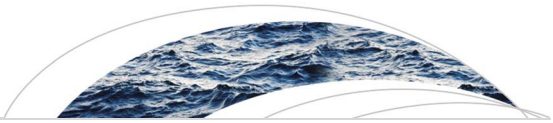
Check for updates

AGU100 ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research



RESEARCH ARTICLE
10.1029/2018WR023060

Special Section:
Big Data & Machine Learning in Water Sciences: Recent Progress and Their Use in Advancing Science

Key Points:
- Global inland water has a dramatic seasonal variation
- Short duration water bodies, sea level rise effects, and various types of rice field use can be detected

Correspondence to:
P. Gong,
penggong@tsinghua.edu.cn

Citation:
Ji, L., Gong, P., Wang, J., Shi, J., & Zhu, Z. (2018). Construction of the 500-m resolution daily global surface water change database (2001–2016). *Water Resources Research*, *54*, 10,270–10,292. https://doi.org/10.1029/2018WR023060

Received 2 APR 2018
Accepted 19 NOV 2018
Accepted article online 26 NOV 2018
Published online 26 DEC 2018

©2018. The Authors.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Construction of the 500-m Resolution Daily Global Surface Water Change Database (2001–2016)

Luyan Ji[1], Peng Gong[1,2], Jie Wang[2,3], Jiancheng Shi[3], and Zhiliang Zhu[4]

[1]Ministry of Education Key Laboratory of Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China, [2]AI for Earth Laboratory, Tsinghua University, Beijing, China, [3]State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China, [4]United State Geological Survey, Reston, VA, USA

Abstract Surface water is the most dynamic land-cover type. Transitions between water and nonwater types (such as vegetation and ice) can happen momentarily. More frequent mapping is necessary to study the changing patterns of water. However, monitoring of long-term global water changes at high spatial resolution and in high temporal frequency is challenging. Here we report the generation of a daily global water map data set at 500-m resolution from 2001 to 2016 based on the daily reflectance time series from Moderate Resolution Imaging Spectroradiometer. Each single-date image is classified into three types: water, ice/snow, and land. Following temporal consistency check and spatial-temporal interpolation for missing data, we conducted a series of validation of the water data set. The producer's accuracy and user's accuracy are 94.61% and 93.57%, respectively, when checked with classification results derived from 30-m resolution Landsat images. Both the producer's accuracy and user's accuracy reached better than 90% when compared with manually interpreted large-sized sample units (≥1,000 m × 1,000 m) collected in a previous global land cover mapping project. Generally, the global inland water area reaches its maximum (~3.80×10^6 km^2) in September and minimum (~1.50×10^6 km^2) in February during an annual cycle. Short-duration water bodies, sea level rise effects, different types of rice field use can be detected from the daily water maps. The size distribution of global water bodies is also discussed from the perspective of the number of water bodies and the corresponding water area. In addition, the daily water maps can precisely reflect water freezing and help correct water areas with inconsistent cloud flags in the MOD09GA quality assessment layer.

Plain Language Summary Daily global inland surface water maps are produced from more than 1.9 million frames of satellite images for the period of 2001–2016 with a spatial resolution of 500 m. From this time series of maps, we found that the inland surface water on Earth varies greatly in area within an annual cycle. It can reach more than 3.8 million square kilometers in September and reduce to approximately 1.5 million square kilometers in February. We demonstrate that (1) short-duration water bodies in arid areas, which are particularly important to life, can be detected from these daily water maps; (2) sea level rise effects on land submersion can be detected over some gentle-slope coasts like west Florida; and (3) different types of rice field use exist in the world. For example, in California, United States, rice fields are filled with water after harvest to help create a wetland environment for wild birds in the winter.

1. Introduction

Abundance and distribution of surface water are direct indicators of availability and accessibility of water in a region. Surface water is an essential variable in ecological, hydrological, climate, and even carbon cycling and biodiversity studies (Downing et al., 2006; Jung et al., 2010; Pereira et al., 2013; Raymond et al., 2013; Schuur et al., 2015; Seekell et al., 2014). Since water is a highly variable surface parameter on Earth, it is particularly important to have an accurate, frequent, and long-term record of its spatial-temporal distribution over large areas (Sun et al., 2014). For applications like land surface modeling and water management, it is desirable to have daily or more frequent water information (Howells et al., 2013; Pfister et al., 2011; Wada et al., 2014). To have such information over large areas, satellite remote sensing is the only possible technology (Gong et al., 2013).

At the global scale, although satellite remote sensing has the global coverage in the 1970s, it was not until the 1990s that global water data with kilometer-scale resolution derived from the NOAA weather satellite data became available in general land cover classification maps (e.g., Hansen et al., 2000; Loveland et al., 2000).

Water surface is always a land cover category in general purpose land cover maps in 500-, 300-, 250-, and 30-m resolution products (e.g., Arino et al., 2008; Friedl et al., 2002; Gong et al., 2013; Wang et al., 2015). However, general purpose land cover maps are primarily produced from optical satellite data acquired at best in annual scales that are subject to cloud contamination and topographic effects, and due to the limitation of classification-based algorithms, accuracies of the water category from land cover map products are not optimized (Ji et al., 2015). Therefore, specialized water extraction algorithms have been developed to take advantages of more frequent optical remotely sensed data, more weather robust microwave data, or the integration of optical and microwave data. A 250-m resolution global land water mask was developed by integrating 90-m resolution Shuttle Radar Topographic Mission (SRTM) data and Moderate Resolution Imaging Spectrometer (MODIS) data (Carroll et al., 2009). Because SRTM was acquired in 2000, this water mask used MODIS data acquired around 2000 and 2001, so it can be regarded as a global water mask for 2000. Coarse resolution (~25 km) water masks have been produced from microwave radiometer data at annual (Papa et al., 2010) and monthly (Prigent et al., 2012) basis covering a time span of 1993–2007. Ancillary topographic and hydrological data were used to downscale those coarser resolution data to 500-m resolution annual global inundation data (Fluet-Chouinard et al., 2015). With intensive manual editing, a 30-m resolution global water mask was produced using single-date circa 2010 Landsat Thematic Mapper (TM) data (Liao et al., 2014). More automated data products on global lake inventory and global water mask were generated from circa 2000 Landsat TM data (Feng et al., 2016; Verpoorter et al., 2014). By using multitemporal Landsat images, a global 3 arc-second water body map, which can separate permanent water bodies from the temporal water-covered areas, was conducted for four Global Land Survey epochs (Dai et al., 2015). A 30-m resolution monthly global water surface product was produced using classification methods in Google Earth Engine based on all Landsat TM data during 1984–2015 (Pekel et al., 2016).

From the above review, it can be seen that creation of daily water surface data at the global scale is still a challenge. The reasons include the restrictions of data availability due to the limited spatial coverage and low temporal frequency of higher spatial resolution data (finer than 100 m) or the lack of spatial detail due to low spatial resolution (poorer than 10 km) of more frequently available data. It seems that despite of the cloud and snow/ice effects over various geographical locations on Earth, the freely accessible, daily repeat-cycle MODIS data (two bands of 250 m with five more bands of 500-m resolution optical data) from the United States and the Medium Resolution Imaging Spectrometer data (300-m full resolution) from Europe are the only choices for producing daily surface cover products at the subkilometer level. Because Medium Resolution Imaging Spectrometer data were only available from 2002 to 2012, MODIS became the only alternative with a continuing capability of data acquisition since 2000. Klein et al. (2015) developed a technique to map daily global surface water using 250-m MODIS data, but they only tested it for the year of 2013 and 2014 (Klein et al., 2015, 2017). However, the 250-m MODIS data are acquired from only two spectral bands in visible (VIS) and near infrared. The unique niche of 500-m resolution MODIS data, which not only have spectral coverage expanding into the short wave infrared (SWIR) spectral region but also are continuously available on a daily basis covering nearly two decades of the Earth surface, makes it worthwhile to continuously generate daily global water cover data for both the water resource management and the scientific community at large in various other Earth system studies.

The purpose of this research is to develop and evaluate an efficient algorithm for producing global surface water cover and apply it to 500-m resolution MODIS data and ancillary topographic and land surface temperature (LST) data to create an archive of daily global water surface data product from 2001 to 2016. The cloud and ice effects have been reduced as much as possible. In the rest of this paper, we present the algorithms, the validation, and some characteristics of the produced daily global water maps through some preliminary analysis.

2. Data

2.1. A MODIS Daily Reflectance Time Series

The 500-m MODIS surface reflectance product, MOD09GA (version 6, https://lpdaac.usgs.gov/) from 2001 to 2016, is selected for daily global water mapping. MODIS data are organized in 10° × 10° image tiles. There are more than 1.9 million titles of images during the period of 2001 and 2016. For each tile, it provides seven reflectance bands from the VIS to SWIR spectral ranges in a daily gridded level-2 product in Sinusoidal

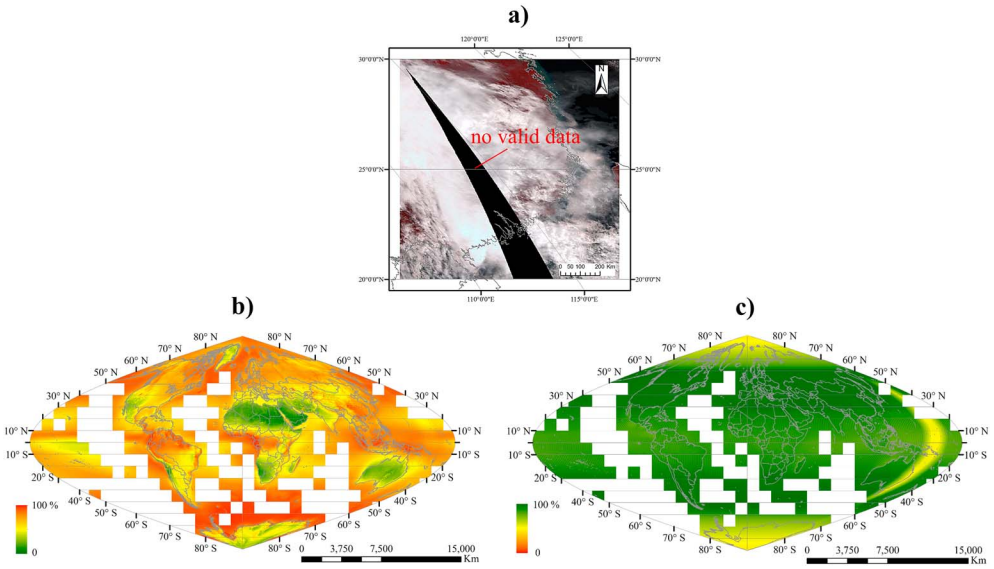


Figure 1. Illustration of pixels without valid data (a, R: Band 4, G: Band 3, B: Band 2, hv = "h28v06", day of year [DOY] = 096, year = 2014). The average percentage of cloud cover (b) and valid data (c) from 2001 to 2016. The image stretching mode is *Minimum-Maximum*.

projection. Besides, it also includes a 500-m State quality assessment (QA) layer, which describes the data quality of the seven reflectance bands, and a 1-km State QA layer to indicate whether a pixel contains cloud, cloud shadow, snow, aerosol, and so forth.

Two major data problems exist in this product: cloud cover and the no valid data (NVD), as shown in Figure 1a. The average occurrence of cloud and valid data for these 16 years are given in Figures 1b and 1c. It can be seen that the tropics and the east side of the Tibet Plateau suffer from severely high cloud cover throughout a year, while the high latitude area is subject to the lack of valid data.

2.2. Auxiliary Data

Three different types of auxiliary data sets are used. (1) Topographic information is derived from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) 30-m elevation data and SRTM 90-m elevation data. The ASTER Digital Elevation Model (DEM) data are used as supplementary data in slope calculation for areas from 60°N to 80°N where SRTM DEM does not cover. (2) MODIS daily LST time series. The LST is important for differentiating between water and low-reflectance snow-cover or ice area. The MODIS LST daily product, MOD11A1 (version 6, https://lpdaac.usgs.gov) at 1-km spatial resolution, is utilized as an additional information for water and snow/ice classification. (3) Improved FROM-GLC water layer. Finer Resolution Observation and Monitoring of Global Land Cover (FROM-GLC) is the first 30-m resolution global land cover product. Its water mask has been refined mainly by eliminating the misclassification of cloud and mountain shadows (Ji et al., 2015). This product has been used as a reference map, especially in preserving small-sized water bodies.

2.3. Validation Data

Two kinds of validation samples, FROM-GLC training and test samples and the higher resolution Landsat imagery, were used in this study.

