Supplementary Information Appendix

Divergent trends of open surface water body area in the contiguous United States from 1984 to 2016

Zhenhua Zou, Xiangming Xiao, Jinwei Dong, Yuanwei Qin, Russell B. Doughty, Michael A. Menarguez, Geli Zhang, Jie Wang

Contents

Text S1. Thresholds to classify year-long, seasonal and ephemeral water bodies

Year-long water bodies should have an annual water frequency of 100% when we observe them from the ground. However, when we detect water using observations from space-borne satellites (e.g., Landsat, ~795 km above the ground), the water frequency of year-long water bodies could be decreased by many factors. First, cloud cover and shadow can affect observations. Some of the clouds, especially optically thin clouds, have a chance of being omitted by the CFMask cloud-screening algorithm (1). Thus, the omitted clouds over the surface water body will be classified as non-water and eventually reduce the annual water frequency of year-long water pixels. Second, although the L1T (terrain corrected, mismatch < 12m) surface reflectance product used for water detection are suitable for pixel-level time series analysis, for the year-long water pixels close to shorelines, the small geometric mismatch may introduce some mixed pixels into the time series data in a year and reduce the annual water frequency. Third, the surface reflectance data are affected by atmospheric correction algorithm. For example, the Landsat 8 SR algorithm may introduce some artifacts over certain geographic areas, including inland water bodies, area of high relief, and areas with high aerosols (1-3). These artifacts in the surface reflectance data may also reduce the annual water frequency of year-long water pixels.

There are a couple approaches to deal with this issue. One approach is used in the "permanent water body map" by the JRC group. They used the frequency with which a pixel occupies the unequivocal portion of the water hull to estimate the likelihood of it actually being water (4). Specifically, "If a pixel sits unequivocally within a water hull for some of the time, then there is a high likelihood it will actually be water even if it occasionally occupies a hull where overlap occurs with other cover types." Their approach manually adjusts the annual water frequency, which may increase the water frequency of some pixels to 100%. In comparison, our

approach is to keep the original frequency values and assume a maximum error range (25%) in classifying year-long water bodies. In other words, we used 75% as a threshold to classify the pixels that have water most of the time in a year as year-long water pixels. As shown in Fig. 1E, the surface water body area within CONUS and its inter-annual variations using 75% threshold value are quite similar with those from JRC permanent water body area. In addition, we reported detailed analyses on the annual frequency of surface water body in Oklahoma (5), and the 75% threshold value was appropriate for year-long water bodies. Also, we used 5% as a threshold to separate ephemeral water bodies (annual water frequency < 0.05) from seasonal water bodies ($0.05 \leq$ annual water frequency < 0.75).

Text S2. Surface water withdrawal data interpolation

The statewide surface water withdrawal data across the US were reported every five years by the USGS (6). We assumed the changing speed of surface water withdrawals within the 5 year-interval was linear with a constant coefficient. Thus, the water withdrawal for the next year equaled the water withdrawal of the current year plus one fifth of the total change within the five years. For example, the water withdrawals in 1986 can be calculated:

$$W_{1986} = W_{1985} + \frac{(W_{1990} - W_{1985})}{5} \qquad (1)$$

Where W_{1985}, W_{1986}, and W_{1990} were the statewide surface water withdrawal in 1985, 1986, and 1990, respectively. The surface water interpolation after 2010 uses the change between 2005 and 2010.

Text S3. Definitions of water and vegetation indices

The Modified Normalized Difference Water Index (mNDWI), Normalized Difference Vegetation Index (NDVI), and Enhanced Vegetation Index (EVI) were calculated for water detection using the equations below:

$$\text{mNDWI} = \frac{(\rho_{green} - \rho_{SWIR1})}{(\rho_{green} + \rho_{SWIR1})} \quad (1)$$

$$\text{NDVI} = \frac{(\rho_{NIR} - \rho_{Red})}{(\rho_{NIR} + \rho_{Red})} \quad (2)$$

$$\text{EVI} = 2.5 \times \frac{(\rho_{NIR} - \rho_{Red})}{(1 + \rho_{NIR} + 6\rho_{Red} - 7.5\rho_{blue})} \quad (3)$$

Where ρ_{blue}, ρ_{green}, ρ_{red}, ρ_{NIR}, and ρ_{SWIR1} are the surface reflectance values of Landsat blue band (0.45–0.52), green band (0.52–0.60), red band (0.63–0.69), near-infrared band (0.77–0.90), and shortwave infrared band (1.55–1.75) μm, respectively.

Text S4. Algorithm development

Thirty-one out of 459 Landsat tiles that overlap with the CONUS were selected (Fig. S5). For each tile, 1–2 rectangle sampling plots were randomly selected in the locations that have both a Landsat image and a high-resolution Google Earth image within a time window of about one month. Each sampling plot should consist of ~50% water and ~50% other land cover pixels. The total sampling pixels of each tile is ~ 40,000. Altogether, 32 sampling plots were selected, of which 14, 9 and 9 were from Landsat 5, 7 and 8, respectively (Fig. S5A). In terms of time distribution, the sampling plots were selected across the time range of each satellite (Fig. S5B). In terms of land cover, 18 had the major land cover types of vegetation and water, 7 had built-up land and water, 5 had cropland and water, and 2 had bare land and water (Fig. S5C). In terms of terrain, 5 sampling plots were in mountainous areas while 27 were in relatively flat areas (Fig. S5D). For each rectangle sampling plot, all water features and non-water features were visually

delineated referring to the high-resolution image in Google Earth. The 32 sampling plots contained ~1.26 million pixels, of which 368,850 were water and 886,496 were non-water according to visual delineation.

