UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MULTI-SCALE REMOTE SENSING OF OPEN SURFACE WATER BODY AREA AND QUALITY

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

ZHENHUA ZOU
Norman, Oklahoma
2019

MULTI-SCALE REMOTE SENSING OF OPEN SURFACE WATER BODY AREA AND QUALITY

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF MICROBIOLOGY AND PLANT BIOLOGY

BY

Dr. Xiangming Xiao, Chair

Dr. Jeffrey Basara

Dr. Yang Hong

Dr. Heather McCarthy

Dr. Lara Souza

This dissertation is dedicated to my parents, my wife, and my daughter for their unconditional love and support.

Acknowledgements

Great thanks to my academic advisor Dr. Xiangming Xiao for his patient guidance, valuable suggestions, and unconditional support during my Ph.D. program. With profound knowledge, he shows me the direction when I am confused, solves the problems that I could not overcome, and encourages me when I am frustrated. I am also appreciated for the freedom he gives me in my research so that I could try all kinds of ideas and strategies to solve a problem. In this way, I make mistakes, learn from my mistakes, and develop the ability little by little to carry out independent research, which I think is the most valuable experience in my Ph.D. program.

I would like to thank my committee members Dr. Jeffrey Basara, Dr. Yang Hong, Dr. Heather McCarthy, and Dr. Lara Souza for their unconditional service, help, and support in my Ph.D. program. Their insightful commons and suggestions have guided me through the difficulties and challenges in my Ph.D. journey.

I would also like to thank my colleagues in EOMF lab. They are nice collaborators and good friends. They have made my research easier and my life happier.

Last but not least, great thanks to my beloved family.

Table of Contents

List of Figures

List of Tables

Abstract

Water is one of the most important resources for life. Climate change and climate variability have caused dramatic variations and significant trends in surface water resources, while global population growth and increased food demand have greatly stressed and modified global surface water systems. These changes in surface water resources have huge consequences to human society, natural environment, and global biodiversity. Landsat satellites have scanned the entire earth in every 16 days since the 1980s. The historical information of surface water body spatial distribution, temporal variation, and multi-decadal trends documented in remote sensing images can aid in water resource research, planning, and management, yet it is not well explored. This dissertation aims to develop algorithms and generate open surface water body maps at state, national, and global scales. Based on these maps, the interannual variations and long-term trends of surface water body area were analyzed while their climatic and anthropogenic drivers were examined. The joint analysis of both surface water body area and land water storage was carried out to explore the consistency and divergence between surface and land water resource dynamics. The potential of satellite images in water chlorophyll-a concentration estimation was also evaluated. Chapter 2 used ~16,000 Landsat images to analyze surface water body dynamics in Oklahoma and found significant decreasing trends in both surface water body area (the maximum, year-long, seasonal, and average water body area) and water body number (maximum and year-long water body numbers) during 1984–2015. The decrease of water body area was mainly attributed to the shrinking of large water bodies (>1 km^2) while the decrease in water body number was mainly caused by the vanishing of some small water bodies. Smaller water bodies have a higher risk of

drying up under climate-warming scenarios. Chapter 3 used ~ 370,000 Landsat images and the Gravity Recovery and Climate Experiment (GRACE) land water storage data to analyze changes in surface water area and groundwater across the contiguous US (CONUS) during 1984–2016. Divergent trends of surface water area were found across the CONUS with water-poor regions of the Southwest and Northwest US getting poorer, while the water-rich regions of the Southeast US and far north Great Plains getting richer. In the 2012-2014 prolonged droughts, surface water body shrinkage had led to massive groundwater mining and the rapid decline of land water storage in California and the Southern Great Plains. Chapter 4 used ~3.8 million Landsat images and GRACE land water storage data to analyze surface water area and land water storage jointly during 1984–2017 at 0.01° grid cells, 0.5° grid cells, and 5° tiles across the globe. About 8.5 million 0.01° grid cells had significant increasing or decreasing trends in surface water area over the past decades, forming interesting spatial patterns in northern Greenland, Tibetan Plateau, western US, the Great Lakes, Gulf of Bothnia, central South America, etc. The interannual variations, magnitude of variability, and the multi-decadal trends of regional surface water area were analyzed and visualized at 5° tiles. Divergent trends between land water storage and regional surface water area occurred in Greenland, China, the Indus Basin, and central Africa, mainly driven by climate and anthropogenic activities. Chapter 5 used ~10,000 chlorophyll-a field measurements to evaluate the potential of Landsat 5/7/8 and Sentinel 2 satellite images in water chlorophyll-a content estimation. Regression models of Landsat data have different performance in various water sampling sites and water bodies across Oklahoma, with relatively good performance in Eufaula Lake, Keystone Lake, Copan Lake, Hugo Lake, Foss Reservoir,

and Atoka Reservoir. The brightness temperature band of Landsat satellites showed great potential in chlorophyll-a estimation, which indicated that temperature was among the most important factors of algal blooms in Oklahoma. The Red Edge 2 band of Sentinel 2 Satellite also showed great potential in chlorophyll-a estimation among different water sampling sites and water bodies across Oklahoma. The findings in this dissertation can be used in water resource research, planning, and management in coping with water scarcity and food security associated with climate change and population growth.

Chapter1: Introduction

1.1 Research background

Terrestrial surface water resources are critical for human society, natural environment, and global biodiversity (Pekel et al., 2016; Vorosmarty et al., 2010). Climate change and increased climate variability have strong impacts on surface water resources (Aherne et al., 2006; Ferguson and Maxwell, 2012; Tulbure et al., 2016), including dramatic spatial and temporal variability (Hall et al., 2014; Mercier et al., 2002), and severe droughts and floods (Melillo et al., 2014), which have a variety of consequences on human societies and ecosystems (Bates et al., 2008; Brown and Lall, 2006). Global population increased from about 5 billion in 1984 to 7.6 billion in 2017 (DESA, 2017). Population growth and increased food demand have also stressed and modified global surface water systems through water diversion for irrigation, dam construction, and land reclamation (Vorosmarty et al., 2010). Agriculture development and domestic sewage discharge have contaminated and eutrophicate surface water bodies, leading to algal blooms and the decrease of water quality (Gons et al., 2008; Matsushita et al., 2015). Mapping terrestrial open surface water bodies is critically important to water resource management, agricultural and industrial production, flood and drought assessment, and aquatic ecosystem protection (Wood et al., 2011). Analysis of the interannual variation of open surface water body areas can reveal the stability of water resources while examining the driving factors of water body variability can reveal the reasons behind the significant trends of surface water body area (Zou et al., 2017; Zou et al., 2018). Chlorophyll-a concentration is one of the most commonly used indicators of water quality, including trophic status, clarity, and algal biomass (Matsushita et al., 2015).

Chlorophyll-a estimation using remote sensing images can serve as an alternative and supplement to the expensive field chlorophyll-a sampling and measurement (LaPotin et al., 2001; Matsushita et al., 2015).