(1) A total of 91,433 training and 38,664 test sample units was produced for FROM-GLC circa 2010 (Zhao et al., 2014). The locations of training samples were manually determined by interpreters traversing all Landsat images used in producing FROM-GLC 2010, while the locations of test samples were fixed by preset in a systematic unaligned manner. Both training and test samples are globally distributed, except for Arctic and Antarctic regions (Gong et al., 2013). Five years later, a multiseason sample set for FROM-GLC circa 2015 was developed, which had the same sample location as in the 2010 version but contain dynamic land cover types for four seasons (C. Li et al., 2017). The attributes for each sample point include location, date, sample size, and land cover type. The date of sample units in the FROM-GLC 2010 product ranges from 13 September

Table 1
The Number of Sample Units in the Four FROM-GLC Sample Sets for All Land Cover Types (Except Cloud) and Water

FROM-GLC sample set	All types	Water	Percent
FROM-GLC training, 2015, multiseason	305,877	33,345	10.90
FROM-GLC test, 2015, multiseason	128,584	4,040	3.14
FROM-GLC training, 2010	73,839	9,979	13.51
FROM-GLC test, 2010	31,827	1,204	3.78
Total	540,127	48,568	8.99

1984 to 09 August 2011, and 69.39% of the samples were between 2009 and 2010, while the FROM-GLC 2015 sample units range from 26 February 1985 to 06 July 2015, and 99.85% of the sample units are between 2014 and 2015. Considering the consistence in time and exclusion of cloud cover, the sample used in this study was picked from different FROM-GLC sample sets (Table 1). It can be observed that the percentage of water for test samples (i.e., 3.14% for 2015 and 3.78% for 2010) is close to the actual percentage of the inland water on Earth ($\approx 380 \times 10^4$ km^2/149×10^6 km^2 = 2.55%). The percentage of 2015 is lower because the there is more inland water on Earth in 2010 than 2015 (referring to Figure 13).

(2) Seven hundred twenty globally distributed Landsat-8 images were selected for further validation (Figure 2). The selection criteria for these Landsat images are

1. The Landsat image should contain water pixels;
2. For the convenience, each Landsat image should fall inside one MODIS tile; and
3. Landsat images that are partially or totally falling into those NVD gaps in the MODIS data were not chosen.

The percentage of valid Landsat-8 images that can be used for validation is shown as the background in Figure 2. Finally, time series of 55 validation sites (red dots in Figure 2) consist of a total of 720 Landsat tiles, including 196 in spring, 203 in summer, 176 in autumn, and 145 in winter, the cloud covers of which are all less than 50%. Moreover, Landsat images of saline lakes, small-size ponds and rivers, frozen lakes, and flooded wetland were deliberately chosen for closer assessment of the accuracy of our MODIS water product in those typically difficult areas.

3. Method

Global water time series are produced through an automated, three-stage procedure (Figure 3). First, water is identified on each single-date MODIS reflectance image by a rule-based algorithm. Then, an object-based postclassification method is applied to each water classification result to remove the misclassification from shadow (including cloud and mountain shadows). At this stage, prior knowledge about global water distribution and DEM information is used. Finally, cloud and NVD areas are removed using time series processing methods.

3.1. Preprocessing for Cloud, Ice/Snow, and NVD

The 1-km MOD09GA State QA layer contains both cloud and ice/snow flags that can be applied to the 500-m reflectance data. For cloud, it contains two flags: *cloud state* (bit 0-1; 00: clear, 01: cloudy, 10: mixed, and 11: not set, assumed clear) and *internal cloud algorithm flag* (bit 10; 1: cloud and 0: no cloud). We combined the

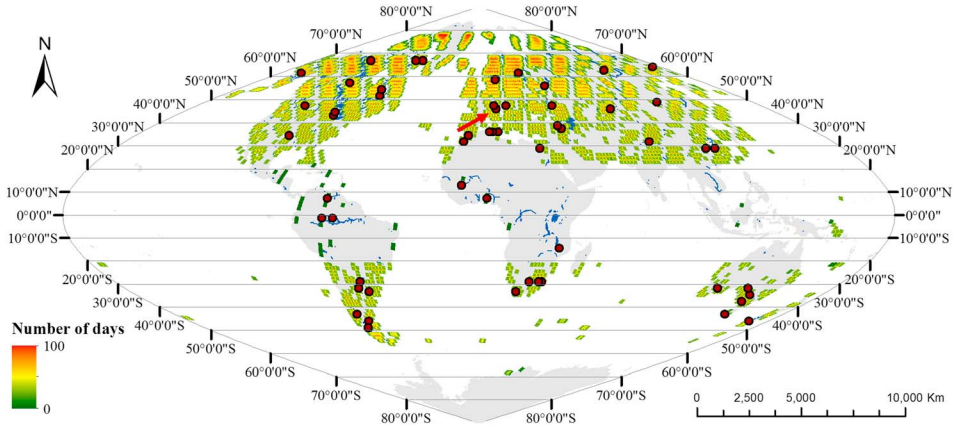


Figure 2. The Landsat image-based validation sites. The background is the percentage of the valid Landsat images for validation. An example where the arrow refers is found in Figure 9.

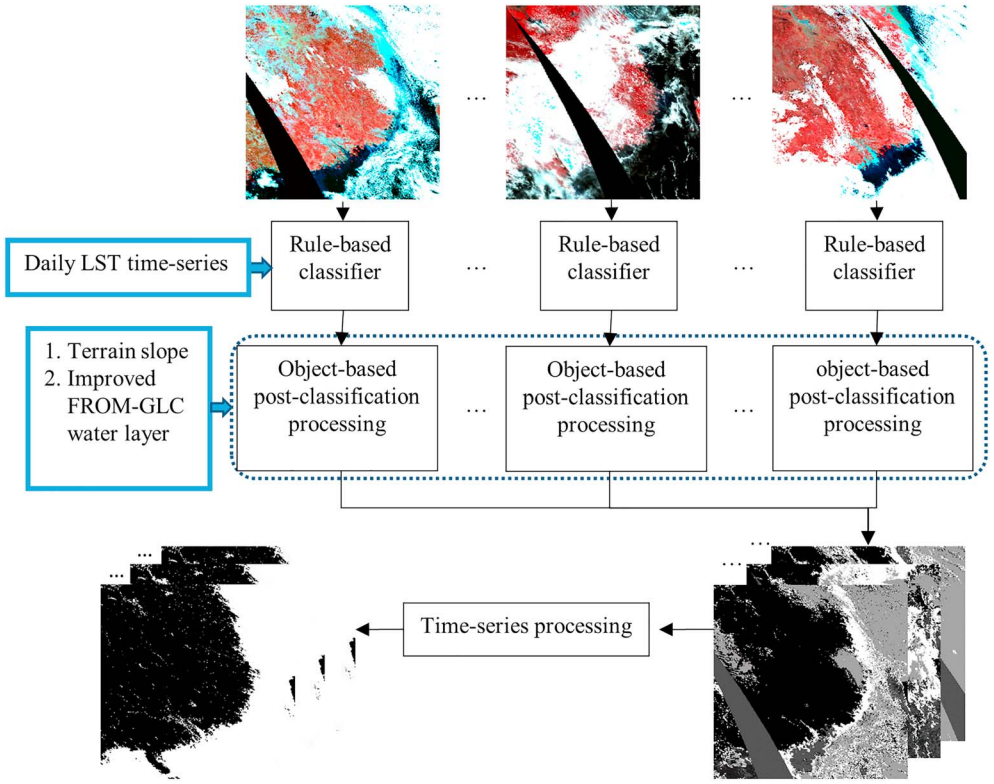


Figure 3. The flow chart of MODIS daily reflectance time series water mapping. LST = land surface temperature.

two flags to form a new cloud flag that contains as describe in Figure 4a, where the *uncertain* flag means the inconsistent result for cloud in MODGA State QA layer. These will be subsequently processed.

Similarly, the State QA layer also have two ice/snow flags: *MOD35 snow flag* (bit 12; 1: yes and 0: no) and *internal snow mask* (bit 15; snow, 0: no snow). We also combined the two flags to form a new ice/snow flag as shown in Figure 4b, where the uncertain flag shows the inconsistent result for ice/snow in MODGA State QA layer. For those inconsistent cloud or ice/snow pixels, they will be reclassified in the following single-tile water classification process. In the case that a pixel is simultaneously labeled as cloud or ice/snow, ice/snow has a higher priority.

However, overclassification of ice/snow is found globally, even in tropical area (Figure 5). To overcome this problem, the strategy is to use the surface temperature. Therefore, the MODIS daily LST time series is applied here to identify the real snow/ice pixels. That is, a pixel that has a LST value $LST > T_{LST}$ will be excluded (in this study, the threshold is set to $T_{LST} = 1$ °C).

Similar to the MODIS reflectance data, the LST data also have the cloud and NVD problem, the interpolation for these gaps will be introduced in section 3.4.

When a pixel has NVD on a certain day (Figure 1a), it is labeled as NVD for that day. Technically, it includes two kinds of situations:

1. when the image is not acquired, all pixels of that image are labeled as NVD on that day. For the absent image, the corresponding water cover map is interpolated with time series processing, which is described in section 3.2.2.
2. when the image is not absent, the 500-m MODIS State QA layer is used. If the first two bits of the State QA layer are equal to 10 or 11, then the pixel is referred to NVD.

3.2. Rule-Based Water Classification

In the solar spectral range, a typical spectral characteristic of water is that it has higher reflectance at the VIS bands than at the SWIR range (Ji et al., 2015). However, due to atmospheric correction error, some low

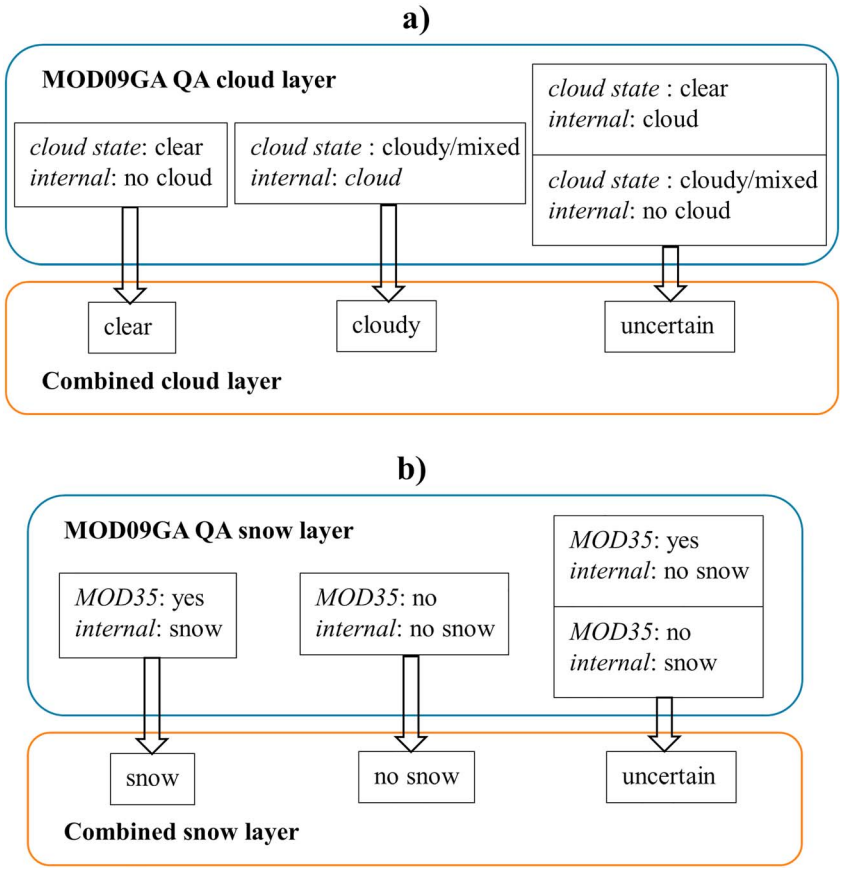


Figure 4. Combined layers for cloud (a) and ice/snow (b) generated from 1-km MOD09GA State QA layer. Internal in (a) = internal cloud algorithm flag; MOD35=MOD35 snow flag; internal in (b) = internal snow mask.

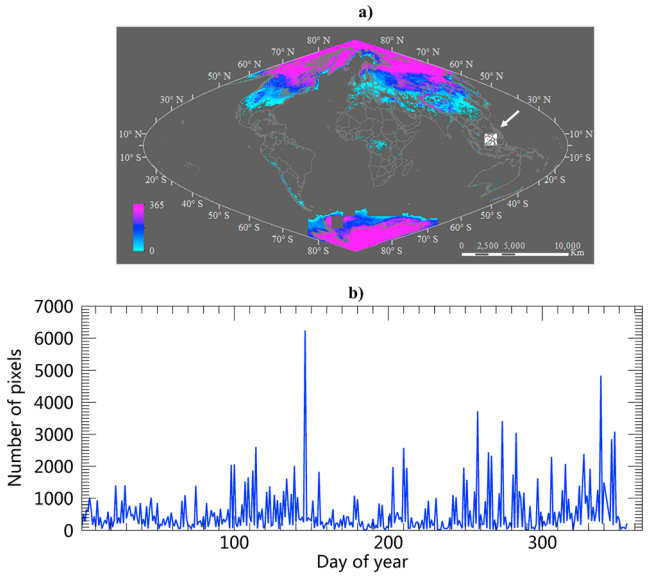


Figure 5. The global ice/snow cover days derived from 1-km MOD09GA State QA layer in 2015 (a) and the number of pixels labeled as ice/snow for tile "h29v08" (b), which is located in the tropical area where there is no ice/snow.

reflectance water pixels may fail to exhibit this spectral property. Therefore, we first divide the water into two types, high-reflectance and low-reflectance water based on VIS reflectance. Suppose *maxVIS* = max {Band 1, Band 2, Band3}, representing the maximum reflectance value in the VIS bands. Then the definition for the three types of water are

$$\begin{cases} \text{high} - \text{reflectance}: & maxVIS \geq T_{VIS} \\ \text{low} - \text{reflectance}: & maxVIS < T_{VIS} \end{cases}, \quad (1)$$

where the thresholds are set to $T_{VIS} = 0.05$ in this study. The strategies to detect the two types of water are different, and the detailed description is as follows.