Water and non-water frequency curves cross each other around 0 in the distribution of mNDWI-EVI values of 1.26 million sampling pixels across the US (Fig. S6A). 97.36% of the water pixels show mNDWI > EVI while 99.29% of the non-water pixels show mNDWI < EVI. Thus, mNDWI >EVI is a good criterion to detect water. Also, 93.2% of the water pixels show mNDWI > NDVI while 99.44% of the non-water pixels show mNDWI < NDVI (Fig. S6B). Therefore, mNDWI > NDVI can be used as a supplementary criterion to separate water from non-water pixels. Furthermore, 98.4% of the water pixels show EVI < 0.1. Thus, EVI < 0.1 can be used to exclude mixed pixels of water and vegetation. The final water detection formula is ((mNDWI > EVI or mNDWI > NDVI) and EVI < 0.1). This formula can be divided into two parts, (mNDWI > EVI and EVI < 0.1) and (mNDWI > NDVI and EVI < 0.1), whose scatter density plots of 1.26 million sampling pixels were shown in Fig. S6D and E, respectively.

Text S5. Algorithm verification

The annual water body maps were based on water frequency, which made use of all good observations across the year. However, it is very difficult to directly verify the annual water frequency maps. Thus, we verify the algorithms instead. According to the water and non-water boundaries from the National Land Cover Database 2011 (7), 1600 water and 1600 non-water sampling points were randomly selected within the CONUS, among which 200 were selected in Great Lake Region of the 2-digit Hydrologic Units Code (HUC-2), while 3000 were selected in the rest of the CONUS (Fig. S7). In the Google Earth Engine platform, all available high-resolution images from US National Agriculture Imagery Program (NAIP) at the specific

location for each point were selected. For each NAIP image, one Landsat image was randomly selected from all Landsat images (TM, ETM+, OLI) acquired within ±5 days of the NAIP image. Each sampling point had up to 10 pairs of NAIP and Landsat images distributing across 2003–2016. One pair of those images was randomly selected. Water detection algorithms were performed on the Landsat image while the NAIP image was used as ground reference data to verify the water detection. At each sampling point location, 4 adjacent Landsat pixels were selected. For each Landsat pixel, its boundary was added to the NAIP image and then record the water covering percentage in reference to the NAIP image and record whether this pixel was classified as water or not in our algorithms. Among the 3200 sampling points, 3197 were able to find adequate NAIP and Landsat data for verification. Out of the 3197 Landsat images, 1623 were from Landsat 5 (51%), 899 were from Landsat 7 (28%), and 675 were from Landsat 8 (21%). The selected images distributed evenly across 2003–2016. The 3197 sampling points correspond to 12,788 Landsat pixels. According to NAIP image, there were 4767 pure water pixels (37%), 7084 pure non-water pixels (56%), and 937 mixed pixels of water and other land cover types (7%).

The confusion matrix of water detection at 11851 pure water and non-water pixels is shown in Table S1. The producer accuracies of water and non-water were 93.39% and 99.28%, respectively. The overall accuracy is 96.91% with a kappa coefficient equal to 0.94 (Table S1).

Among the 937 mixed pixels, 843 contained water and vegetation (90%), 77 contained water and bare land or sand (8%), and 17 contained water and built-up area (2%). The water detection rates of mixed pixels with various non-water coverage were shown in Fig. S8. Generally, as water coverage increased, water detection rates (pixel percentage classified as water) increased. The water detection rates of mixed pixels of water and vegetation were all very low because

water detection had to meet the criteria, “EVI < 0.1”. The unstable trends of water detection rates of mixed pixels of water and bare land, sand, and built-up area were probably caused by limited sampling pixels.