Satellite-based remote sensing images have been widely used to identify and map open surface water bodies over the past decades (Boland, 1976; Heimhuber et al., 2017; Pekel et al., 2016; Proulx et al., 2013; Tao et al., 2015; Xu, 2006). Many water detection algorithms were based on different spectral water indices (Bhagat and Sonawane, 2011; Crist, 1985; Feyisa et al., 2014; Fisher et al., 2016; McFeeters, 1996; Wang et al., 2015; Xiao et al., 2002; Xu, 2006). The modified Normalized Difference Water Index (mNDWI) is one of the most widely used water indices because of its good performance in different landscapes (Ji et al., 2009; Yamazaki et al., 2015). As United States Geological Survey (USGS) made all Landsat images available to the public, mapping of open surface water bodies was carried out in continental (Mueller et al., 2016; Tulbure et al., 2016) and global scales (M Feng et al., 2016; Pekel et al., 2016; Verpoorter et al., 2014; Yamazaki et al., 2015). Google Earth Engine (GEE) cloud computing platform provides satellite image archives, user-friendly programing interface, and free computing power for the public to carry out data analyzing. It opens an unprecedented opportunity for the remote sensing community to explore the potential of water body mapping.

Many water body maps were generated using single-temporal or multi-temporal images gathered at a specific time or period of the year, normally the wet season (Liu et al., 2013; Schaffer-Smith et al., 2017; Yamazaki et al., 2015). However, because of the seasonal variations of water bodies, these water body maps can only reveal a snapshot of the dynamic water body conditions. Water body maps based on water body frequency

within a year on the other hand can capture the water body dynamics (Mueller et al., 2016; Pekel et al., 2016). Water body extents detected in selected years were often used to indicate the inter-annual trends of water body conditions (Homer et al., 2015; Necsoiu et al., 2013; Tao et al., 2015). However, because of inter-annual variations of water bodies (Gaupp et al., 2015), the selection of comparing years may lead to misleading trends. In comparison, the continuous annual water body conditions over a long period can better reveal the inter-annual variations and changing trends of open surface water bodies (Pekel et al., 2016; Zou et al., 2017).

Surface water body dynamics are affected by both climate anomalies and anthropogenic activities (Bates et al., 2008; Pekel et al., 2016). Precipitation and temperature are two significant climate drivers of water body area changes (Krueger et al., 2017; Liu et al., 2013; Sellinger et al., 2008; Tao et al., 2015). Various anthropogenic activities were also found related to the changes of surface water body areas, including reservoir construction (Du et al., 2012; Pekel et al., 2016), irrigation (Liu et al., 2013), thermoelectric power production (van Vliet et al., 2012), coal, oil and gas mining (Murray, 2013; Tao et al., 2015), and urbanization (Steele et al., 2014). Surface water bodies are also related to groundwater (Brunner et al., 2009) and total land water storage (Proulx et al., 2013). The withdrawal of ground water could change the surface water and streamflow (Dale et al., 2015; Ferguson and Maxwell, 2012; Krueger et al., 2017).

More and more lakes are becoming eutrophic and hypereutrophic (OWRB, 2015). Harmful algal blooms are likely to occur in eutrophic and hypereutrophic lakes, which would lead to fish mortality, jeopardize drinking water quality, and cause health problems (Matsushita et al., 2015). Landsat images have been used to quantify water chlorophyll-

a concentration in previous studies (Dall'Olmo et al., 2003; O'Reilly et al., 1998). Compared with Landsat images, the newly launched Sentinel 2 satellites have four additional red edge bands, designed to capture the reflectance of green vegetation and chlorophyll. Field sampling and laboratory analysis of chlorophyll-a concentration is costly and time-consuming. Satellite-image based water quality estimation has the potential to provide Chlorophyll-a concentration information in high temporal frequencies and high spatial resolution with relatively low cost.

1.2 Overall research objectives

The overall objectives of this dissertation are to (1) develop efficient and easy-to-use algorithms to identify and map open surface water bodies, (2) generate annual maps of open surface water body at state, national, and global scales, (3) analyze the inter-annual variation and long-term trends of surface water body areas, (4) examine the climatic and anthropogenic drivers of water body area variability and reveal the effects of severe droughts on surface water body area, (5) assess the interannual variations and trends of both surface water area and land water storage to shed new light on the consistency and divergence between surface and land water resource dynamics, and (6) evaluate the potential of Landsat 5/7/8 and Sentinel 2 satellite images in water chlorophyll-a concentration estimation.

1.3 Organization of the dissertation

This dissertation has 6 chapters, with chapter 1 as introduction and chapter 6 as summary. Chapters 2 and 3 have been published while chapter 4 is under review. The manuscript of chapter 5 is in preparation.

Chapter 2: Continued decrease of open surface water body area in Oklahoma during 1984–2015 from analysis of time series Landsat images. Oklahoma contains the largest number of manmade lakes and reservoirs in the U.S., but vulnerable with a typical spatial transition and temporal variation in climate. In the context of climate change and variability, how these surface water bodies changed in the past three decades and what consequences would these changes bring to the society and ecosystems? These questions have not been well addressed because studies on this subject are limited in Oklahoma. This chapter aims to map open surface water bodies of the entire Oklahoma during 1984–2015, investigate the spatiotemporal dynamics of surface water body area and number, and explore their relationship with climate variables (precipitation and temperature) and human water withdrawals.

Chapter 3: Divergent trends of open surface water body area in the contiguous US from 1984 to 2016 from analysis of time series Landsat images. Affected by climate change and human activities, severe dynamics in open surface water body areas have impacted US agriculture, economy, society, and ecosystems. The historical information of surface water bodies stored in satellite images can aid in water resource planning and management, yet it is not well explored. Also, how the climate and human development have affected the variability of surface water body area at a state scale across the CONUS have not been examined. This chapter aims to map open surface water bodies of the CONUS during 1984–2016, analyze the interannual variations and trends of water body area at state and watershed scales, examine the impacts of climatic and anthropogenic drivers on water body area change, explore the relationships between water body area and GRACE land water storage, and investigate the impacts of severe

drought events on temporal dynamics of surface water body area and land water storage in California and the Southern Great Plains.

Chapter 4: Variations and trends of global surface water body area and land water storage in the past decades from analysis of time series Landsat images. Under the impacts of climate variability and anthropogenic activities, strong variations and significant changes of water resources, such as the dry-up of surface water bodies and the loss of groundwater, has threatened municipal water supplies, reduced crop production, and decreased aquatic biodiversity. While global surface water body and land water storage have been studied, a global water resource dynamic analysis through a combination of the two have not been investigated yet. This chapter aims to generate a new global multi-decadal surface water body dataset during 1984–2017 at 30-m resolution using millions of Landsat images, detect locations where surface water area had significantly changed in the past decades, analyze and visualize the interannual variations and multi-decadal trends of regional surface water area, and assess the interannual variations and trends of both regional surface water area and land water storage during 2002–2016 to shed new light on the consistency and divergence between surface and land water resource dynamics.

Chapter 5: The potential of time series Landsat and Sentinel-2 images for estimating chlorophyll-a content of open surface water bodies. Chlorophyll-a concentration is directly related to algal blooms, which has caused health problems, fish death, and the decrease of drinking water quality in Oklahoma. Oklahoma Water Resource Board carried out chlorophyll-a field sampling and lab measurement since 2002. However, the chlorophyll-a measurement is not frequent because of the time-consuming

and costly process. This chapter aims to evaluate the potential of Landsat 5/7/8 and Sentinel 2 satellite images in chlorophyll-a estimation, explore the relationships between chlorophyll-a field measurement and surface reflectance data from satellite images, and try to build robust chlorophyll-a estimation models using multiple stepwise regression analysis.