3.2.1. High-Reflectance Water Classification

The spectral characteristic of water that its VIS bands has higher reflectance than the SWIR range can be represented using the following simple water index (WI),

$$WI = \begin{cases} 0, & if\ maxVIS \leq maxSWIR \\ 1, & if\ maxVIS > maxSWIR \end{cases}, \quad (2)$$

where variable *maxSWIR* is defined as max {Band 6, Band 7}, meaning the maximum reflectance value in SWIR spectral ranges.

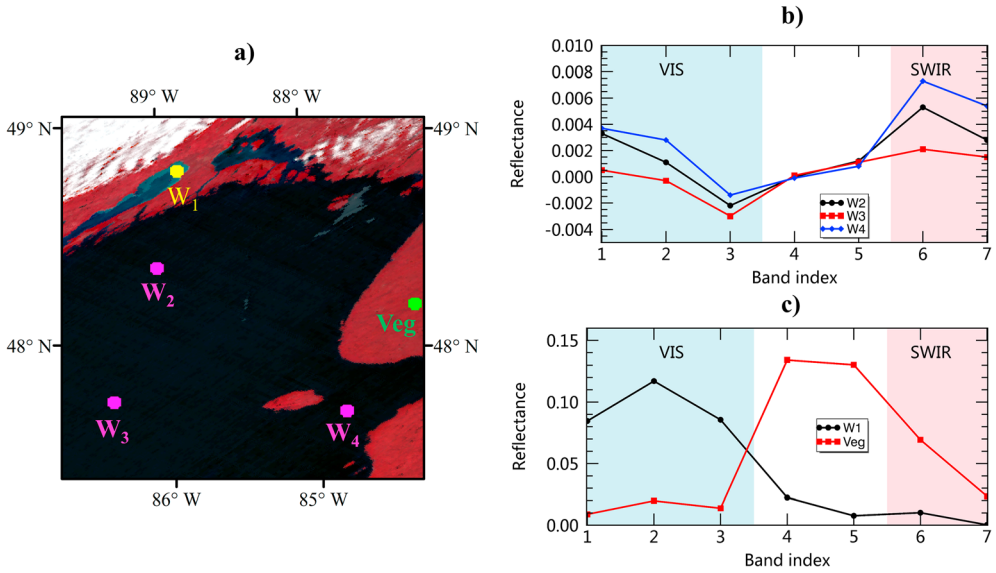


Figure 6. Illustration of water spectra from erroneous atmospheric correction, as $W_{2\sim4}$ and the confusion between water (W_1) and vegetation (Veg) in visible bands. (a) Locations of water and vegetation samples and (b and c) the corresponding spectral signatures. The image covers the northern part of Lake Superior. VIS = visible; SWIR = short wave infrared.

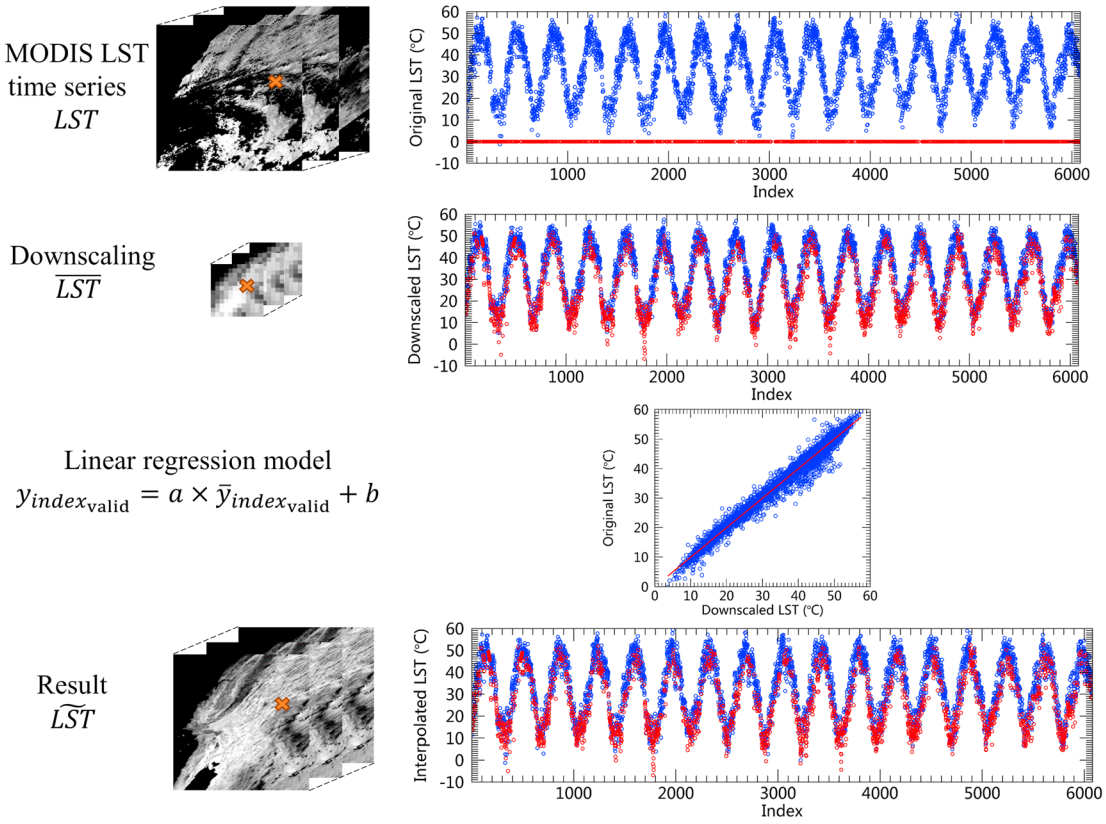


Figure 7. Illustration of the time series process for MODIS LST data (valid data in blue and originally no valid data in red). LST = land surface temperature; MODIS = Moderate Resolution Imaging Spectroradiometer.

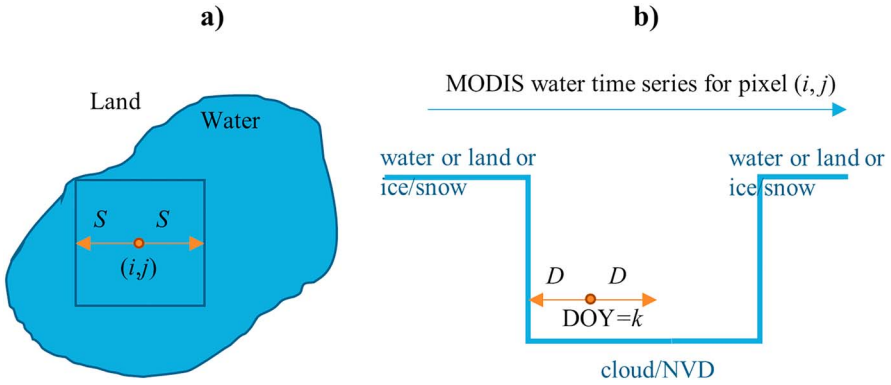


Figure 8. Illustration for the definition of spatial limit, *S* (a) and temporal limit, *D* (b). MODIS = Moderate Resolution Imaging Spectroradiometer; DOY = day of year; NVD = no valid data.

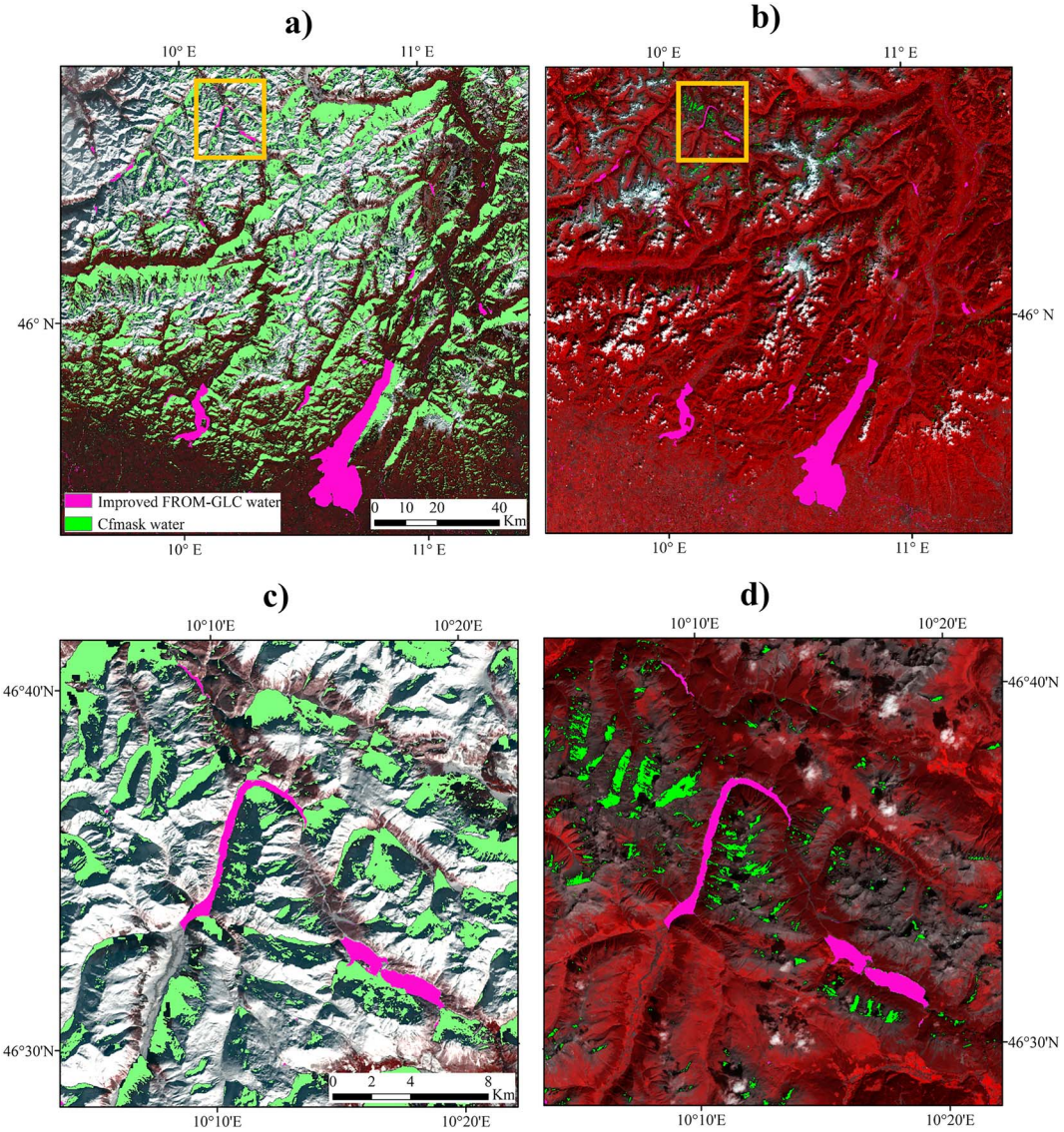


Figure 9. Illustration of the problem in the water map in the Cfmask layer for Landsat image (path = 203, row = 028 with location corresponding to the arrow in Figure 2). (a) False color image with day of year = 001 (2014, R: NIR band, G: Red band, B: Green band) and (b) for day of year = 257. (c and d) Zoom images for the squares in (a) and (b).

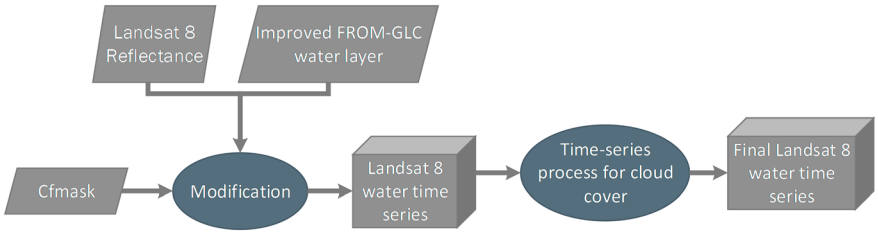


Figure 10. Flow of the water classification for Landsat-8 time series.

In addition, due to the strong water absorption in the SWIR range, the SWIR bands usually have very low reflectance for water. The SWIR reflectance is also used for water classification in this study. The classification rule for high-reflectance water is

$$\text{condition 1 (high} - \text{reflectance water)}: \; maxVIS \geq T_{VIS} \text{ and } WI = 1 \text{ and } maxSWIR < T_{SWIR}, \text{ where the thresholds are set to } T\text{SWIR} = 0.1.$$

3.2.2. Low-Reflectance Water Classification

However, when the total water reflectance is low, its VIS and SWIR bands may not satisfy *WI* = 1 because of atmospheric correction errors. An example is shown in Figure 6 to illustrate this phenomenon. For water pixels $W_{2\sim4}$, the reflectance of their SWIR bands should be lower based on the absorption characteristics of water. Yet, as shown in Figure 6b, $W_{2\sim4}$'s SWIR bands have a higher reflectance value than the VIS bands. Accordingly, their *WI*s are equal to 0, which is obviously wrong. Therefore, another important feature of this type of water is used here, that is, water generally has low reflectance in VIS and SWIR bands.