Fig. S1. Study area

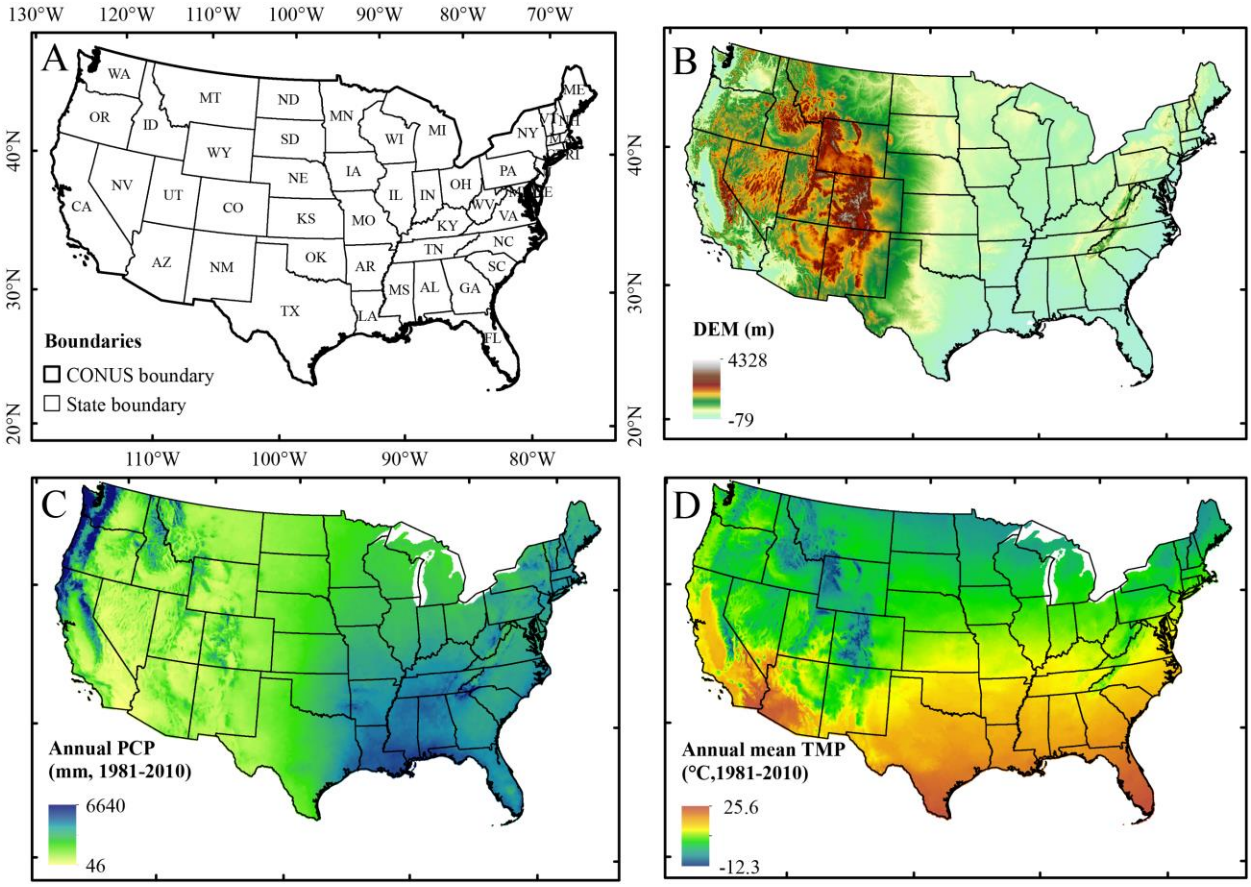

Fig. S1. Study area. (A) State and CONUS boundary. (B) Digital elevation model (DEM). (C) Annual precipitation (PCP) during 1981–2010. (D) Annual mean temperature (TMP) during 1981–2010. DEM is from USGS EarthExplorer (http://earthexplorer.usgs.gov/), while the precipitation and temperature are from the Parameter-elevation Relationships on Independent Slopes Model (PRISM) Climate Group, Oregon State University (http://prism.oregonstate.edu/).