1.4 List of Publications from the Dissertation

Chapter 2

Zou, Z., J. Dong, M. A. Menarguez, X. Xiao, Y. Qin, R. B. Doughty, K. V. Hooker, and K. David Hambright (2017), Continued decrease of open surface water body area in Oklahoma during 1984-2015, Sci Total Environ, 595, 451-460, doi: 10.1016/j.scitotenv.2017.03.259.

Chapter 3

Zou, Z., X. Xiao, J. Dong, Y. Qin, R. B. Doughty, M. A. Menarguez, G. Zhang, J. Wang. (2018), Divergent trends of open surface water body area in the contiguous US during 1984-2016, Proceedings of the National Academy of Sciences, 115, 3810-3815, https://doi.org/10.1073/pnas.1719275115.

Chapter 4

Zou, Z., X. Xiao., Y. Qin, J. Dong., L. Du., R. B. Doughty., X. Wang., X. Wu., R. Bajgain., J. Wang., Q. Chang., M. A. Menarguez. (2019), Significant trends of global water resources in the past decades (Under review in Nature Geoscience).

Chapter 5

Zou, Z., X. Xiao., Estimation of chlorophyll-a content in surface water bodies from Landsat and Sentinel 2 satellite images (In preparation).

Chapter 2: Continued decrease of open surface water body area in Oklahoma during 1984–2015 from analysis of time series Landsat images

Abstract

Oklahoma contains the largest number of manmade lakes and reservoirs in the United States. Despite the importance of these open surface water bodies to public water supply, agriculture, thermoelectric power, tourism, and recreation, it is unclear how these water bodies have responded to climate change and anthropogenic water exploitation in the past decades. In this study, we used all available Landsat 5 and 7 images (16,000 scenes) from 1984 through 2015 and a water index- and pixel-based approach to analyze the spatial-temporal variability of open surface water bodies and its relationship with climate and water exploitation. Specifically, the areas and numbers of four water body extents (the maximum, year-long, seasonal, and average water body extents) were analyzed to capture variations in water body area and number. Statistically significant downward trends were found in the maximum, year-long, and annual average water body areas from 1984 through 2015. Furthermore, these decreases were mainly attributed to the continued shrinking of large water bodies (>1 km2). There were also significant decreases in maximum and year-long water body numbers, which suggested that some of the water bodies were vanishing year by year. However, remarkable inter-annual variations of water body area and number were also found. Both water body area and number were positively related to precipitation, and negatively related to temperature. Surface water withdrawals mainly influenced the year-long water bodies. The smaller water bodies have a higher risk of drying up under a drier climate, which suggests that small water bodies are more vulnerable under climate-warming scenarios.

2.1 Introduction

Climate change and increased climate variability can strongly impact surface water resources (Aherne et al., 2006; Ferguson and Maxwell, 2012; Tulbure et al., 2016), causing dramatic intra-annual and inter-annual water variability (Hall et al., 2014; Mercier et al., 2002), which has been shown to have wide-ranging consequences on human societies and ecosystems (Bates et al., 2008; Brown and Lall, 2006). Previous studies using remote sensing approaches have documented strong relationships between water body extent (area and number) and both climate variability and anthropogenic impacts (Liu et al., 2013; Pekel et al., 2016; Tao et al., 2015; Tulbure and Broich, 2013; Tulbure et al., 2014).

Water body monitoring with remote sensing techniques has advanced along with an increase in freely available high-resolution satellite data. Many surface water detection methods were developed (Table S2.1). They can be divided into general feature classification methods and thematic water body extraction methods (Li et al., 2013). The general feature classification methods include maximum likelihood classification (Henits et al., 2016), decision tree classification (Zhang et al., 2015), support vector machine classification (Li et al., 2015), artificial neural network models (Evora et al., 2008), genetic algorithms (Song et al., 2013), etc. General feature classification methods usually include data sample selection and algorithm training processes, which requires human expertise and knowledge of local area. Thus, these methods have difficulties to quickly detect water bodies in multi-temporal images at large scales (Li et al., 2013). The thematic water body extractions were usually based on different kinds of spectral water indices (Gao, 1996; Gond et al., 2004; McFeeters, 1996; Xiao et al., 2002; Xu, 2006). Different

spectral water indices have different emphasis according to their definition. The Automated Water Extraction Indexes ($AWEI_{nsh}$ and $AWEI_{sh}$) can suppress the classification noise from shadow and non-water dark surface (Feyisa et al., 2014). The Enhanced Water Index focus on the sub-pixel level analysis of water surface proportion mapping (Wang et al., 2015). Water Index 2015 (WI_{2015}) can better delineate surface water than other water indices in eastern Australia (Fisher et al., 2016). The accuracy of different water indices depends on the color, turbidity, depth, size and shape of the water bodies, the amount of bare land and urban areas, and the presence of cloud and mountain shadow (Feng et al., 2016; Fisher et al., 2016; Verpoorter et al., 2012). For large study areas, the accuracy difference among different water indices is not large due to the complexity of water features and background noise. Fisher et al. (2016) compared the performance of seven Landsat water indices (WI_{2006}, WI_{2015}, $AWEI_{nsh}$, $AWEI_{sh}$, NDWI, mNDWI, TCW_{Crist}) in water classification in eastern Australia (~3 million km^2) and found that all water indices are excellent classifiers for pure water or non-water pixels and they have similar accuracy in classifying the mixed pixels, with no index performing the best across all types of water and non-water pixels.

McFeeters (1996) defined the Normalized Difference Water Index (NDWI) using green and near infrared band to delineate open water features. Xu (2006) modified the NDWI into mNDWI by replacing the near infrared band with short-wave infrared band to suppress the noise of built-up land. The mNDWI is one of the most widely used water indices due to its good performance in water body delineation across diverse landscapes (Du et al., 2012; Feyisa et al., 2014; Hui et al., 2008; Ogilvie et al., 2015; Tao et al., 2015).

Previous remote sensing approaches have inconsistent capabilities of capturing water body variability. Many surface water bodies have strong intra-annual dynamics, during for example, wet and dry seasons (Alsdorf et al., 2007; Tulbure and Broich, 2013). But some studies estimated water body extent from satellite images gathered at a single time of the year, typically in the wet season (Feng et al., 2011; Homer et al., 2015; Liu et al., 2013). However, it is difficult to define the proper period due to uncertainties in intra-annual variability of climate and anthropogenic effects. Some studies compared the difference of water body area between the same time of selected years to indicate the increasing or decreasing trends of water body area among those years (Du et al., 2012; Homer et al., 2015; Necsoiu et al., 2013; Tao et al., 2015). However, due to the strong interannual dynamics of open surface water bodies (Hall et al., 2014; Mercier et al., 2002; Tulbure et al., 2016), the selection of different years for comparison could lead to very different results and inaccurate inference of trends in water body area and number. Thus, a comprehensive analysis considering different phases or extents of surface water bodies is important. To get a more complete picture of water body variability in Oklahoma, USA, this study explored four indicators of surface water body extents based on the annual water body frequency: 1) the maximum water body extent in a given year, 2) the persistent year-long water body extent, 3) seasonal changes in water body area, which is the difference between the maximum and year-long water body extents, and 4) the annual average water body extent.