However, vegetation-dominated pixels (like wetland) could also have low reflectance in these bands (Veg in Figure 6c). So the normalized difference vegetation index (NDVI) is also introduced to reduce the confusion between water and "dark" vegetation. The NDVI index in the form of MODIS band designation is expressed in the following:

$$NDVI = (\text{Band 4} - \text{Band 3})/(\text{Band 4} + \text{Band 3}). \tag{3}$$

To sum up, if a pixel satisfies the following condition, it will also be classified as water,

$$\text{condition 2 (low} - \text{reflectance water)}: maxVIS < T_{VIS} \text{ and } maxSWIR < T_{SWIR} \text{ and } NDVI < T_{NDVI}, \text{ where the NDVI threshold is set to } T\text{NDVI} = 0.2.$$

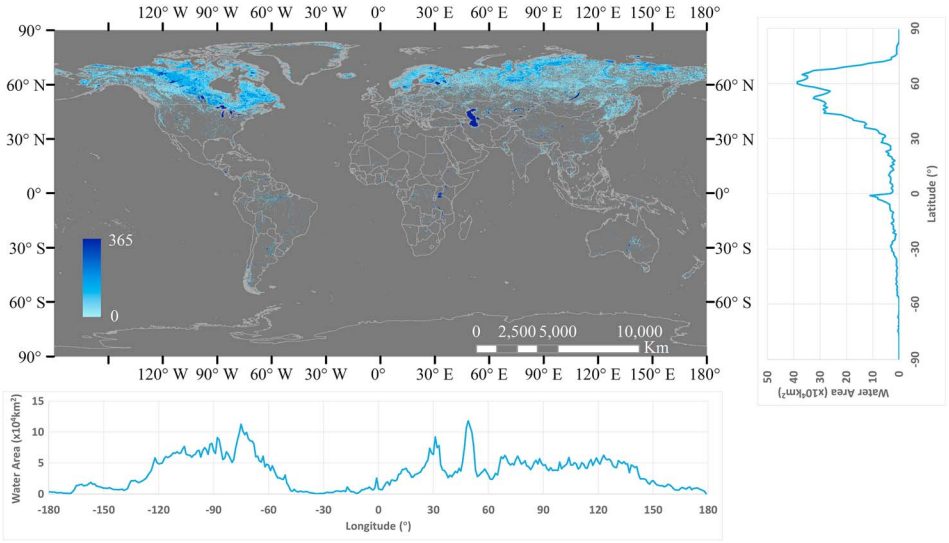


Figure 11. Distribution of global water cover days for 2010. The image stretching mode is *Minimum-Maximum*.

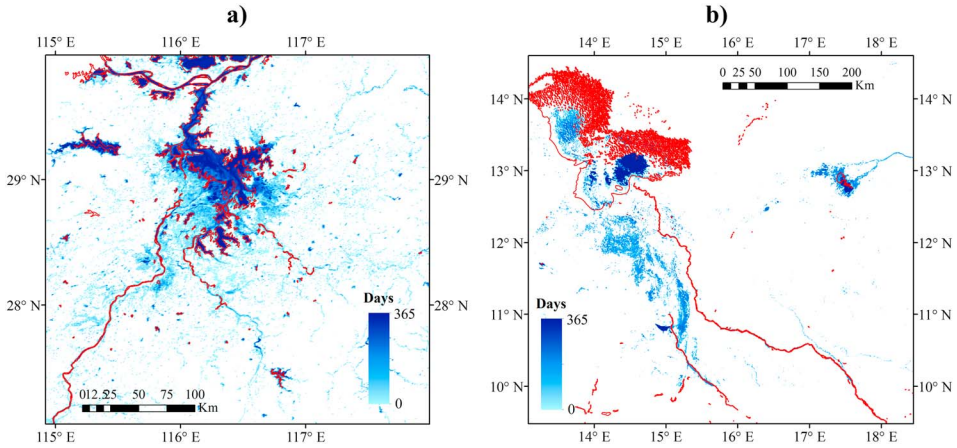


Figure 12. Comparison between our result (2010 water cover days) and global lakes and wetlands database (red) in (a) Poyang Lake and (b) Lake Chad.

It should be noted here that the inconsistent cloud- or ice/snow-covered pixels (i.e., cloud or ice/snow flag equal 1 or 2) will be relabeled as water if it satisfies either condition (1) or (2). Thereby, water that has been misrecognized by MODIS cloud and ice/snow flags can be corrected by our algorithm.

3.3. Object-Based Post-Classification

Shadows caused by cloud and mountains also have relatively low reflectance as water. For many prevalent classifiers, shadows are easily misclassified as water, which becomes the largest source of error in water mapping. An object-based postclassification method is adopted here, which has been successfully applied to remove incorrectly classified water pixels in Landsat water mapping (Ji et al., 2015). To remove these shadows, this method computes the topographical feature, spectral feature, and geometrical relation with cloud for each water object and sets specific rules to determine whether the water object is misclassified.

Since spectral features (WI) have already been used in the rule-based classification procedure, we only compute the average terrain slope and cloud shadow probability for each water object. The cloud shadow flag for each pixel is directly derived from the MODIS State QA layer. Next, a water object satisfies any of the following conditions will be ruled out as shadows:

Exclusion criterion 1 (mountain shadow): the average slope $> T_{\text{slope}}$,
Exclusion criterion 2 (cloud shadow): the cloud shadow probability $> T_{\text{shadow}}$.

The improved 30-m FROM-GLC water map is used for better threshold setting. Looser criteria will be set if a water body has a greater overlap with the FROM-GLC water layer. Here we use the sigmoid function to adjust the thresholds. That is, $T_{\text{slope}} = 4 \times S(10 \times [p_{\text{FROM}-\text{GLC}} - 0.5]) + 6$, where $S(x) = 1/(1 + e^{-x})$ is the sigmoid function, and $p_{\text{FROM}-\text{GLC}} \in [0, 1]$ is the overlap extent of the water body with the FROM-GLC water layer. Similarly, $T_{\text{shadow}} = 0.2 \times S(10 \times [p_{\text{FROM}-\text{GLC}} - 0.5]) + 0.8$. As a result, some small-sized water bodies in mountainous or cloudy areas will have a larger chance to avoid being mistakenly deleted in this procedure.

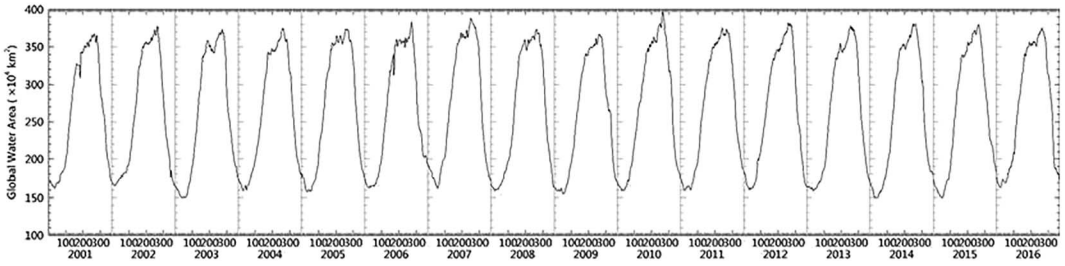


Figure 13. Daily global water area dynamics from 2001 to 2016.

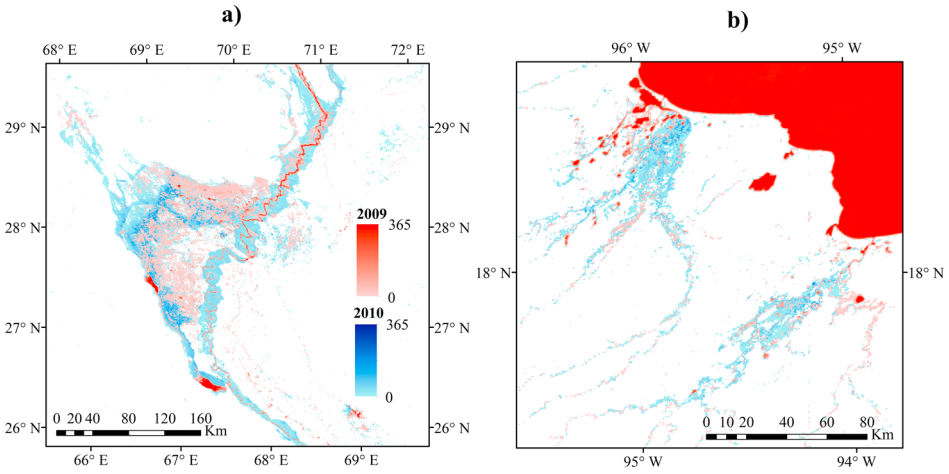


Figure 14. Comparison of water cover days between 2009 and 2010 in Pakistan (a) and Mexico (b).

3.4. Time Series Processing

To solve the cloud cover and NVD problem (Figure 1), a temporal-spatial filtering algorithm is adopted for both MODIS daily LST and water classification time series. Since the former is continuous while the latter is discrete, their detailed processing steps are different.

3.4.1. MODIS LST Time Series Processing

An example diagram for MODIS LST time series process is presented in Figure 7, which aims to provide continuous values for cloud/NVD gaps in LST data. Suppose $LST_{M \times N \times L}$ is the input MODIS LST time series, where M and N represent numbers of lines and columns and L stands for the number of days (here $M = N = 1{,}200$ and $L = 12 \times 365 + 4 \times 366$). First, we downscale the data in the spatial dimension with a factor of K pixels using a spatial average method, and the downscaled data are noted as $\overline{LST}_{\overline{M}\times\overline{N}\times L}$, where $\overline{M} = fix(M/K)$, and $\overline{N} = \mathrm{fix}(N/K)$ (the operator fix(.) means returning a integer value). The selection criterion for K is to choose the minimum integer which can guarantee that the time series for each pixel in $\overline{LST}_{\overline{M}\times\overline{N}\times L}$ does not contain invalid value (here $K = 100$). Then a linear regression method is applied for each pixel to fix the cloud and NVD problem. For simplicity, let $y_{1 \times L} = LST_{i,\,j,\,:}$ represent the LST time series for pixel (i, j). So the corresponding downscaled LST time series $\overline{y}_{1\times L} = \overline{LST}_{\overline{i},\overline{j},:}$, where $\overline{i} = fix(i/K)$, and $\overline{j} = fix(j/K)$. Suppose $index_{\mathrm{valid}}$ and $index_{\mathrm{non\ -\ valid}}$ represent the indices for the valid and invalid values in y. Then a linear regression model is established as

$$y_{index_{\mathrm{valid}}} = a \times \overline{y}_{index_{\mathrm{valid}}} + b. \tag{4}$$

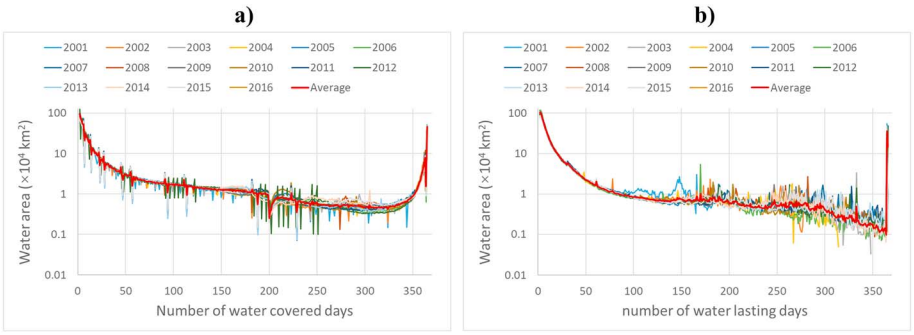


Figure 15. The distribution of water area with the number of water covered days (a) and the number of water lasting days (b). The y axis is plotted in a logarithmic scale.

Table 2
Confusion Matrix Using 720 Landsat Images

		Landsat		
		Non-water	Water	UA (%)
MODIS	Non-water	15,393,169	62,166	
	Water	75,022	1,091,581	93.57
	PA (%)		94.61	99.17

Note. MODIS = Moderate Resolution Imaging Spectroradiometer; UA = user's accuracy; PA = producer's accuracy.

Those nonvalid values in y can be interpolated by

$$\tilde{y}_{index_{invalid}} = a \times \overline{y}_{index_{invalid}} + b. \quad (5)$$

Therefore, the final cloud and NVD-free LST time series $\widetilde{LST}_{M\times N\times L}$ is combined by

$$\widetilde{LST}_{i,j,index_{valid}} = y_{index_{valid}} \text{ and } \widetilde{LST}_{i,j,index_{invalid}} = \tilde{y}_{index_{invalid}}. \quad (6)$$

3.4.2. MODIS Water Time Series Processing

In order to produce a gap-free MODIS water time series, all the cloud/NVD pixels are eliminated based on certain rules considering both spatial and temporal consistency. Two parameters, the spatial limit S and temporal limit D, are first calculated for each pixel (i, j), which are defined as

1. Spatial limit S. The improved FROM-GLC water layer is used, noted as $W^{FROM-GLC}$. For pixel (i, j), if it is a water pixel in the FROM-GLC water layer, the spatial limit S is defined as the minimum value that the S neighborhood of pixel (i, j) is all water (Figure 8a). On the other hand, if pixel (i, j) is a land pixel in the FROM-GLC water layer, then $S = 1$. In the following interpolation process, for pixel (i, j), only its $(2S + 1) \times (2S + 1)$ neighborhood pixels are involved.
2. Temporal limit D. For pixel (i, j, k) (where i and j stand for the spatial coordinates and k is the temporal coordinate), D represents the lasting cloud/NVD days around pixel (i, j, k) in the temporal dimension, as illustrated in Figure 8b. It is the minimum value that the D temporal neighborhood of pixel (i, j, k) is all cloud/NVD. The temporal limit D is used to determine the temporal range for interpolation. The detailed interpolation method is as follows.