Fig. S2. Multiple stepwise regression

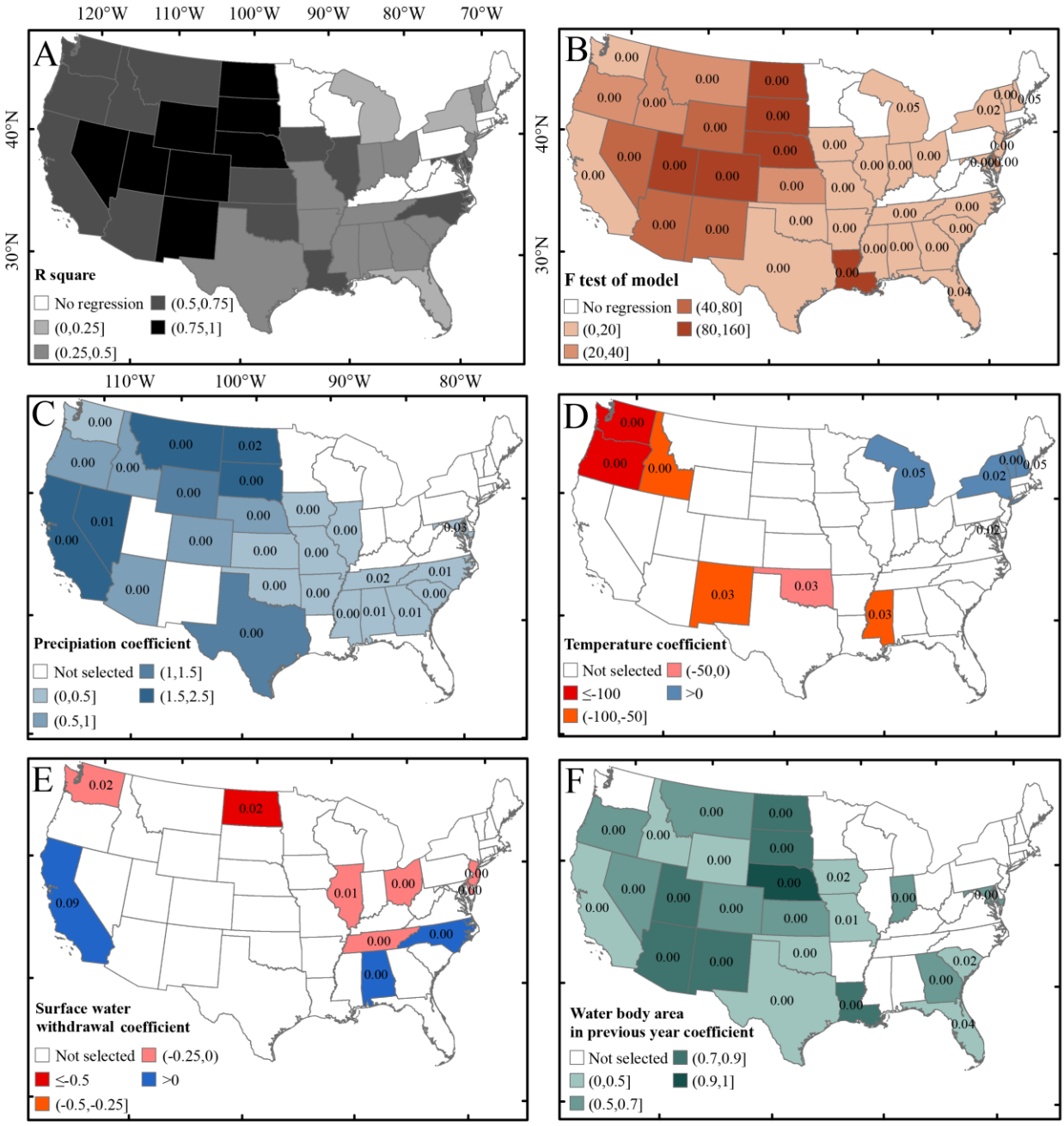

Fig. S2. Multiple stepwise regression models. (A) R square of regression models. (B) F test and p value of regression models (the filled color is F, while the number is p value). Coefficient (slope) and p value of factor annual precipitation (C), annual average temperature (D), surface water withdrawal (E), and year-long water body area in the previous year (F) (the filled color is coefficient, while the number is p value).

Fig. S3. Spatial and temporal distribution of major dams

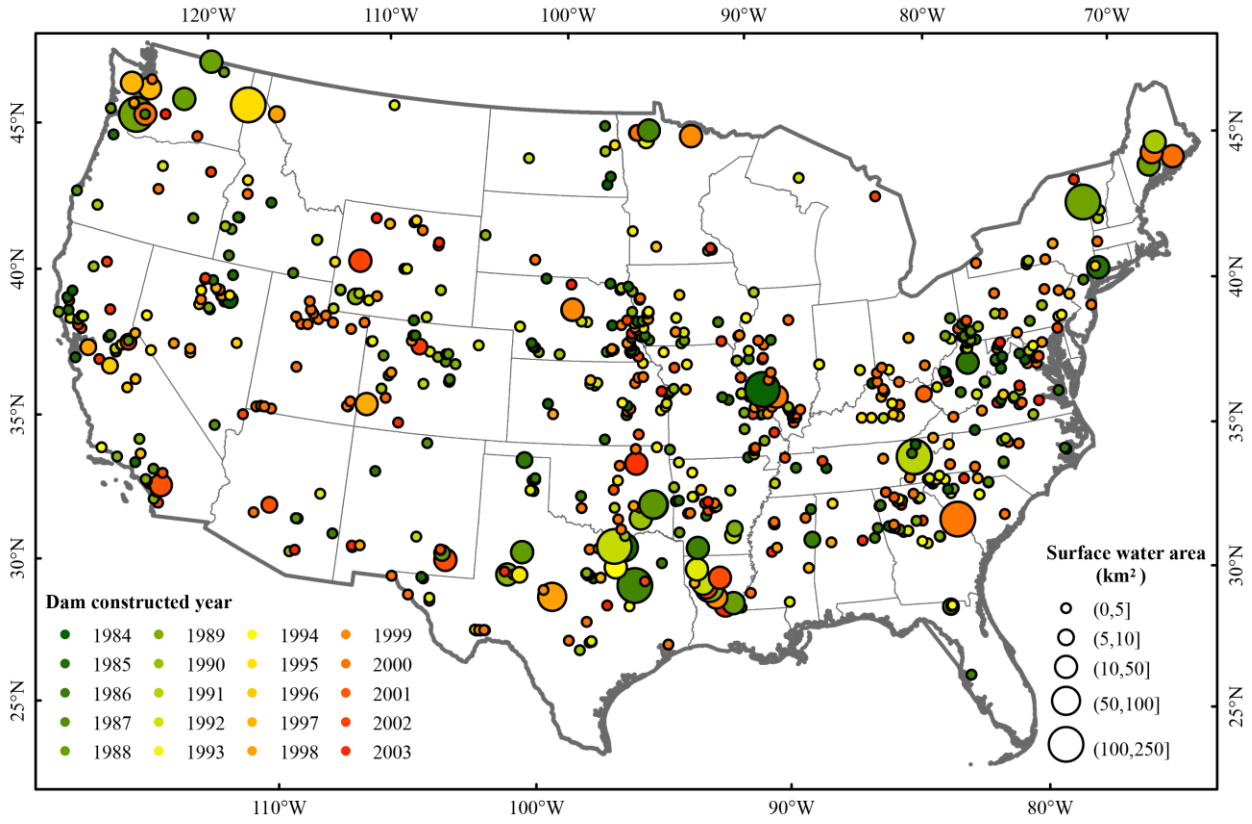

Fig. S3. Spatial and temporal distribution of major dams constructed during 1984–2003 within the CONUS (8) (the circle filled color represents dam constructed year, while the circle size represents the surface water body area of the impoundment at its normal retention level).