Oklahoma is located in a climatic transition zone characterized as sub-humid in the south and east to cold and dry in the north and west, which causes a widely variable seasonal fluctuations in weather (Gibson, 1981). The few natural water bodies in

Oklahoma are temporary oxbows and playas (Johnson and Luza, 2008), but more than 200 large reservoirs have been built in Oklahoma to meet current and projected water demand (OWRB and ODWC, 2015). Between 1985 and 2010, 47% of the total annual water withdrawals in Oklahoma came from these open surface water bodies (USGS, 2010). Some water-use sectors rely more heavily on open surface water bodies as a percentage of their total water use than others, especially thermoelectric power (99%), public water supply (82%) and livestock (65%) (USGS, 2010). From 1980 to 2009, Oklahoma's soil moisture had continuously declined due to decreased precipitation and increased land surface net radiation and temperature (Lin et al., 2013). The prevailing climate models have predicted more frequent and intense droughts in the Southern Great Plains due to changes in precipitation intensity and frequency (Shafer et al., 2014). In the context of climate change and variability, changes in water body area and number would undoubtedly affect human society and ecosystems. However, these questions have not been well addressed because studies on this subject are limited in Oklahoma. It is still unclear how open surface water bodies, mostly manmade, have and will respond to a changing climate.

The objective of this study is to investigate the spatial-temporal dynamics of open surface water bodies and analyze their relationship with climate variability and anthropogenic water exploitation in Oklahoma. We used all of the Landsat 5 and 7 surface reflectance images and a water index- and pixel-based algorithm to detect surface water body changes from 1984 through 2015. Four water body extent maps (maximum, year-long, seasonal, and average) for each year were generated based on annual water body frequency at the pixel level, which better represented water body status in a more

comprehensive way. We analyzed the trends and variations of both water body area and number of four different water body extents. With a continuous long record of water body area and number in a climatic transition zone of widely variable weather, the relationship between water body variability, and climate factors (precipitation and temperature) and anthropogenic water exploitation (surface water withdrawals) were analyzed. This study aims to develop a systematic approach to monitor changes in the area and number of water bodies in Oklahoma using remote sensing techniques and to understand the effects of climate change and water exploitation on water bodies.

2.2 Materials and methods

2.2.1 Study area

Oklahoma is located in the south-central United States (Fig. S2.1), with an area of ~181,000 km^2 (U.S. Census Bureau, 2010). Its altitude decreases gradually from the high plains in the west to the forest dominated landscape of the east. Oklahoma's temperature decreases from south to north while its precipitation decreases from east to west. The statewide long-term annual average temperature and annual total precipitation are 15.4 ℃ and 857.8 mm, respectively (Oklahoma Climatological Survey, 2016). The 147 most imperative lakes and reservoirs (OWRB and ODWC, 2015) were built between 1902 and 1997. Ninety-four percent of these water bodies existed before the beginning (1984) of our study period. Oklahoma has approximately 3,000 lakes, reservoirs, and ponds that are 4 ha or larger, among which, 53 lakes are larger than 400 ha (OWRB and ODWC, 2015).

2.2.2 Data

We made use of the Landsat 5 and 7 surface reflectance data archive in Google Earth Engine (GEE) (Google Earth Engine, 2017), which is ~16,000 images of our study area.

These datasets were computed from the Landsat standard Level 1 Terrain-corrected (L1T) images in USGS using Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) algorithms (USGS, 2012). Observations of unacceptable quality, caused by invalid pixels, cloud, and snow, were excluded in our calculation based on the corresponding masks in the 8-bit quality band in each image. The availability of high-quality observations were crucial to the generation of the annual water body maps. Landsat tiles and the total image number of each tile in the last 32 years were shown in Fig. 2.1a. The numbers of available images in each year are shown in Fig. 2.1b, and the distribution of average annual good observations of all pixels from 1984 through 2015 are shown in Fig. 2.1c. More than 99.9% of the pixels had 14 or more good observations per year. On average, there were ~25 good observations per pixel in a year. Total observations, total good observations, and percentage of good observations by pixel in 2015 and 1984–2015 were shown in supplementary materials (Fig.S2.2, Fig. S2.3).

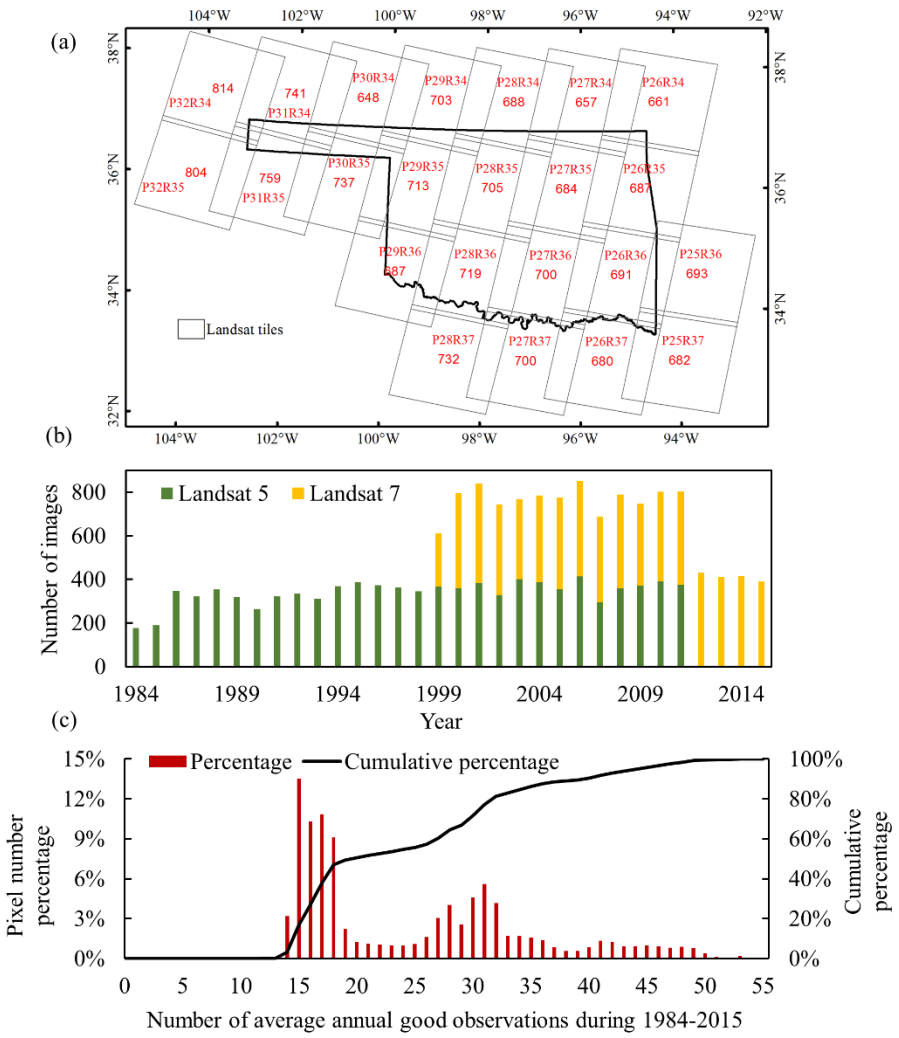


Figure 2.1 Landsat data. (a) Landsat tiles and total Landsat 5 and 7 image numbers of each tile from 1984 through 2015, (b) total Landsat 5 and 7 images of the study area in each year, (c) distribution of the average annual good observations from 1984 through 2015, including pixel number percentages of good observations in red bars and cumulative percentages in black curve.