Then, monthly downscaled water-cover time series $\overline{W}^{year}_{M\times N\times \overline{K}}$ were calculated in time, where year = 2001, …, 2016, and $\overline{K} = 12$. For each month from January 2001 to December 2016, a voting method is applied by comparing the probability of pixel (i, j) being land, ice/snow, and water with only valid classification values. If pixel (i, j) does not have any valid label for a whole month, $\overline{W}_{i,j,\overline{k}}$ is kept as cloud/NVD.

To process pixel (i, j, k) that is labeled as cloud/NVD, if $D < T_D$, which indicates a short-term cloud/NVD period (here $T_D = 7$), a spatial temporal nearest neighborhood method is applied. That is, $W_{i,j,k}$ is replaced by the nearest valid label in the $W_{i-S:i+S,\, j-S:j+S,\, k-D-1:k+D+1}$. For the same distance, the priority rank is *snow/ice* > *water* > *land*. On the other hand, if $D \geq T_D$, which implies a long-term cloud/NVD period, $W_{i,j,k}$ is assigned to $W_{i,j,k} = \overline{W}^{year}_{i,j,\overline{k}}$. If $\overline{W}^{year}_{i,j,\overline{k}}$ is labeled as cloud/NVD, we applied the nearest neighborhood method on the time series $\left[\overline{W}^{2001}_{i,j,\overline{k}}, \ldots, \overline{W}^{2016}_{i,j,\overline{k}}\right]$.

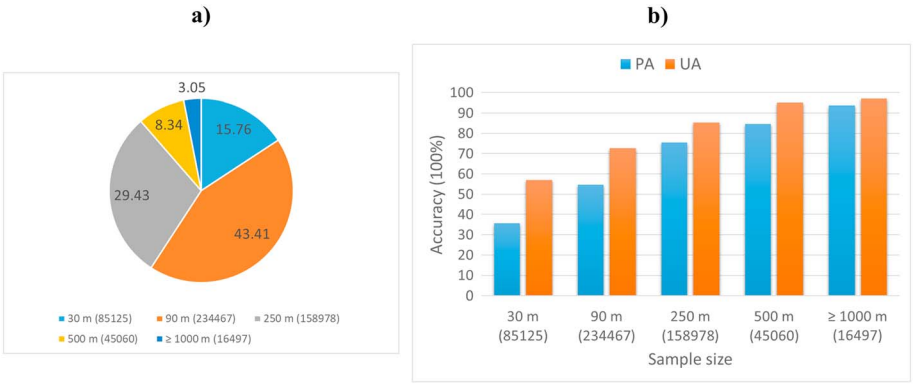


Figure 16. The accuracy of Moderate Resolution Imaging Spectroradiometer water product based on FROM-GLC samples. (a) The distribution of FROM-GLC sample units of different spatial sizes and (b) corresponding PA and UA of the Moderate Resolution Imaging Spectroradiometer water product. The number in the bracket represents the number of sample units. PA = producer's accuracy; UA = user's accuracy.

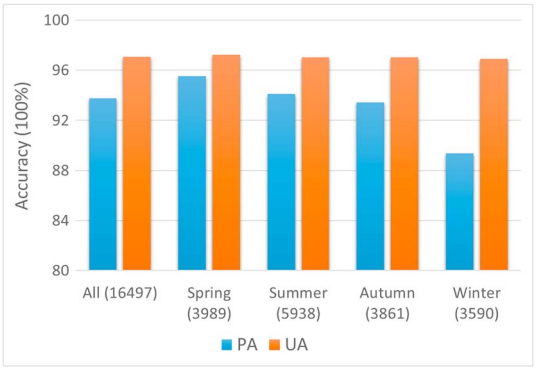


Figure 17. Accuracy assessment for Moderate Resolution Imaging Spectroradiometer water product for spring, summer, autumn, and winter with large FROM-GLC samples (≥1,000 m × 1,000 m). In the north (south) hemisphere, spring: March–May (September–November), summer: June–August (December–February), autumn: September–November (March–May), winter: December–February (June–August). The number in the bracket represents the number of sample units. PA = producer's accuracy; UA = user's accuracy.

3.5. Landsat Image Processing for Validation

The Cfmask in the Landsat surface reflectance package, which is mainly produced for cloud and cloud shadow identification, also includes classes for water and snow/ice (Zhu & Woodcock, 2012). Therefore, the water layer from Cfmask for each Landsat image can be used to validate the MODIS result. For Landsat data acquired in the summer with relatively flat terrain, the Cfmask water result is quite reliable. However, for mountainous areas, shaded slopes covered by snow/ice are easily misclassified as water, even in the summer season (Figure 9). To solve this problem, the improved FROM-GLC water results are introduced as an auxiliary data to exclude mountain shadows with steep slopes. Based on the algorithm for mountainous shadow removal (Ji et al., 2015), a simplified method is applied here. Water objects satisfy either of the following two conditions will be relabeled as nonwater,

1. average slope $>$ T_{slope} and the percentage of FROM-GLC being water $<$ $T_{FROM\text{-}GLC}$;
2. average NDVI $>$ T_{NDVI} and the percentage of FROM-GLC being water $<$ $T_{FROM\text{-}GLC}$,

where the percentage of FROM-GLC being water is calculated by the number of pixels being labeled by FROM-GLC as water divided by the total number of pixels in the water object. The thresholds used here are set as T_{slope} = 8°, T_{NDVI} = 0.2, and $T_{FROM\text{-}GLC}$ = 30%.

In addition, the cloud cover problem is implemented by using the time series interpolation (Figure 10). Finally, careful manual inspection for each Landsat water mapping result is done to ensure a high accuracy validation data set.

4. Results

Our daily 500-m water map is a global raster data set in the same sinusoidal projection as in MODIS data products. There are four discrete values as 0-Land, 61-Water (unfrozen), 62-Water (frozen), and 255-Fill (areas of no value). A data quality layer is also produced for all water pixels in each tile in two states: high quality and low quality (cloud or NVD pixels). Figure 11 presents the global distribution of water cover days (both frozen and unfrozen states) in 2010, and some details are given in Figure 12 for two highly dynamic lakes, Poyang Lake and Lake Chad. Compared with the static water product, like the global lakes and wetlands database (Lehner & Döll, 2004), our result is able to reveal the temporal changes for those dynamic water systems.

Table 3
Comparison of Global Water Area for Different Water Products

Product name	Main data sources	Spatial resolution	Temporal range	Temporal resolution	Water area ($\times 10^4$ km^2)
MOD44W (Carroll et al., 2009)	MOD44C (optical)	250 m	2000–2007	Single-period	373
GIW (Feng et al., 2016)	Landsat (optical)	30 m	2000	Single-period	365
Improved FROM-GLC (Ji et al., 2015)	Landsat (optical)	30 m	2010	Single-period	407
Pekel et al. (2016)	Landsat (optical)	30 m	1984–2015	Monthly	Extent[a]: 452 Permanent[b] (2015): 278
Prigent et al. (2012)	SSMI (passive microwave) ERS (active microwave) AVHRR (optical)	~25 km	1993–2004	Monthly	Min[c]: 210 Max[d]: 590
This study	MOD09GA (optical)	500 m	2001–2016	Daily	Min: 150 Max: 380 Extent: 490 Permanent (average): 125

[a]All the locations ever detected as water over a period of time. [b]12 months of water over a year. [c]The water area with the least water cover over a period of time. [d]The water area with the most water cover over a period of time.

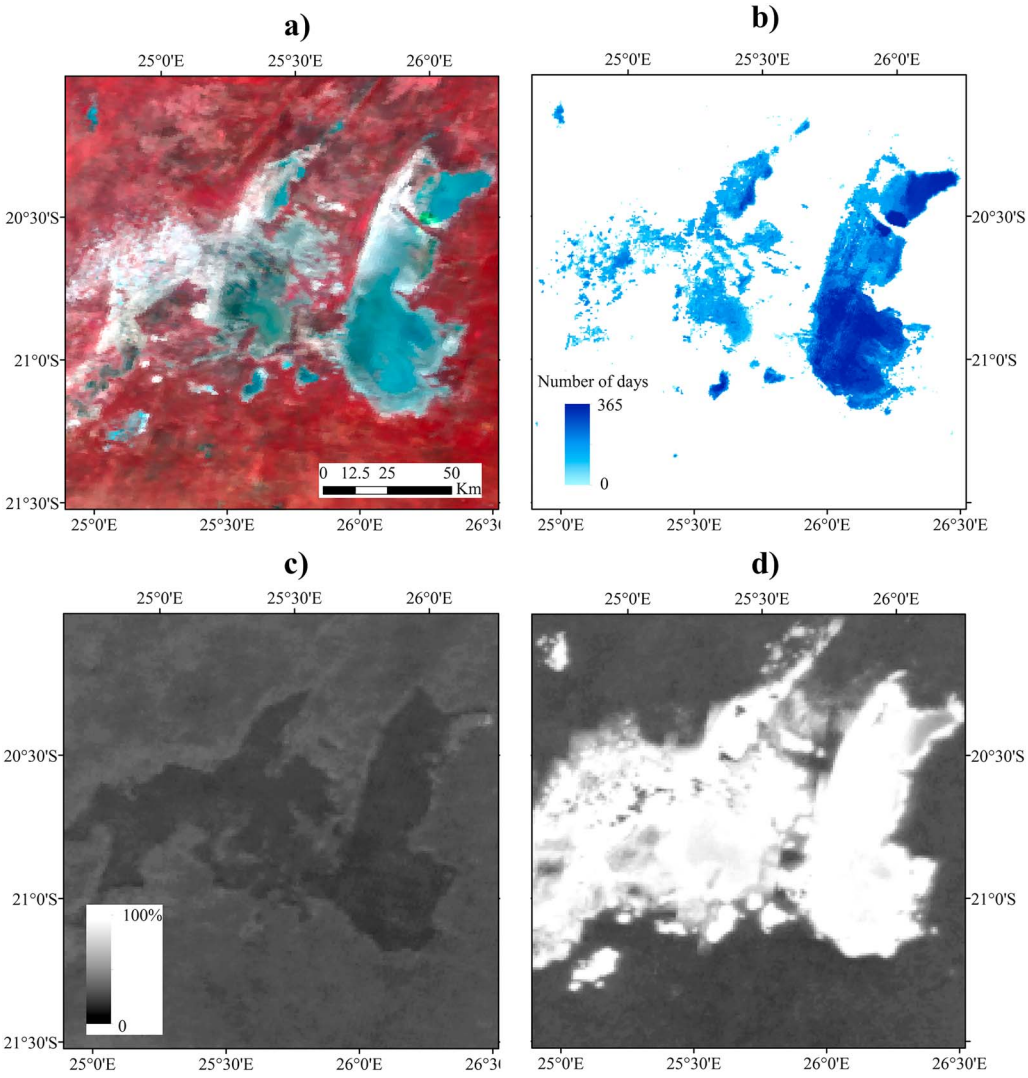


Figure 18. The cloud detection problems of Moderate Resolution Imaging Spectroradiometer State QA layer at high-reflectance water bodies. (a) False color image at SaltPans (year = 2015, day of year = 003). (b) The number of water cover days in 2015 of our product. (c) The percentage of cloud cover labeled as "cloudy" or "mixed" in the *cloud state* layer, and (d) the percentage of cloud cover labeled as "cloud" in the *internal cloud algorithm flag* for 2015.

Daily change of global inland water areas is shown in Figure 13. Generally, global water area reaches its maximum (~380 × 10^4 km^2) in September and its minimum (~150 × 10^4 km^2) in February. It is mainly because (a) water will freeze in winter in the northern hemisphere and (b) most of the water is distributed in the northern hemisphere. Our result also reveals the importance of water mapping in a daily frequency. It can be found that water covers more area in some year (like 2010) than other years (like 2009). Examples of water cover days for 2009 and 2010 of Pakistan and Mexico are shown in Figure 14. Both areas have suffered severe flood disasters in 2010.

The distributions of water area versus the number of water covered days and the number of water lasting days are shown in Figure 15. Generally, water area decreases as the number of water covered/lasting days increases, except when the number of water covered/lasting days ≥360. It is because some large lakes are mostly covered by water, such as the Caspian sea, Lake Victoria, the Great Lakes in North America, and so forth. From Figure 15, we can see that high temporal resolution is important to mapping the water dynamics.