Fig. S4. Landsat data

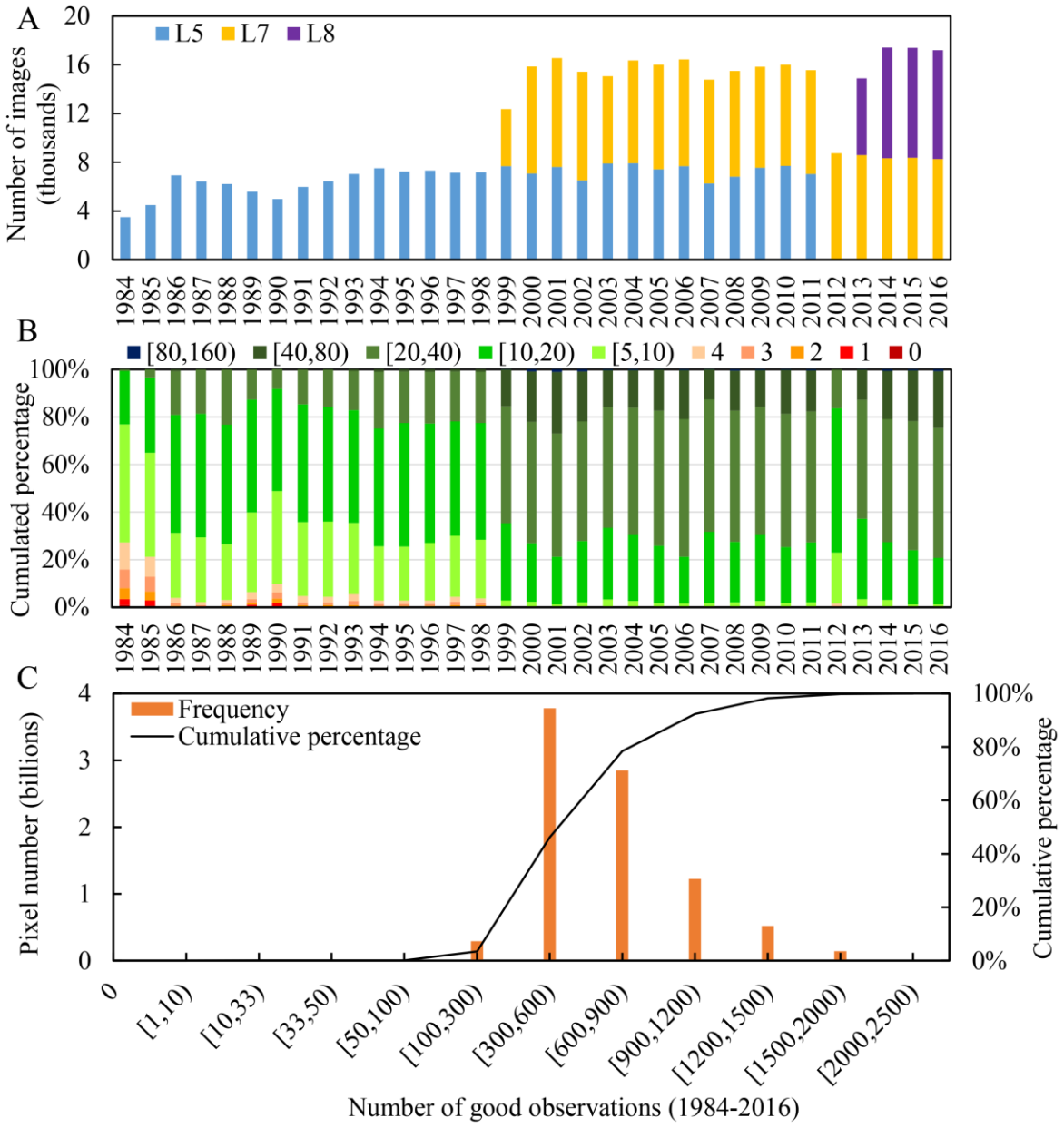

Fig. S4. Landsat data availability within the CONUS during 1984-2016. (A) Number of Landsat 5, 7 and 8 images in each year. (B) Cumulative percentage of pixels with good observations of 0, 1, 2, 3, 4, [5, 10), [10, 20), [20, 40), [40, 80), and [80, 160), respectively. (C) Number of pixels with good observations of 0, [1, 10), [10, 33), [33, 50), [50, 100), [100, 300), [300, 600), [600, 900), [900, 1200), [1200, 1500), [1500, 2000), and [2000, 2500), respectively, in 33 years.