2.2.3. Algorithms to identify open surface water bodies

Despite the advantage of mNDWI over NDWI in the remote sensing of water bodies (Feyisa et al., 2014; Ji et al., 2009; Xu, 2006), the mNDWI approach still has commission error in the mixed pixels of water and other land cover types. In particular, vegetation over a wet surface is one of the major causes for commission error in open surface water

body mapping (Santoro et al., 2015). In this study, we combined mNDWI and vegetation indices (NDVI and EVI) to reduce the effects of vegetation on water body mapping algorithm (Equations 2.1-2.3), which has been addressed in our previous studies (Dong et al., 2015; Xiao et al., 2006; Xiao et al., 2005). Specifically, we detected only pixels with stronger water signal than vegetation signal as actual water pixels (mNDWI > NDVI or mNDWI > EVI). In order to further remove the noise caused by vegetation, EVI was applied to exclude the wetland pixels with vegetation (EVI < 0.1). Therefore, only those pixels that met the criteria ((mNDWI > NDVI or mNDWI > EVI) and (EVI < 0.1)) were classified as open surface water body pixels. The remaining pixels were classified as non-water pixels (Fig. 2.2). The water detection algorithm ((mNDWI>NDVI or mNDWI>EVI) and EVI<0.1) in this study can be divided into two parts (mNDWI-NDVI>0 and EVI<0.1) and (mNDWI-EVI>0 and EVI<0.1). A pixel meets the criteria of either part is classified as water pixel. Scatterplots of (mNDWI-NDVI) vs EVI and (mNDWI-EVI) vs EVI were used to show the pixel distribution in the built-up area (Fig. S2.4) and vegetated area (Fig. S2.5).

$$\text{mNDWI} = \frac{(\rho_{green} - \rho_{SWIR1})}{(\rho_{green} + \rho_{SWIR1})} \quad \text{(Equation 2.1)}$$

$$\text{NDVI} = \frac{(\rho_{NIR} - \rho_{Red})}{(\rho_{NIR} + \rho_{Red})} \quad \text{(Equation 2.2)}$$

$$\text{EVI} = 2.5 \times \frac{(\rho_{NIR} - \rho_{Red})}{(1 + \rho_{NIR} + 6\rho_{Red} - 7.5\rho_{blue})} \quad \text{(Equation 2.3)}$$

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.

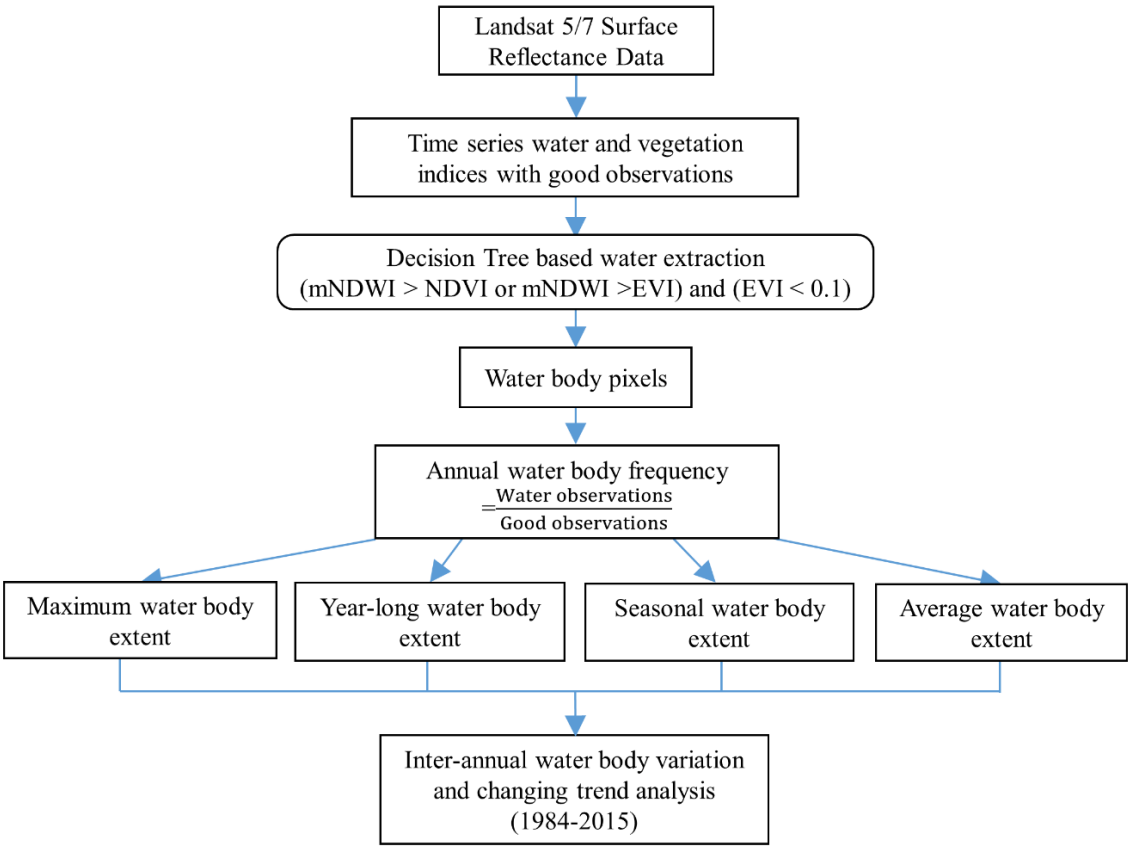


Figure 2.2 Workflow of this study.

The algorithm was validated using 1-m spatial resolution images as ground reference data in the platform of ArcMap 10.3.1. A stratified random sampling method was used to generate ground reference points (Fig. S2.6). One thousand water random points and 1500 non-water random points were generated within Oklahoma water and non-water boundaries from National Land Cover Database 2011 (Homer et al., 2015). Each of the 2500 random points was added to the single-temporal water body map and then extracted the 4 adjacent pixels around it. Altogether, 9942 pixels were extracted and added to the 1-m resolution NAIP images (when a point falls in the edge of missing data area, it may extract less than 4 pixels, thus the total number of extracted pixels is less than 10,000). Each pixel was visually interpreted as water or non-water pixel according to

NAIP images in the platform of ArcMap 10.3.1. The 9942 water/non-water pixels served as ground reference data to generate the confusion matrix of the single-temporal water body map (Table S2.2). According to the confusion matrix, the overall accuracy of the water detection method is 94%, the kappa coefficient is 0.86 and the producer's accuracy is 83%.