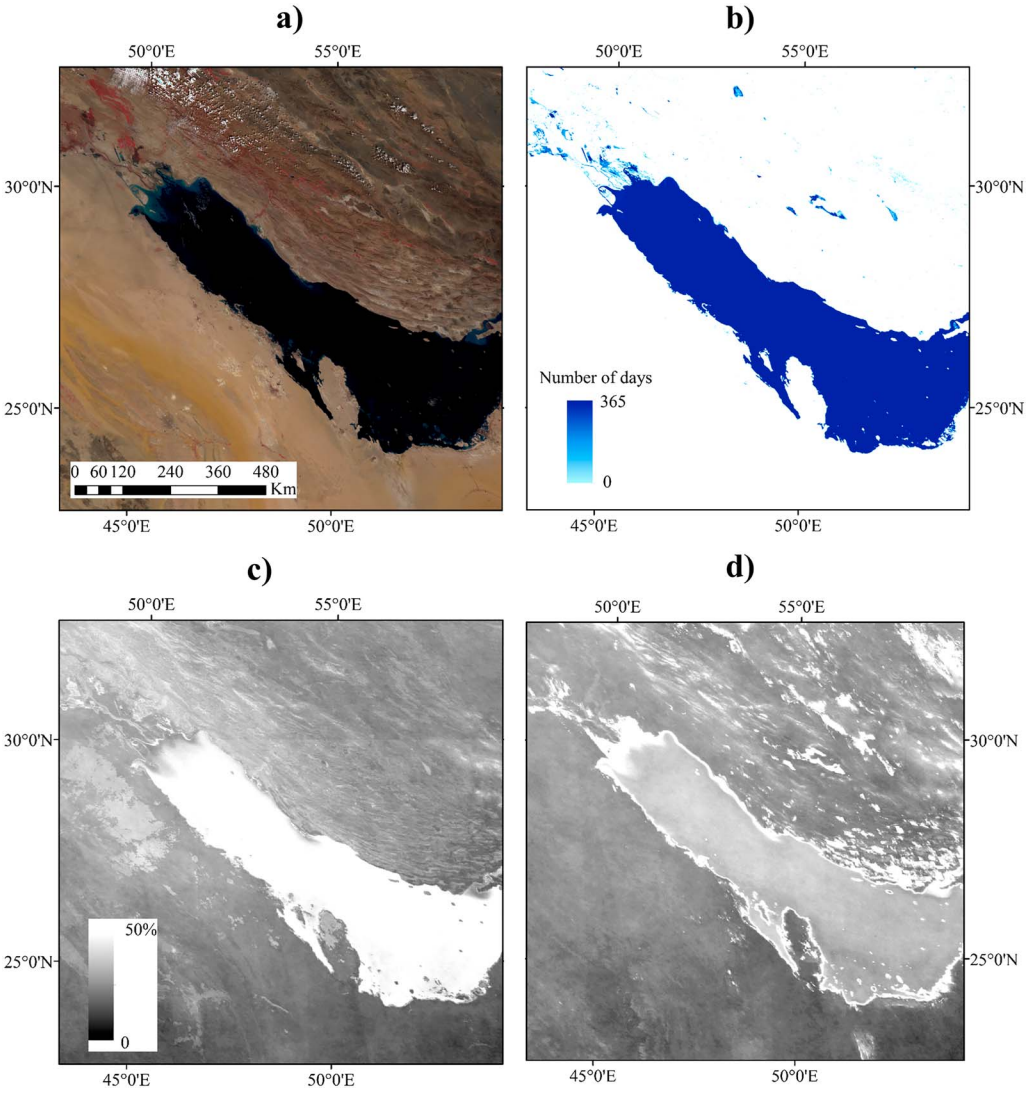


Figure 19. The cloud detection problems of Moderate Resolution Imaging Spectroradiometer State QA layer at ocean. (a) False color image at Persian Gulf (year = 2015, day of year = 282). (b) The number of water cover days in 2015 in our product. (c) The percentage of cloud cover labeled as "cloudy" and "mixed" in the *cloud state* layer, and (d) the percentage of cloud cover labeled as "cloud" in the *internal cloud algorithm flag* layer for 2015.

5. Validation

The accuracy of our product is assessed by using both the image-based data (Landsat-8 images) and the point-based data (FROM-GLC sample sets). Each Landsat image is first classified into two classes, water and nonwater, and then resampled into a 500-m resolution water fraction map in Sinusoidal projection. Finally, the reference Landsat-8 water map is produced by labeling all the pixels with water fraction no less than 50% as water. A pixel-to-pixel comparison with the MODIS result is performed, and the producer's accuracy (PA), user's accuracy (UA), and overall accuracy are computed (Table 2). It can be seen that both PA and UA are higher than 93%.

FROM-GLC sample units with different sizes are used to evaluate our product separately (Figure 16). It can be seen that more than half of the sample units are smaller than 90 m × 90 m in area. Only about 11% of the sample units are ≥500 m × 500 m. With respective to the accuracy, both PA and UA increase as the size of sample units increases, and both reach 90% when the area of sample units are ≥1,000 m × 1,000 m. The result indicates that our product has higher performance at larger-sized water bodies. Due to misregistration

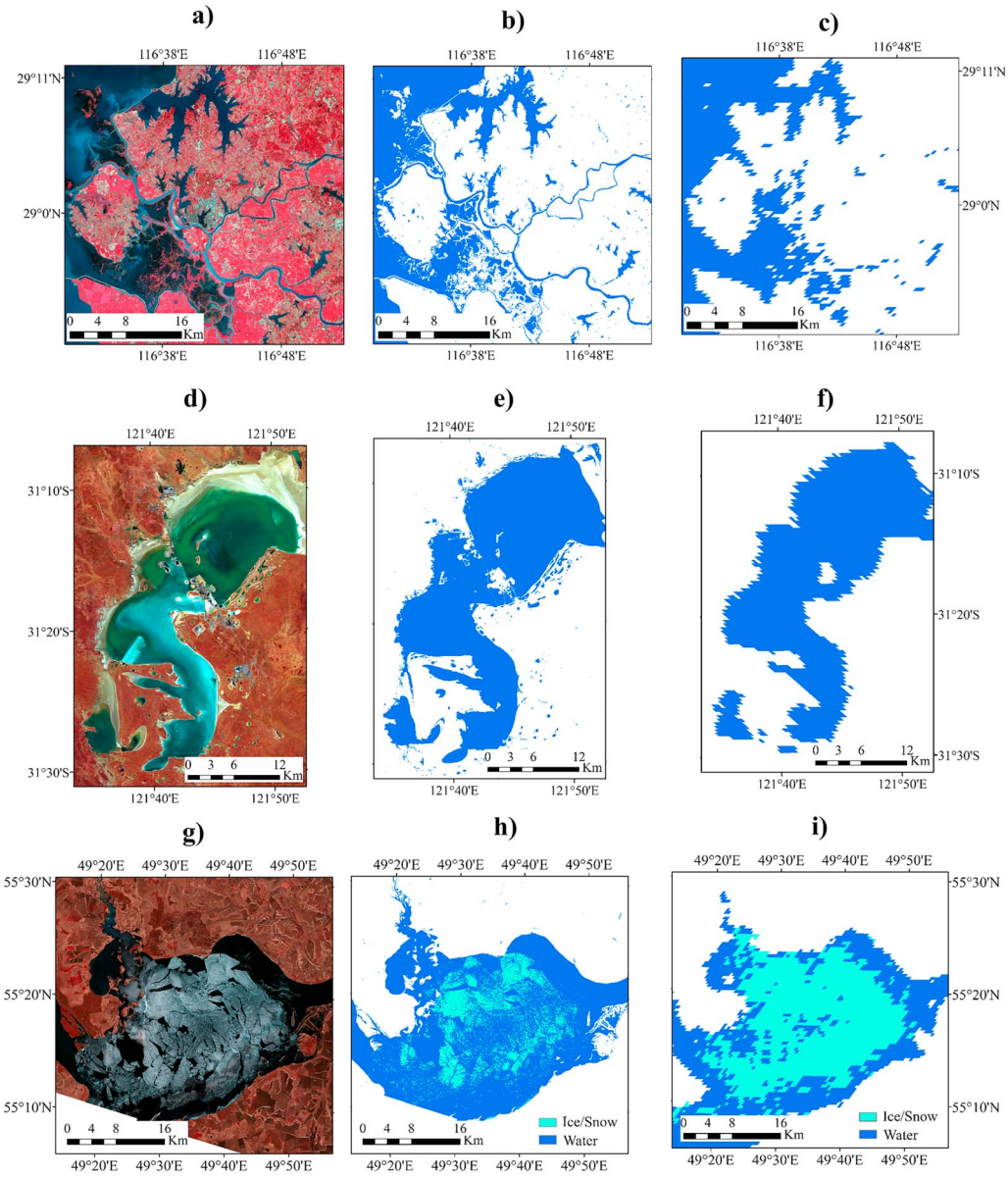


Figure 20. The spatial limits of MODIS data in water mapping at flooded wetland (a–c), saline lake (d–f), and freezing lake (g–i). Flooded wetland: (a) false color image of Landsat 8 data (path = 121, row = 040, year = 2014, DOY = 281), (b) Landsat result, and (c) MODIS result for the same day. Saline lake: (d) false color image of Landsat 8 data (path = 109, row = 082, year = 2014, DOY = 037), (e) Landsat result, and (f) MODIS result for the same day. Freezing lake: (g) false color image of Landsat 8 data (path = 170, row = 021, year = 2014, DOY = 112), (h) Landsat result (the ice classification is derived from Cfmask), and (i) MODIS result for the same day. DOY = day of year; MODIS = Moderate Resolution Imaging Spectroradiometer.

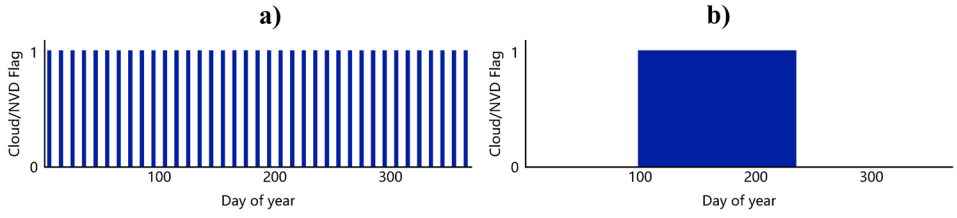


Figure 21. Illustration of cloud/NVD distribution over time. (a) Evenly distributed; (b) connectedly distributed. NVD = no valid data.

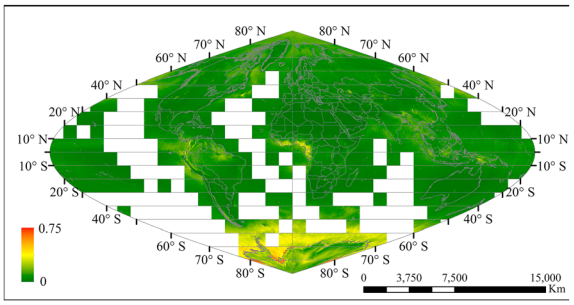


Figure 22. Global distribution of the average cloud index from 2001 to 2016. The image stretching mode is *Maximum-Minimum*.

uncertainties, 500-m resolution data cannot be perfectly overlaid. Therefore, a lower accuracy is expected for sample units smaller than 1,000 m × 1,000 m. Limited by the resolution of our data set, we will only use those sample units ≥1,000 m × 1,000 m in area in subsequent analysis.

Next, we consider the temporal effect on the accuracy of our product using only sample units ≥1,000 m × 1,000 m. The corresponding PA and UA are shown in Figure 17. It can be seen that the lowest PA in winter is approximately 89%, at least 3% lower than those in other seasons. The reasons could be explained as follows: (1) the low solar incidence angle in the higher northern latitudes causes low data quality of MODIS data in winter. As a result, the water detection accuracy is lower. (2) There are errors in frozen water bodies of FROM-GLC sample units. Thirty-five winter FROM-GLC sample units are classified as ice/snow by our product, among which 23 can be confirmed as ice/snow by careful manual checking. It indicates the effectiveness of using LST in distinguishing ice from water. The remaining 12 sample units are partially frozen lakes or unfrozen lakes or sea water surrounded by snow/ice covered land.

Finally, a comparison between our data set with other global water mapping products from different data sources and with different spatial/temporal resolution is tabulated in Table 3. It can be found that compared with single-period products, multitemporal water mapping can always map more water with respect to *Extent* or *Max* water area, which infers that it is necessary to develop dynamic water mapping. Among them, Prigent's product, though with the coarsest spatial resolution, can achieve the largest maximum water area. The most probable reason is that they used microwave data to detect water, which can capture inundated areas even covered with vegetation, and thus can detect more water area.