Fig. S5. Sample plot distribution

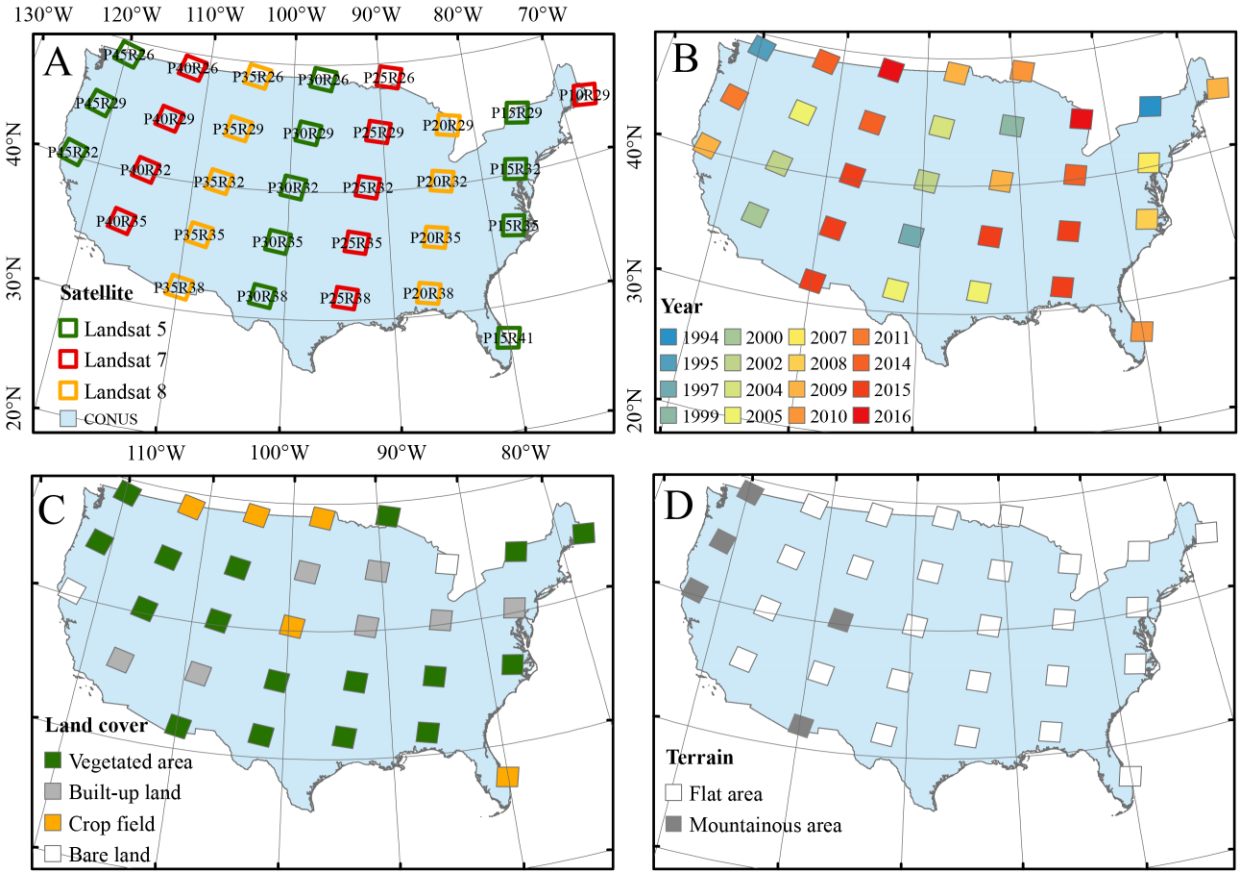

Fig. S5. Landsat tile (path/row) and sampling plot distribution. (A) Distribution of selected Landsat tiles (path/row) by satellites. (B) Distribution of years of sampling plots. (C) Distribution of land cover types of sampling plots. (D) Distribution of terrain types of sampling plots. These sampling plots were used for visual delineation of water body and non-water features, which were further used to analyze their spectral characteristics (see Text S4 for more information).

Fig. S6. Distribution of 1.26 million sampling pixels.