We used these mapping methods and GEE cloud-computing platform to identify open surface water bodies on all of the 16,000 Landsat images during 1984-2015. For each pixel, we counted the number of observations within a year it was identified as open surface water body, and then divided it by the total number of good observations in that year. We termed the resultant ratio as water body frequency. When the pixel had an annual water body frequency greater than or equal to 0.25, it was classified as effective open surface water pixels. This frequency standard (>=0.25) was chosen because we need to reduce the potential error from the uncertainty in image data quality flags and other small-probability problems in image preprocessing (Fig.S.2.7). All of the effective open surface water pixels in a year formed the maximum water body extent. Water pixels with an annual water body frequency greater than or equal to 0.75 were classified as year-long water pixels since they have water most of the year. The remaining water pixels, with a water body frequency spanning from 0.25 to 0.75, were classified as seasonal water pixels. For each year, we generated annual maps of maximum water bodies (water body frequency >=0.25%), year-long water bodies (water body frequency >=75%), and seasonal water bodies (25% <= water body frequency <75%), respectively, and then calculate the areas of maximum, year-long and seasonal water bodies. Annual average

water body area is calculated as a product of all the effective water body pixels and the length of water body (water body frequency) (Equation 2.4).

$$\text{Annual average water body area} = \sum_i^n (F_i) \times A \qquad \text{(Equation 2.4)}$$

Where F_i is the water body frequency of water pixel *i* ranging from 0.25 to 1; *n* is the total number of water pixels within Oklahoma; *A* is the area of one pixel, which is 900 m^2.

We also compared these resultant maps with previous studies for the purpose of inter-comparison. Since most of the available water body data and maps were static or in a specific year (Lehner and Döll, 2004; Verpoorter et al., 2014), we used the annual maximum water body extent in 2000, 2001, 2006, and 2011 from our study to compare with the Global Inland Water (GIW) dataset of 2000 (M Feng et al., 2016), National Land Cover Database (NLCD) of 2001 (Homer et al., 2004; Homer et al., 2007; Vogelmann et al., 2001), NLCD of 2006 (Fry et al., 2011), and NLCD of 2011 (Homer et al., 2015). According to the confusion matrixes (Table S2.3), the maximum water body maps in this study showed strong agreement with GIW and NLCD, with overall accuracy > 99% and Kappa coefficients > 0.8.

2.2.4. Interannual variations and trends of open surface water bodies

The area of maximum, year-long, seasonal, and annual average water body extent in all years (1984-2015) were calculated. Their inter-annual variations during 1984-2015 were analyzed using anomaly analysis while their changing trends were analyzed through linear regression analysis. Water pixels adjacent to each other in the maximum and year-

long water body raster maps were merged and converted into vectors, respectively. The number of maximum and year-long water bodies in each of the last 32 years were counted and their inter-annual variations and changing trends from 1984 to 2015 were also analyzed through anomaly and linear regression analysis.

2.2.5. Effects of climate and water exploitation on surface water bodies

Multiple linear regression analyses were conducted to explore the relationship between climate, anthropogenic water exploitation, and the variability in the number and area of water bodies. The dependent regression variables of the six regression models were maximum water body area and number, year-long water body area and number, seasonal water body area, and annual average water body area from 1985 to 2015. The climate factors included statewide annual total precipitation and annual average temperature from the Oklahoma Climatological Survey (McPherson et al., 2007). The water exploitation was represented by the statewide annual surface water withdrawal data, which was gathered every 5 years by U.S. Geological Survey (USGS, 2010). This surface water withdrawal data was interpolated into annual water withdrawal data spanning from 1985 through 2015 (Equation 2.5). Water body condition of current year changes from the water body condition of the previous year because of the legacy effect. Thus, the dependent variable in the previous year of each model was used as an independent variable of the subsequent year, serving as the base of water body change in the subsequent year.

$$W_{1986} = W_{1985} + \frac{(W_{1990} - W_{1985})}{5} \qquad \text{(Equation 2.5)}$$

Where W_{1985}, W_{1986}, and W_{1990} were the statewide surface water withdrawal in 1985, 1986, and 1990, respectively.

2.3 Results

2.3.1. Open surface water bodies in 2015 and 1984-2015

There were 3.3 million water pixels in both the annual water body frequency map of 2015 and the long-term water body frequency map of 1984-2015, which represented a maximum water body extent of 2980 km^2 and accounted for ~1.6% of the entire state area of Oklahoma (Fig. 2.3a-b). The distribution of different water body frequency levels of 2015 and 1984-2015 (Fig. 2.3c-d) showed that about 70% of the water pixels had a water body frequency greater than or equal to 0.75. These water pixels formed the interior portions of large lakes, reservoirs, and major rivers that were able to maintain water throughout the year. However, water at the shallow edges of these large water bodies would dry up at times due to fluctuations in the water level. For example, the center region of Keystone Lake (Fig. 2.3a-b insets) had water body frequency values close to 1, which meant that it always had water and is the deepest part of the lake. The upper part of the lake had water body frequency values around 0.7, which indicated that this portion of Keystone Lake could dry up at times and was not very deep. The water body frequency values where the river joined the lake were below 0.5, which indicated that this area only had water during the wet seasons and that these portions of the lake were shallow. A large number of small water bodies had a low water body frequency, meaning that they only existed for several months during the wet season or became so small that they could not be detected. The number of water pixels observed in each of the last 32 years were distributed across 8 water body frequency levels in Fig. 2.3e. The majority of water body pixels had a big water body frequency. There were also some interannual variations among different water body frequency levels.

Out of the 3.3 million water pixels in 2015 (Fig. S2.8), there were about 2.3 million year-long water body pixels, which formed the central part of large lakes, reservoirs, and major rivers. The remaining 1.0 million pixels indicated seasonal water bodies, which was comprised of small ponds, minor rivers, and the edges of large water bodies. In 2015, 70% and 30% of the maximum water body area was year-long and seasonal water body area, respectively.