6. Discussions

6.1. Effects of Cloud Contaminations and Spatial Resolution

The MODIS09GA product contains a 16-bit 1-km State QA layer, which has two flags for cloud description: the *cloud state* (bit 0–1) and the *internal cloud algorithm flag* (bit 10). Figures 17 and 18 show the cloud cover percentage of both flags in two areas. It can be seen that the internal cloud algorithm flag overdetect clouds at

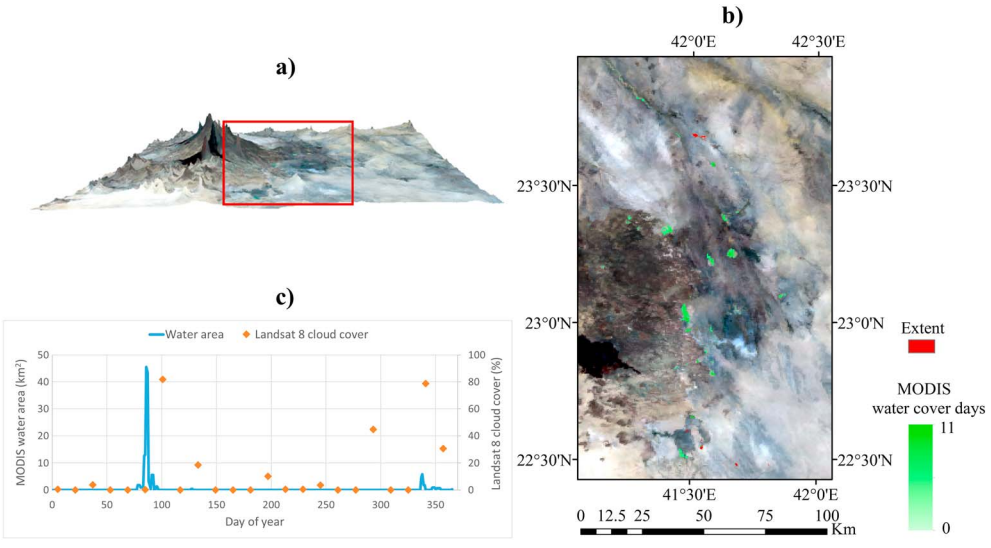


Figure 23. Illustration of the advantage of daily water mapping in capture fast changes. (a) The 3-D model of our study area (MODIS HV = "h21v06", year = 2015 and DOY = 086) based on Shuttle Radar Topographic Mission Digital Elevation Model. (b) MODIS water cover days (green) and Pekel's 30-m Landsat water mapping product (extent, red) and (c) the daily water cover area curve and the cloud cover for all the Landsat-8 images in 2015. MODIS = Moderate Resolution Imaging Spectroradiometer.

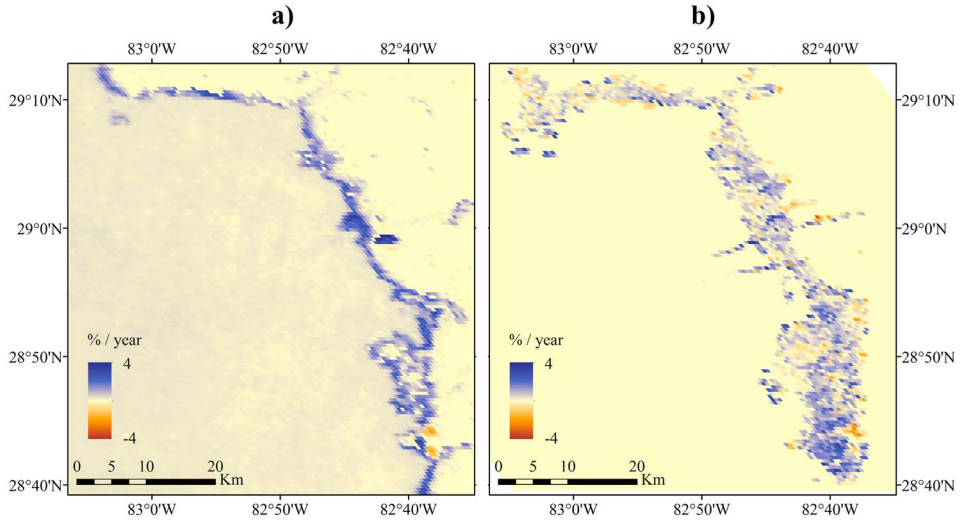


Figure 24. (a) Annual change rate determined from our daily Moderate Resolution Imaging Spectroradiometer product in west Florida (b) determined from Landsat images as in (W. Li & Gong, 2016).

shallow water areas (Figure 18), while the cloud state overdetect clouds in large open water areas (Figure 19). Our algorithm corrected a large proportion of these areas with inconsistent cloudy flags in MODIS State QA layer.

Limited by resolution, MODIS data are not able to extract small-sized lakes, sinuous rivers, narrow water bodies, and fine pieces of ice (Figures 20a–20c). Compared with the Landsat-derived water mask, the loss of those fine features is the main source of omission error in the MODIS water maps. On the other hand, small-sized islands or ices in a large water area are more favorably recognized in the MODIS water maps, which increases the commission error. For the case of saline lakes, as exemplified in Figures 20d–20f, the MODIS water map can capture the water area well for the major water body except small water bodies. For partially frozen lakes, the MODIS water extraction produces more accurate results than the Cfmask result from the Landsat product (Figures 20g–20i).

6.2. Accuracy of Water Time Series Processing

The time series processing in section 3.2.2 is mainly to solve the cloud/NVD problem as shown in Figure 1. It can be found that the accuracy of the applied temporal-spatial filtering method is largely dependent on the cloud/NVD distribution over time. When the cloud/NVD is evenly distributed, as shown in Figure 21a, the cloud/NVD gaps are much easier to fill. However, if the cloud/NVD is distributed more continuously as shown in Figure 21b, pixels within the cloud/NVD window are hard to recover.

In order to demonstrate the influence of cloud/NVD cover to the accuracy of our product, we first design the following cloud index (I_{cloud}) as

$$I_{cloud} = {p_{covered}}^{0.5} \times {p_{lasting}}^{0.5}, \tag{7}$$

where $p_{covered}$ is the percent of cloud cover over time, which can be calculated by $p_{covered}$ = (number of days labeled as cloud/NVD) ⊠ (total number of days), and $p_{lasting}$ corresponds to the maximum cloud/NVD duration days, which can be calculated by $p_{lasting}$ = (maximum cloud/NVD duration days) ⊠ (total number of days). The data range of the cloud index is I_{cloud}∈[0,1].

The global distribution of the cloud index is shown in Figure 22, where a higher value indicates lower time series interpolation accuracy. It can be seen that I_{cloud} is low in most parts of the land areas in the world. However, the cloud index at northern part of South America, West Africa, and Southeast Asia around the Equator and Antarctica is relatively high, implying that the time series processing accuracy at these regions cannot be guaranteed. This problem should be improved to some extent when MODIS time series by Aqua is included. This will be assessed in a future study.

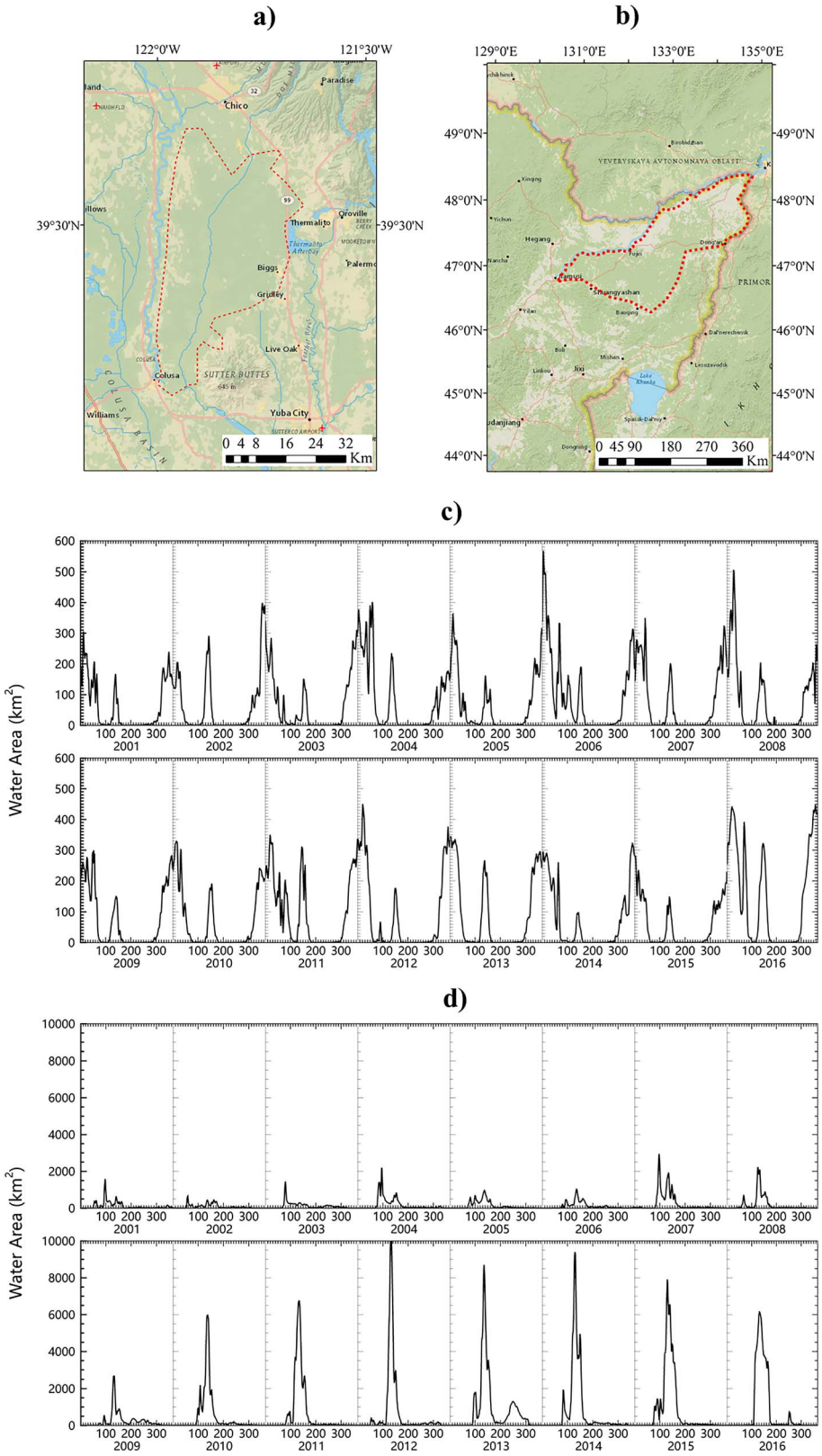


Figure 25. Two study areas of Paddy fields in the United States of America (a) and China (b). Corresponding daily water area changes are plotted in (c) and (d).

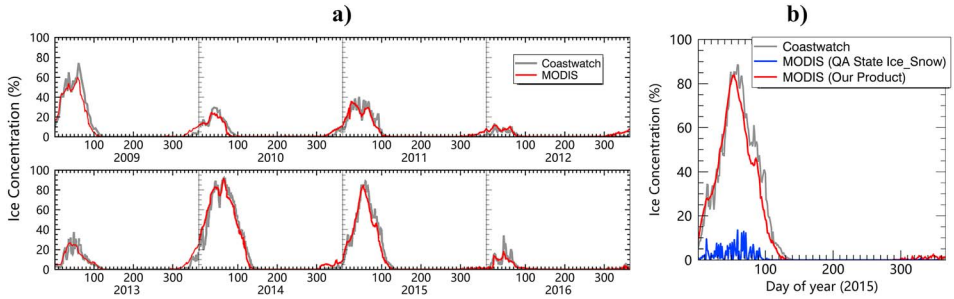


Figure 26. Comparison of ice concentration for the Great Lakes from 2009 to 2016 using Coastwatch and MODIS result (a) and the comparison with MODIS State QA ice/snow flag in 2015. MODIS = Moderate Resolution Imaging Spectroradiometer.

6.3. The Daily Water Time Series Captured Short Duration of Surface Water

Figure 23 shows the ability of our daily water maps in detecting short-term water existence. The region is located at the dry area of the Middle East, where the corresponding cloud index is low (Figure 22), so it is easy to grab reliable flash flood results. From the daily water area curves, we can see that there exist three water accumulation in this area for 2015, day of year (DOY) = 78–96, DOY = 127–128, and DOY = 337–354. On the other hand, due to its 16-day revisit period and cloud cover, Landsat may fail to capture these short life water surfaces, not to mention their maximum extent during the peaking time. The Landsat-based water product produced by Joint Research Centre with 32 years of data also fails to obtain these short-lived waters.

6.4. Sea Level Rise Effects on Land Submerging

On average, the sea level has risen at an annual rate of approximately 3.2 mm/yr. Therefore, it is natural to imagine that land will be gradually submerged by sea water for relatively stable coastal areas. However, the sea level rise is not rising evenly over the entire globe with the low-latitude Western Pacific region being the highest from 1996 to 2010 (Yang et al., 2013). Due to the tide effect, sea level variation is one of the most frequent surface phenomena on Earth. Therefore, the trend of sea level rise can only be estimated when there are abundant number of scenes. W. Li and Gong (2016) found in a vegetated mud flat area of west Florida that the more Landsat images, the better is the precision in determining the land-sea proportional change. To study the long-term changing trend of sea level, it is suitable to set the time interval for estimation of land-sea proportion on an annual basis. Our daily water mapping result can be used for detecting land-sea proportional changes over gently sloped coasts where subtle sea level rise could cover a large tract of lands.

We computed the probability of water for each pixel location in each year. The annual change rate is then determined. Figure 24 shows the annual change rate determined from our daily MODIS data and that derived with the method proposed in W. Li and Gong (2016) applied to Landsat data during 2001–2013. It can be seen that the general patterns do agree, and the majority of pixels affected are showing land loss and sea gain.