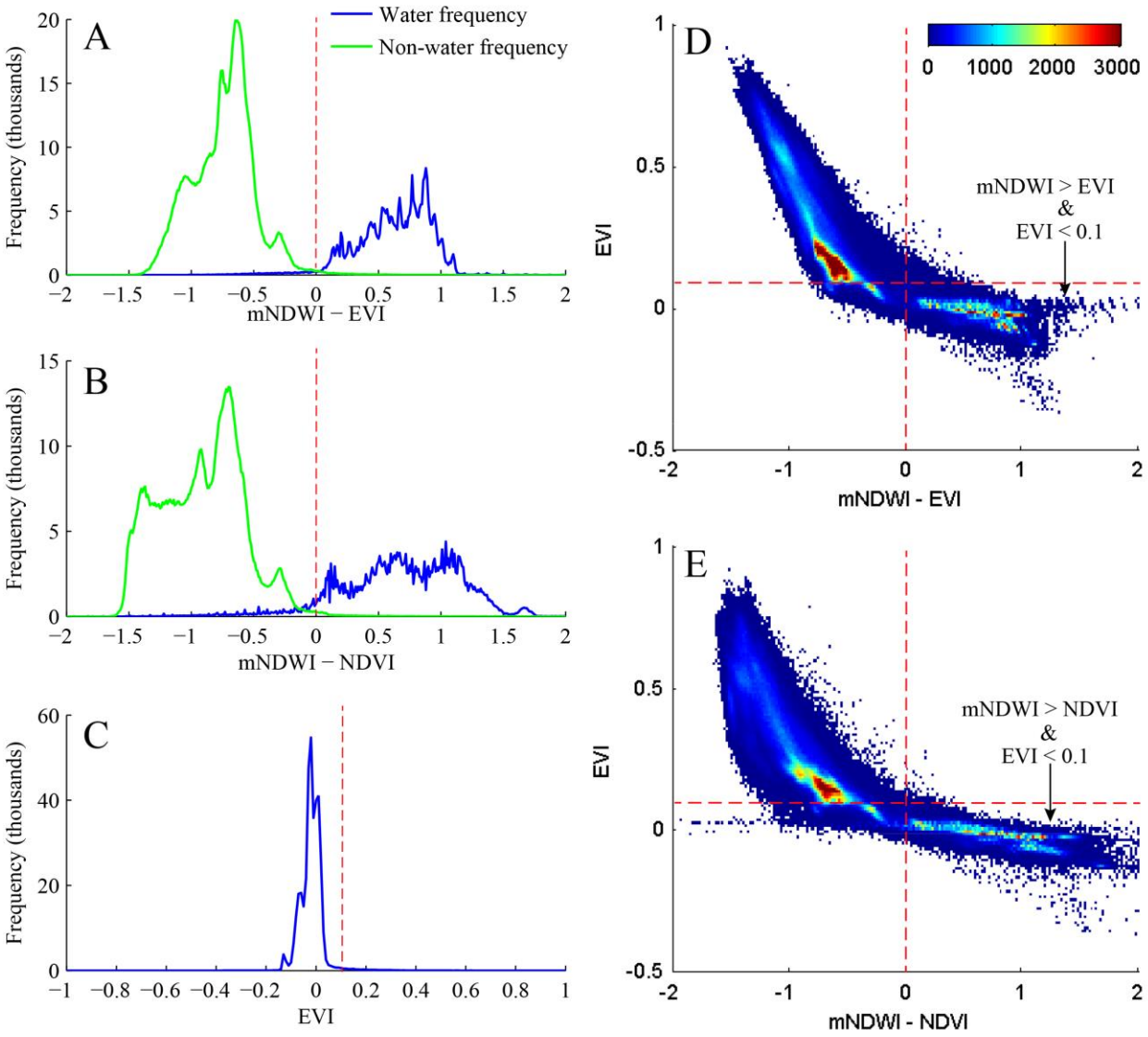

Fig. S6. Spectral characteristics of 1.26 million sampling pixels. (A) Frequency distribution of (mNDWI-EVI) for water and non-water sampling pixels. (B) Frequency distribution of (mNDWI-NDVI) for water and non-water sampling pixels. (C) Frequency distribution of EVI for water sampling pixels. (D) Scatter density plots of EVI VS (mNDWI-EVI) of all sampling pixels. (E) Scatter density plots of EVI VS (mNDWI-NDVI) of all sampling pixels.

Fig. S7. Sampling points for verification

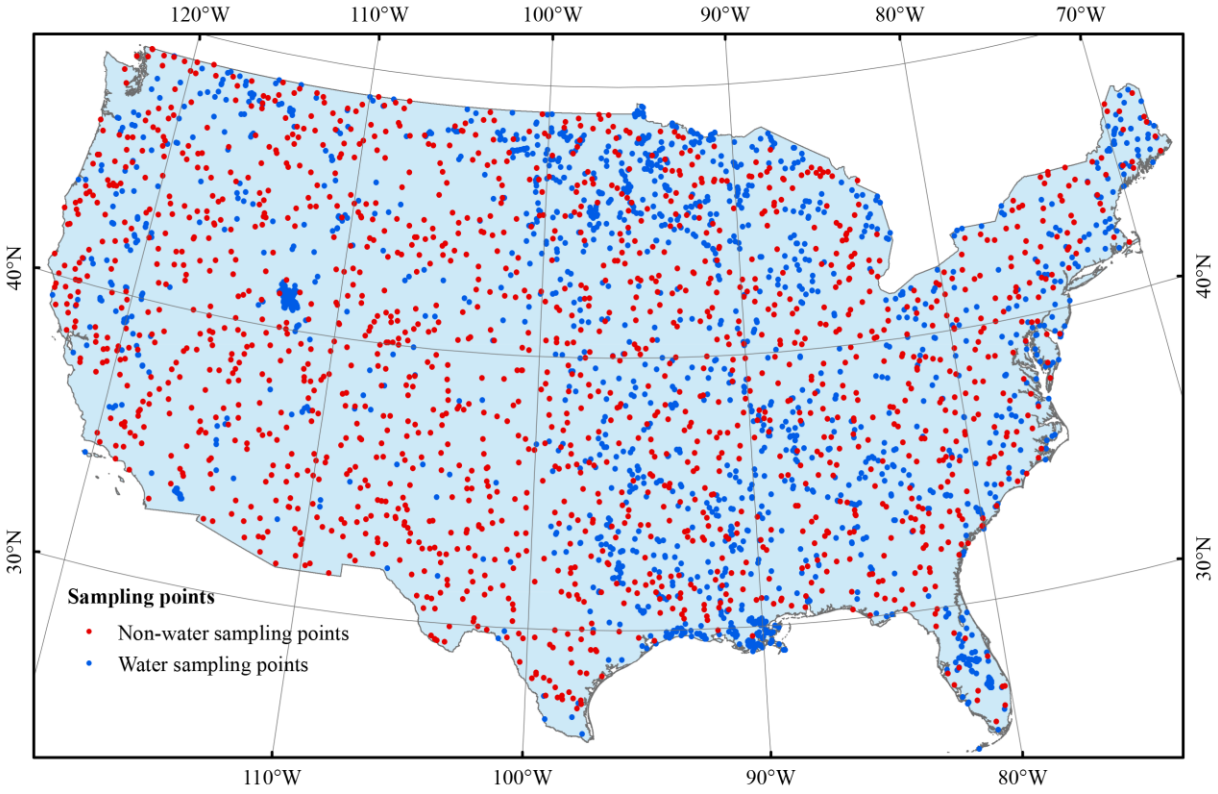

Fig. S7. Spatial distribution of 3200 verification sampling points that were used to select ~12,000 sampling pixels for accuracy assessment (see Text S5 for more information).

Fig. S8. Mixed pixel water detection rate.

