7 images should undergo radiometric calibration to generate the top of atmosphere reflectance product.

2.2. Surface water extraction method

The MLP neural network has been confirmed to be an excellent surface water algorithm for extracting various water body types and suppressing the noise caused by shadows and ice/snow[8]. The flow of the extraction included sample selection, model training and classification. To ensure the accuracy of the classification, all training samples covering various water body types (lake, river, sea water, open pond, turbid water and aquacultural water) and not water body types (land, ice/snow, cloud and cloud shadow) were manually labelled in the ENVI 5.3 software platform based on TOA reflectance. Based on the training samples for each scene, the TOA reflectance values of seven bands in the labelled pixels were input to optimize the weights and minimize the errors of the MLP model[8]. Then, the probability of water and not water types for each pixel were automatically calculated to generate classification results.

2.3. The framework of validation points generation

To conduct the quantitative accuracy assessment of SWMC-2015, a comprehensive accuracy assessment flow is proposed, which is shown in Figure 1. The permanent water product (G3WBM) is employed as reference data to generate the water extent, and not water extent, and a 200 m buffer zone is set based on experience and experiment. The uniformity and spatial representation of the random points are considered, and a grid of $0.35^\circ \times 0.35^\circ$ is created using land vectors in China. Then, the three extents are connected with the grid. Two random points are generated for each grid cell of the water extent-based grid. Moreover, only one random point is generated in each grid cell of the not water extent. In order to eliminate the effect of the reference data, these random points are verified using Google Earth high-resolution satellite remote sensing images. Finally, the error matrix is created using the validation points and the SWMC-2015 product. The overall accuracy and kappa coefficients are calculated to quantitatively assess the accuracy.
Figure 1. The flow of validation points generation.

Based on the flow of validation point generation, the distribution of random validation points is shown in Figure 2(a). These points are verified using Google Earth imagery, which indicates that there are 10,419 random water validation points and 5332 random not water validation points.

Figure 2. The distribution of random points and random validation points in China ((a) the random validation points, and (b) the distribution of water and not water incorrectly points).

3. Results and discussion

3.1. Classification Results and Accuracy Assessment

The SWMC-2015 is generated using the MLP model, and the results are shown in Figure 3. The major lake clusters and river networks are clearly shown in the SWMC-2015. To quantitatively assess the accuracy of the SWMC-2015, the random points of the water and not water are validated. There are 1347 water error points and 263 not water error points (figure 2b). The overall accuracy and kappa coefficient are 90% and 0.78, respectively. This result suggests that the SWMC-2015 performs better when identifying detailed surface water bodies.
Figure 3. The performance of SWMC-2015. (a) the classification result of SWMC-2015, (b), (c), (d), (e) are the extraction performance of river, lake, aquacultural water and sea water, respectively.

3.2. Spatial Characteristics Analysis of SWMC-2015

The SWMC-2015 was used to analyse the spatial characteristics of China's surface waters at national, basin and province scales. The surface water area in China is 154,811.1 km², accounting for 1.63% of the land area, which is less than the officially declared water coverage (2.8%). This difference occurs because the definition of surface water used in this study refers to only inland water, excluding the sea and offshore areas.

Figure 4. The water body area and proportion of water body area at the provincial scale

The spatial characteristics of the SWMC-2015 in 32 provincial administrative regions (excluding Hong Kong and Macao) are analysed, and the result for each province is shown in Figure 4. The three provinces with the highest water body areas are Tibet (31,950.36 km²), Qinghai (19,244.38 km²) and Jiangsu (10,193.45 km²), while the lowest three are Beijing (174.30 km²), Shanghai (345.20 km²) and Ningxia (345.20 km²). These results suggest that water bodies are unevenly distributed across the provinces. The three provinces with the highest proportions of water body areas are Jiangsu (10.43%), Tianjin (9.87%) and Shanghai (5.61%), while the lowest three are Shanxi (0.48%), Shaanxi (0.33%) and Gansu (0.17%). These results show that there is a considerable difference in surface water abundance between coastal and inland provinces.

The water body area and proportion of water body area at the basin scale are shown in Figure 5. The water body area of the continental basin is the largest (49,964.23 km²), followed by the Yangtze River Basin (41,490.21 km²). However, the water body area of the Haihe River Basin is the smallest. The three regions with the highest proportions of water body area are the Huaihe River Basin (3.61%), Southeast Basin (2.21%) and Yangtze River Basin (2.31%), and the three areas with the lowest proportions are the Southwest Basin (0.73%), Yellow River Basin (0.88%) and Haihe River Basin (1.31%).
4. Conclusions

This study introduces an MLP neural network to develop the SWMC-2015 based on Landsat 8 OLI images. Then, validation points is created to assess the accuracy of the SWMC-2015 and the spatial characteristics of surface water in China are analysed at different scales. The conclusions are summarized as follows:

(1) Based on the random validation points, the result shows that the accuracy of the SWMC-2015 is high (overall accuracy is 90%, and kappa coefficient is 0.78). The SWMC-2015 can clearly show the major lake clusters and river networks.

(2) The surface water area in China is 154,811.1 km², accounting for 1.63% of the total land area. At the province scale, the water body area of Tibet is the largest, and the largest proportion of water body area is in Jiangsu. At the basin scale, the water body area of the continental basin is the largest, and the largest proportion of water area is in the Huaihe River Basin.

References

[1] Liao A P, Chen L J, Chen J, He C Y, Cao X, Chen J, Peng S, Sun F D and Gong P 2014 High-resolution remote sensing mapping of global land water Science China-Earth Sciences 57 2305-16

[2] Mueller N, Lewis A, Roberts D, Ring S, Melrose R, Sixsmith J, Lymburner L, McIntyre A, Tan P, Curnow S and Ip A 2016 Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia Remote Sens. Environ. 174 341-52

[3] Pekel J F, Cottam A, Gorelick N and Belward A S 2016 High-resolution mapping of global surface water and its long-term changes Nature 540 418+-

[4] Carroll M L, Townshend J R, DiMiceli C M, Noojipady P and Sohlberg R A 2009 A new global raster water mask at 250 m resolution Int. J. Digit. Earth. 2 291-308

[5] Feng L, Hu C M, Chen X L, Cai X B, Tian L Q and Gan W X 2012 Assessment of inundation changes of Poyang Lake using MODIS observations between 2000 and 2010 Remote Sens. Environ. 121 80-92

[6] He G, Zhang Z, Jiao W, Long T, Peng Y, Wang G, Yin R, Wang W, Zhang X, Liu H, Cheng B and Xiang B 2018 Generation of ready to use (RTU) products over China based on Landsat series data Big Earth Data 2 56-64
[7] Donchyts G, Baart F, Winsemius H, Gorelick N, Kwadijk J and van de Giesen N 2016 Earth's surface water change over the past 30 years Nat. Clim. Chang. 6 810-3

[8] Jiang W, He G, Long T, Ni Y, Liu H, Peng Y, Lv K and Wang G 2018 Multilayer Perceptron Neural Network for Surface Water Extraction in Landsat 8 OLI Satellite Images Remote Sens. 10 755

[9] Huang C, Chen Y, Zhang S and Wu J 2018 Detecting, extracting and monitoring surface water from space using optical sensors — a review Reviews of Geophysics

[10] Feng M, Sexton J O, Channan S and Townshend J R 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. 9 113-33

Acknowledgments
This research was financially supported by the National Natural Science Foundation of China (51779269) and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA190090300).