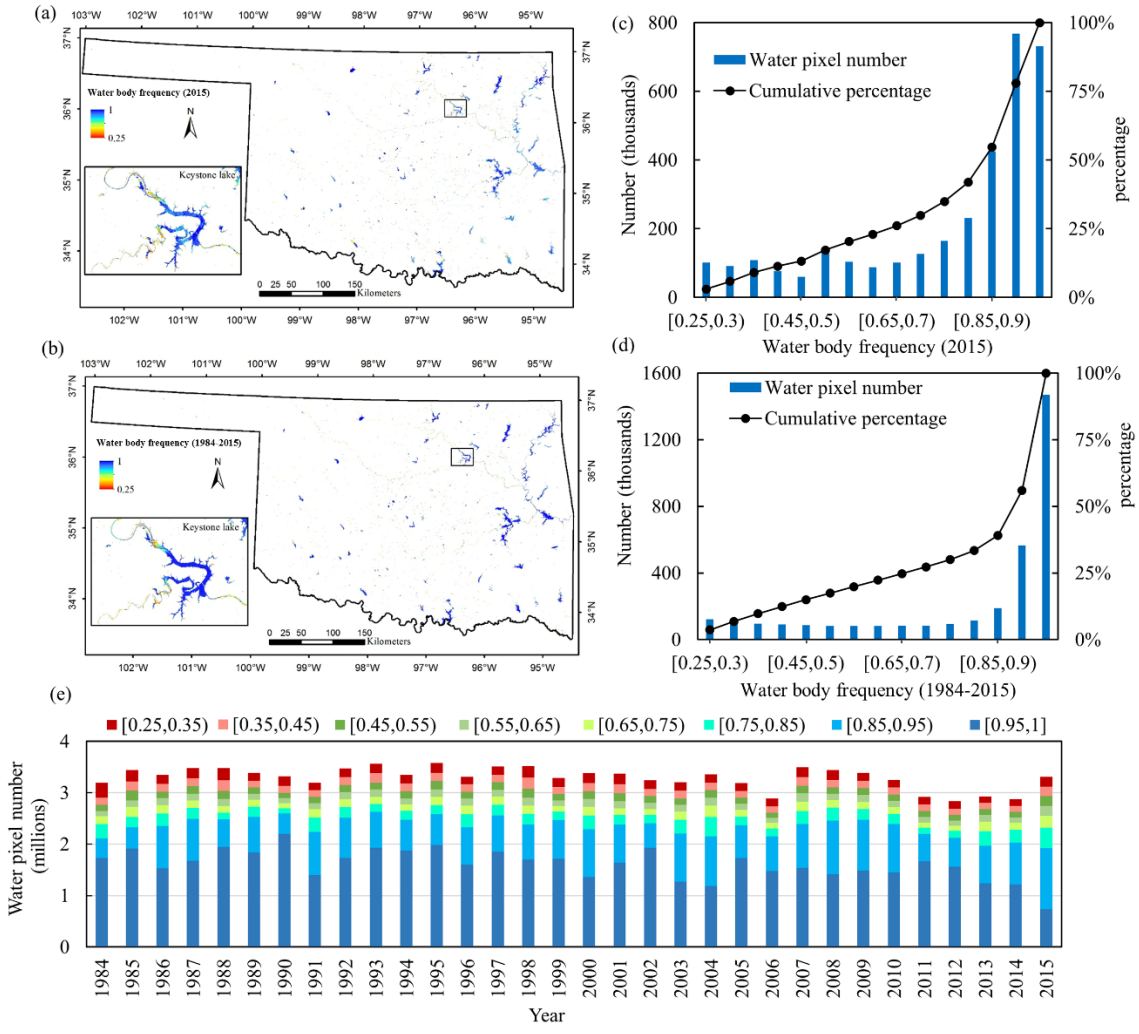


Figure 2.3 Water body frequency distribution in Oklahoma. Water body frequency map of 2015 (a) and 1984-2015 (b). The distribution of different water body frequency levels with a bin of 0.05 in 2015 (c) and 1984-2015 (d). Distribution of different water body frequency levels with a bin of 0.1 in 32 years (e).

2.3.2. Interannual variation of open surface water bodies during 1984-2015

The maximum, year-long, seasonal, and annual average water body areas showed similar patterns of variation from 1984 to 2015, which were also similar to the variability of precipitation (Fig. 2.4). The annual maximum water body area from 1984 through 2015 varied between 2548 and 3224 km^2, which was 14% below to 9% above its average value (~2966 km^2). The year-long water body areas varied between -12% and 9% of its average value (2302 km^2), while the seasonal water body areas had the largest variability, from 23% below to 34% above its average value (665 km^2). The annual average water body area best described the average water body extent within one year since it considered the length of water existence of all effective water pixels. The annual average water body area in the last 32 years varied between 2205 and 2758 km^2, which was 12% below to 9 % above its mean value (2520 km^2). Statistically significant downward trends were found in the maximum water body areas (R^2=0.29, p=0.001), year-long water body areas (R^2=0.28, p=0.002), and annual average water body areas (R^2=0.37, p<0.001) in the last 32 years (Fig. 2.4). These downward trends indicated shrinkage of total statewide water body area. According to the linear regression model, the statewide annual average water body area shrank 10 km^2 each year over the last three decades (Fig. 2.4d).

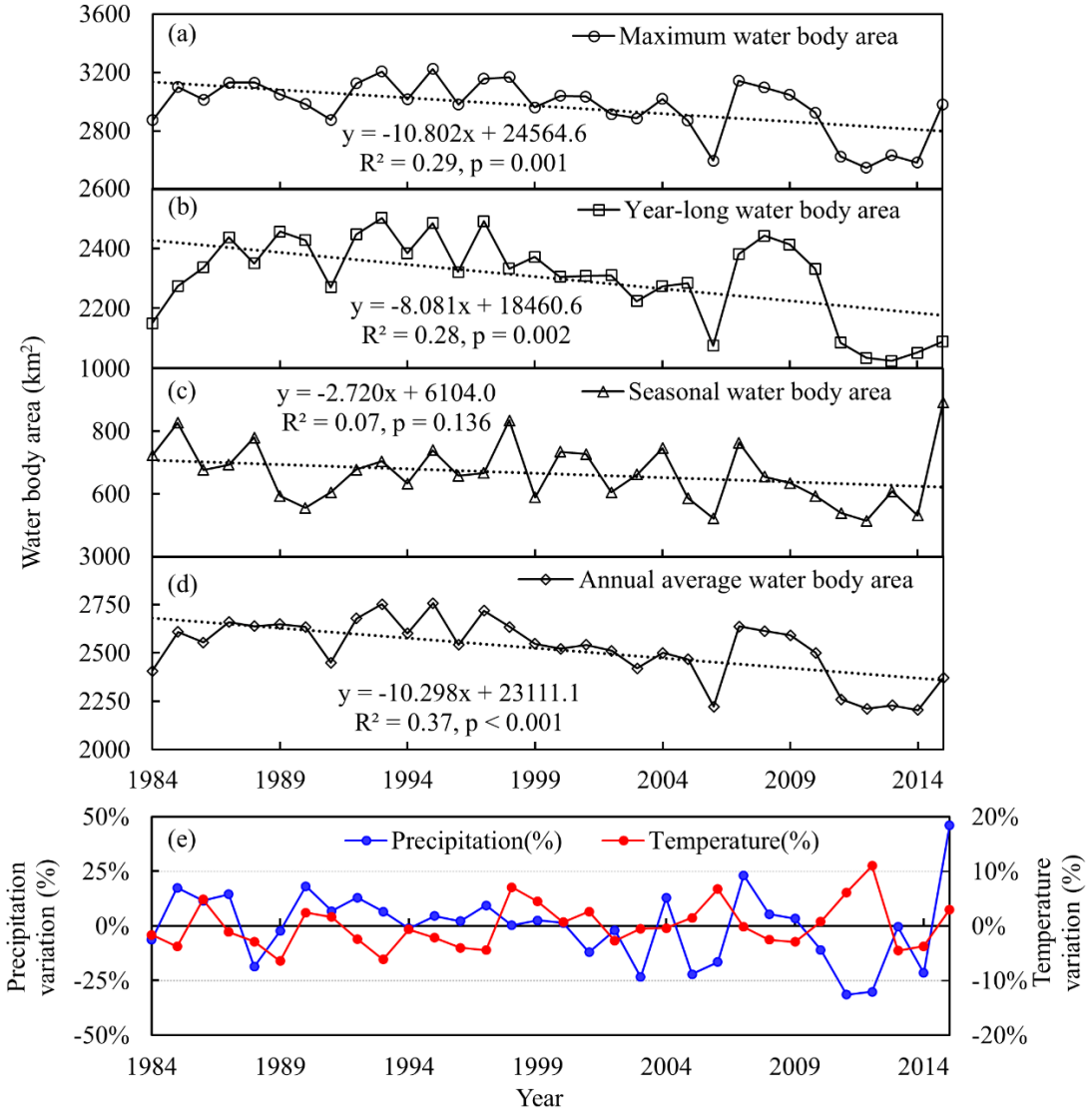


Figure 2.4 Inter-annual variations of water body area in different water extents, including maximum (a), year-long (b), seasonal (c), and average (d) water body extents. (e) Statewide annual total precipitation and annual average temperature.