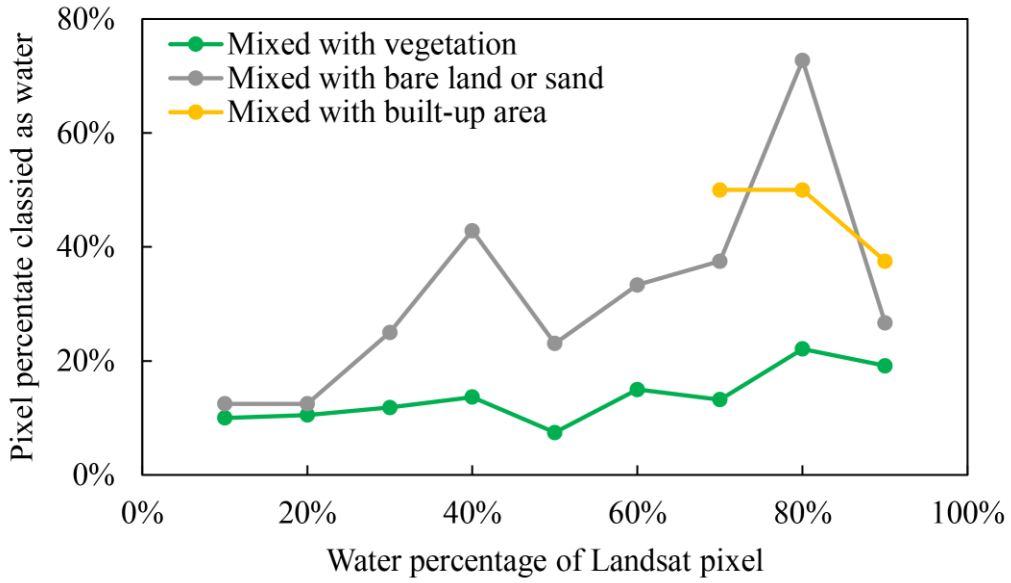

Fig. S8. Mixed pixel water detection rate.

Table S1. Confusion matrix for evaluating water detection algorithms

Classification	Ground Reference		Sum	User accuracy (%)
	Water	Non-water		
Water	4452	51	4503	98.87%
Non-water	315	7033	7348	95.71%
Sum	4767	7084	11851	OA=96.91%
Producer accuracy (%)	93.39%	99.28%		Kappa = 0.94

Table S2 Variance inflation factor (VIF) of all predictor variables.

State	V1	V2	V3	V4	State	V1	V2	V3	V4
Arkansas	1.1	1.1	1.1	1.1	Louisiana	1.2	1.1	1.7	1.7
Arizona	1.7	1.8	1.1	1.2	Mississippi	1.0	1.1	1.2	1.2
Colorado	1.2	1.3	1.2	1.1	South Carolina	1.2	1.1	1.2	1.1
Iowa	1.0	1.1	1.0	1.1	California	1.1	1.4	2.2	1.8
Illinois	1.1	1.2	1.3	1.3	Connecticut	1.1	1.1	1.2	1.0
Indiana	1.1	1.2	1.0	1.2	Washington DC	1.9	1.6	1.3	1.5
Kansas	1.2	1.3	1.1	1.2	Delaware	1.2	1.5	3.0	2.5
Missouri	1.1	1.2	1.1	1.1	Kentucky	1.1	1.3	1.1	1.3
Nebraska	1.3	1.2	1.1	1.2	Massachusetts	1.2	1.4	1.4	1.1
New Mexico	1.6	1.9	3.1	3.3	Maryland	1.2	1.2	1.4	1.4
Nevada	1.1	1.1	2.6	2.5	North Carolina	1.0	1.1	1.9	1.9
Oklahoma	1.2	1.3	1.7	1.5	New Jersey	1.1	1.1	1.4	1.4
Tennessee	1.2	1.3	1.3	1.6	New York	1.0	1.1	1.3	1.1
Utah	1.2	1.3	1.1	1.3	Ohio	1.0	1.2	1.4	1.6
West Virginia	1.0	1.1	1.2	1.2	Pennsylvania	1.0	1.1	1.1	1.1
Idaho	1.1	1.1	1.3	1.2	Rhode Island	1.4	1.2	1.2	1.5
Montana	1.5	1.3	1.4	1.3	Virginia	1.0	1.1	1.0	1.0
North Dakota	1.3	1.4	4.0	3.7	Wisconsin	1.2	1.1	1.2	1.1
South Dakota	1.2	1.3	1.0	1.1	Michigan	1.1	1.4	1.2	1.2
Vermont	1.2	1.1	1.1	1.1	Maine	1.3	1.4	1.3	1.1
Wyoming	1.1	1.3	1.1	1.0	Minnesota	1.0	1.1	1.1	1.0
Florida	1.4	1.2	1.2	1.6	New Hampshire	1.3	1.1	1.1	1.3
Texas	1.2	1.1	1.2	1.2	Oregon	1.1	1.2	1.2	1.4
Alabama	1.2	1.2	1.6	1.6	Washington	1.1	1.2	1.1	1.2
Georgia	1.1	1.3	1.2	1.3					

Multiple stepwise regression models between year-long water body area and four predictor variables by states (Y=f (V1, V2, V3, V4). V1: annual precipitation (mm), V2: annual average temperature (℃), V3: annual total surface water withdraw (million gallons per day), V4: year-long water body area of the previous year (km^2).

References

1. USGS (2017) Landsat 4-7 Climate Data Record (CDR) Surface Reflectance Product Guide (Department of the Interior US Geological Survey, Washington, DC), Version 7.2.
2. Google Earth Engine (2017) *USGS Landsat 5/7/8 Surface Reflectance (pre-Collection)*. (Google Earth Engine, Mountain View, CA).
3. USGS (2017) Landsat 8 Surface Reflectance Code (LaSRC) Product Guide (Department of the Interior US Geological Survey, Washington, DC), Version 3.4
4. Pekel JF, Cottam A, Gorelick N, Belward AS (2016) High-resolution mapping of global surface water and its long-term changes. *Nature* 540:418-422.
5. Zou Z, et al. (2017) Continued decrease of open surface water body area in Oklahoma during 1984-2015. *Sci Total Environ* 595:451-460.
6. US Geological Survey (2010) *Estimated Use of Water in the United States in 1985, 1990, 1995, 2000, 2005 and 2010* (USGS, Reston, VA).
7. Homer C, et al. (2015) Completion of the 2011 National Land Cover Database for the Conterminous United States - Representing a Decade of Land Cover Change Information. *Photogramm Eng Remote Sens* 81:345-354.
8. National Atlas of the United States (2006) Major Dams of the United States (National Atlas of the United States, Reston, VA).