6.5. Differences of Rice Fields Use

Besides wetland, paddy fields have a flood season due to its plantation requirement. Two study areas in California, United States, and Heilongjiang Province, China, are selected for a comparison. The corresponding daily water changes can be seen in Figure 25. Comparing their water curves, we can observe different plantation styles. Both fields are irrigated in April (DOY = 100). Yet in California, the field is irrigated again in September (DOY = 260). From the fall to the next spring, rice fields in California were filled with water to help provide a wetland environment for wild birds. From Figure 25d, we can observe that the water area increases before 2012 and decreases gradually since then. This indicates that the area in Heilongjiang, China, has undergone a process of rice plantation increase up to around 2012, and thereafter, the expansion trend has been arrested. This indicates the potential of our daily water time series in studying different rice planting styles and timing of rice plantation.

6.6. More Precise Freezing Date

The freezing time can be derived from the MODIS State QA layer. Thus, it can be compared with the freezing data derived from our product. Since 2009, daily ice concentration of the Great Lakes in North America has

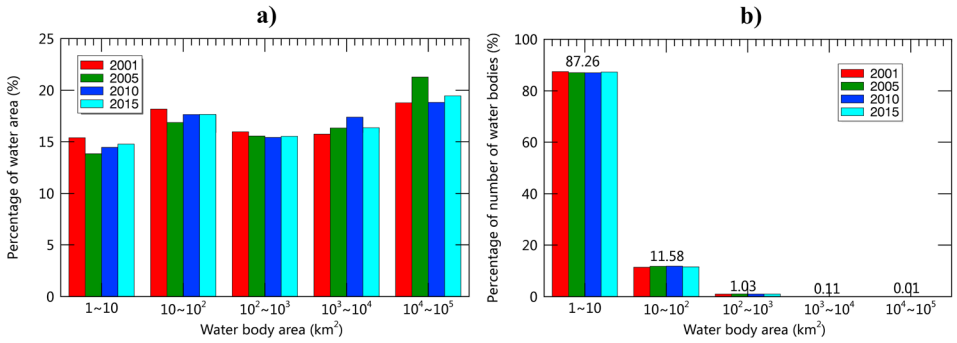


Figure 27. The size distribution of global water bodies in 2001, 2005, 2010, and 2015 for water bodies greater than 1,000 m × 1,000 m. (a) The percentage of number of water bodies = the number of water bodies within the size class / the total number of water bodies × 100%; and (b) the percentage of water area = the total area of water bodies within the size class / the total water area. The numbers in (a) are the average percentage of 2001, 2005, 2010, and 2015.

been posted at the Coastwatch website (https://coastwatch.glerl.noaa.gov/statistic/statistic.html). This can be regarded as a source of reference to compare the accuracy of our daily water product and the MODIS State layer (Figure 26). We can see that our result agrees well with the Coastwatch data. Large frozen areas occurred in 2013–2014, while the lowest happened in 2011–2012. Figure 26b also shows the comparison with MODIS State QA ice/snow flag, and it can be found that due to the cloud cover and NVD data gap, there were much fewer ice/snow pixels recorded by MODIS State QA layer.

6.7. Size Distribution of Water Bodies

The size distribution of global water bodies is shown in Figure 27. As shown in Figure 16b and discussed in section 6.1, due to the limitation of the 500-m resolution, we are unable to map small-sized water bodies in the world. Thus, only water bodies greater than 1,000 m × 1,000 m are used here. It can be seen that the number of water bodies decreases gradually as the size of water body increases, which agrees with previous studies (Downing et al., 2006; Lehner & Döll, 2004). Though the number of water bodies in the size class from 1 to 10 km^2 occupies more than 87%, they only account for ~15% of the total water area on land. On the other hand, though the number of water bodies in the size class of 10^4 to 10^5 km^2 is only 0.01% of the total number of water bodies in the world, they account for more than 15% of the total water area.

7. Conclusion and Future Works

In this paper, we reported a new daily global water data set has been produced in this research covering a 16-year period (2001–2016). It involved the use of more than 1.9 million tiles of MODIS daily reflectance data. The daily global water data set is freely available to potential users through email request to the corresponding author. It will be continuously updated as more recent MODIS data are processed with our algorithm.

Because we did not use auxiliary data in our interpolation of missing or cloud contaminated water areas, there is a limitation in our data in those areas. There is a trend that more satellite data will become available. In particular, high-resolution microwave/radar data are becoming more available. Such data are less affected by clouds, so they could help further reduce cloud or shadow effects and are also helpful for capturing water under vegetation (e.g., swamp forest). Moreover, subpixel approaches need to developed to capture water bodies smaller than a single pixel (Devries et al., 2017), especially for midresolution satellite data like MODIS data. In addition, constellation of small satellites acquiring high-resolution data at high-temporal frequency are also becoming available (e.g., the possibility of better than 10-m resolution daily global coverage from over 100 small satellites from PLANET company). Although there is a huge potential of integrating these various types of data for more precise water extraction from remotely sensed data, there is a lack of algorithms that can handle all data for mapping water surfaces. More efforts need to be made to the development of next generation of algorithms for consistent processing of those large volumes of data. We expect that the 16-year global water dynamics data set will receive a wide range of applications in the future.

Acknowledgments
This research is partially supported by the National Basic Research Program of China (973 Program) under grant 2015CB953701 and the Cross-strait Tsinghua Research Institute. The data is available at http://data.ess.tsinghua.edu.cn/modis_500_2001_2016_waterbody.html.

References

Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., & Weber, J.-L. (2008). GLOBCOVER the most detailed portrait of Earth. *European Space Agency*, *136*, 25–31.

Carroll, M., Townshend, J. R., DiMiceli, C. M., Noojipady, P., & Sohlberg, R. A. (2009). A new global raster water mask at 250 m resolution. *International Journal of Digital Earth*, *2*(4), 291–308. https://doi.org/10.1080/17538940902951401

Dai, Y., Trigg, M. A., & Ikeshima, D. (2015). Development of a global ~ 90 m water body map using multi-temporal Landsat images. *Remote Sensing of Environment*, *171*, 337–351.

Devries, B., Huang, C., Lang, M. W., Jones, J. W., Huang, W., Creed, I. F., & Carroll, M. L. (2017). Automated quantification of surface water inundation in wetlands using optical satellite imagery. *Remote Sensing*, *9*(8), 807. https://doi.org/10.3390/rs9080807

Downing, J. A., Prairie, Y. T., Cole, J. J., Duarte, C. M., Tranvik, L. J., Striegl, R. G., et al. (2006). The global abundance and size distribution of lakes, ponds, and impoundments. *Limnology and Oceanography*, *51*(5), 2388–2397. https://doi.org/10.4319/lo.2006.51.5.2388

Feng, M., Sexton, J. O., Channan, S., & Townshend, J. R. (2016). A global, high-resolution (30-m) inland water body dataset for 2000: First results of a topographic–spectral classification algorithm. *International Journal of Digital Earth*, *9*(2), 113–133. https://doi.org/10.1080/17538947.2015.1026420

Fluet-Chouinard, E., Lehner, B., Rebelo, L.-M., Papa, F., & Hamilton, S. K. (2015). Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data. *Remote Sensing of Environment*, *158*, 348–361. https://doi.org/10.1016/j.rse.2014.10.015

Friedl, M. A., McIver, D. K., Hodges, J. C., Zhang, X., Muchoney, D., Strahler, A. H., et al. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, *83*(1–2), 287–302. https://doi.org/10.1016/S0034-4257(02)00078-0

Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., et al. (2013). Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing*, *34*(7), 2607–2654. https://doi.org/10.1080/01431161.2012.748992

Hansen, M., DeFries, R., Townshend, J. R., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, *21*(6–7), 1331–1364. https://doi.org/10.1080/014311600210209

Howells, M., Hermann, S., Welsch, M., Bazilian, M., Segerström, R., Alfstad, T., et al. (2013). Integrated analysis of climate change, land-use, energy and water strategies. *Nature Climate Change*, *3*(7), 621–626. https://doi.org/10.1038/nclimate1789

Ji, L., Gong, P., Geng, X., & Zhao, Y. (2015). Improving the accuracy of the water surface cover type in the 30 m FROM-GLC product. *Remote Sensing*, *7*(10), 13,507–13,527. https://doi.org/10.3390/rs71013507

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., et al. (2010). Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, *467*(7318), 951–954. https://doi.org/10.1038/nature09396

Klein, I., Dietz, A., Gessner, U., Dech, S., & Kuenzer, C. (2015). Results of the Global WaterPack: A novel product to assess inland water body dynamics on a daily basis. *Remote Sensing Letters*, *6*(1), 78–87. https://doi.org/10.1080/2150704X.2014.1002945

Klein, I., Gessner, U., Dietz, A. J., & Kuenzer, C. (2017). Global WaterPack—A 250 m resolution dataset revealing the daily dynamics of global inland water bodies. *Remote Sensing of Environment*, *198*, 345–362. https://doi.org/10.1016/j.rse.2017.06.045

Lehner, B., & Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, *296*(1), 1–22.

Li, C., Gong, P., Wang, J., Zhu, Z., Biging, G. S., Yuan, C., et al. (2017). The first all-season sample set for mapping global land cover with Landsat-8 data. *Science Bulletin*, *62*(7), 508–515. https://doi.org/10.1016/j.scib.2017.03.011

Li, W., & Gong, P. (2016). Continuous monitoring of coastline dynamics in western Florida with a 30-year time series of Landsat imagery. *Remote Sensing of Environment*, *179*, 196–209. https://doi.org/10.1016/j.rse.2016.03.031

Liao, A., Chen, L., Chen, J., He, C., Cao, X., Chen, J., et al. (2014). High-resolution remote sensing mapping of global land water. *Science China Earth Sciences*, *57*(10), 2305–2316. https://doi.org/10.1007/s11430-014-4918-0

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., & Merchant, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, *21*(6–7), 1303–1330. https://doi.org/10.1080/014311600210191

Papa, F., Prigent, C., Aires, F., Jimenez, C., Rossow, W., & Matthews, E. (2010). Interannual variability of surface water extent at the global scale, 1993–2004. *Journal of Geophysical Research*, *115*, D12111. https://doi.org/10.1029/2009JD012674

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422. https://doi.org/10.1038/nature20584

Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R., Scholes, R. J., et al. (2013). Essential biodiversity variables. *Science*, *339*(6117), 277–278. https://doi.org/10.1126/science.1229931

Pfister, S., Bayer, P., Koehler, A., & Hellweg, S. (2011). Environmental impacts of water use in global crop production: Hotspots and trade-offs with land use. *Environmental Science & Technology*, *45*(13), 5761–5768. https://doi.org/10.1021/es1041755

Prigent, C., Papa, F., Aires, F., Jimenez, C., Rossow, W., & Matthews, E. (2012). Changes in land surface water dynamics since the 1990s and relation to population pressure. *Geophysical Research Letters*, *39*, L08403. https://doi.org/10.1029/2012GL051276

Raymond, P. A., Hartmann, J., Lauerwald, R., Sobek, S., Mcdonald, C., Hoover, M., et al. (2013). Global carbon dioxide emissions from inland waters. *Nature*, *503*(7476), 355–359. https://doi.org/10.1038/nature12760

Schuur, E. A., Mcguire, A. D., Schädel, C., Grosse, G., Harden, J. W., Hayes, D. J., Hugelius, G., et al. (2015). Climate change and the permafrost carbon feedback. *Nature*, *520*(7546), 171–179. https://doi.org/10.1038/nature14338

Seekell, D. A., Carr, J. A., Gudasz, C., & Karlsson, J. (2014). Upscaling carbon dioxide emissions from lakes. *Geophysical Research Letters*, *41*, 7555–7559. https://doi.org/10.1002/2014GL061824

Sun, F., Zhao, Y., Gong, P., Ma, R., & Dai, Y. (2014). Monitoring dynamic changes of global land cover types: Fluctuations of major lakes in China every 8 days during 2000–2010. *Chinese Science Bulletin*, *59*(2), 171–189. https://doi.org/10.1007/s11434-013-0045-0

Verpoorter, C., Kutser, T., Seekell, D. A., & Tranvik, L. J. (2014). A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, *41*, 6396–6402. https://doi.org/10.1002/2014GL060641

Wada, Y., Wisser, D., & Bierkens, M. (2014). Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth System Dynamics*, *5*(1), 15–40. https://doi.org/10.5194/esd-5-15-2014

Wang, J., Zhao, Y., Li, C., Yu, L., Liu, D., & Gong, P. (2015). Mapping global land cover in 2001 and 2010 with spatial-temporal consistency at 250 m resolution. *ISPRS Journal of Photogrammetry and Remote Sensing*, *103*, 38–47. https://doi.org/10.1016/j.isprsjprs.2014.03.007

Yang, J., Gong, P., Fu, R., Zhang, M., Chen, J., Liang, S., et al. (2013). The role of satellite remote sensing in climate change studies. *Nature Climate Change*, *3*(10), 875–883. https://doi.org/10.1038/nclimate1908

Zhao, Y., Gong, P., Yu, L., Hu, L., Li, X., Li, C., et al. (2014). Towards a common validation sample set for global land-cover mapping. *International Journal of Remote Sensing*, *35*(13), 4795–4814. https://doi.org/10.1080/01431161.2014.930202

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, *118*(0), 83–94. https://doi.org/10.1016/j.rse.2011.10.028