The number of maximum and year-long water bodies showed similar patterns of variation from 1984 through 2015 (Fig. 2.5), which again was similar with those of water body area variations (Fig. 2.4). The annual maximum water body number in the last 32 years varied between 54,000 and 92,000, which was 32% below to 16% above the average value (79,000). The average year-long water body number from 1984 through 2015 was 36,000, varying between 24,000 and 45,000, which was 33% below to 25% above its

average. Statistically significant downward trends were found in the maximum water body numbers (R^2 = 0.48, $p < 0.001$) and year-long water body numbers (R^2 = 0.28, p = 0.002) over the last 3 decades. These decreasing trends in water body number indicated that some water bodies were disappearing year by year.

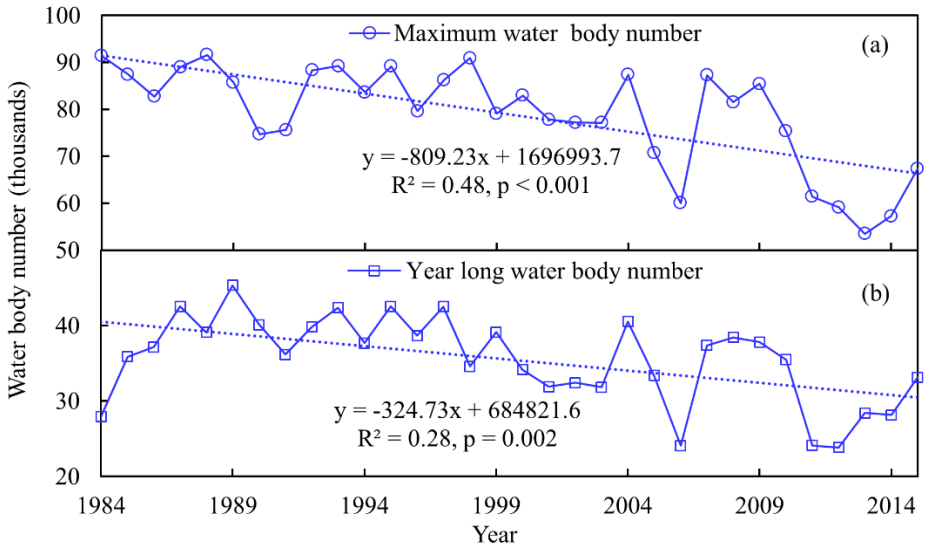


Figure 2.5 Inter-annual variations of the number of (a) maximum water bodies and (b) year-long water bodies.

All of the water bodies in the maximum water extent of each year were classified into 10 ranges based on water body size. The distribution of water body number and area in different classifications were shown in Fig. 2.6a, b. On average, the number of water bodies larger than 100 ha was 138, which made up ~ 0.18% of the total number of water bodies. However, these larger water bodies contributed 80% of the total water body area on average. The inter-annual variation in area of these large water bodies contributed ~68% of the statewide water body area variation. In comparison, water bodies smaller than 0.5 ha accounted for 77% of the total number of water bodies on average, but they comprised only 3.6% of the total water body area. These small water bodies accounted

for ~71% of the statewide interannual variation in the number of water bodies. Therefore, variability in water body area was influenced mostly by the large water bodies while the variation in the number of water bodies statewide was mainly caused by the small water bodies.

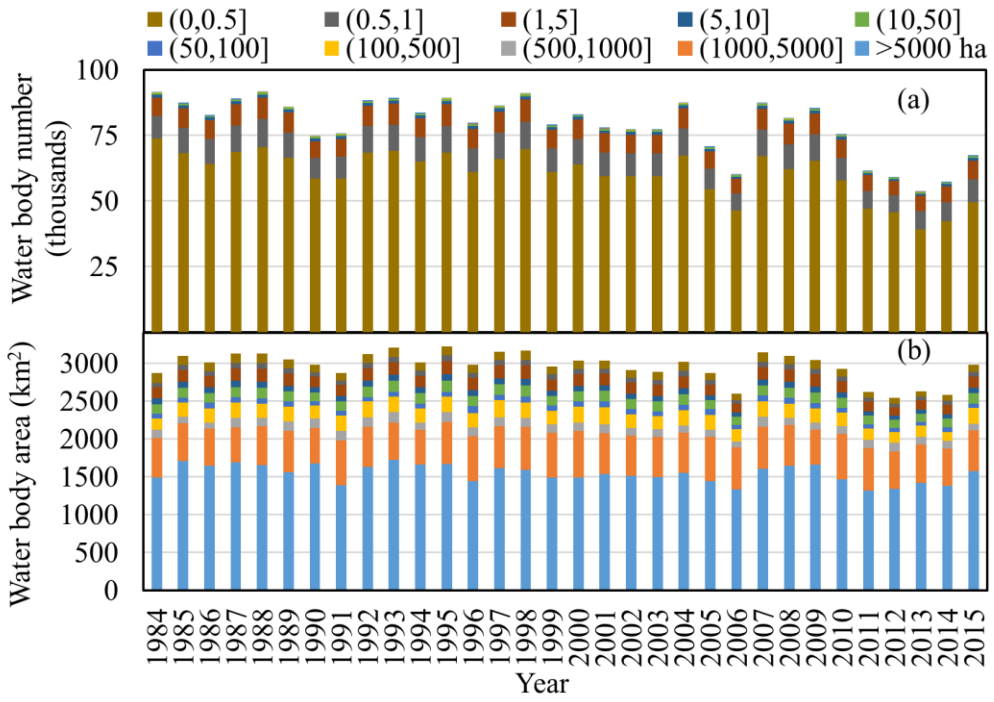


Figure 2.6 Water body number and area distribution at different water body size levels, (a) water body number distribution, and (b) water body area distribution.

2.3.3. Attribution analysis of surface water body area and number

Multiple linear regression was performed with SPSS Statistics 19 using the "stepwise" method for explanatory variable selection. The variance inflation factor (VIF) was used as the collinearity index. The VIFs for all input explanatory variables of each model were below 2.4 (Table S2.4). The results of multiple linear regression analysis were shown in Table 2.1. Precipitation had statistically significant positive effects on all six analyses. Precipitation is the major water source for Oklahoma open surface water bodies.

Basically, more precipitation leads to more water bodies and a larger water body area. Temperature had statistically significant negative effects on the annual average water body area, year-long water body area and number. Higher temperature will increase evaporation in addition to other factors, such as higher wind speed, lower concentration of water vapor in the air, lower air pressure, larger surface area, etc. Higher temperature may also increase agricultural water demands. Thus, higher temperature may reduce water body area and number. Surface water withdrawal had negative effects on the annual average water body area, year-long water body area and number. In Oklahoma, total surface water withdrawal increased from 707 million gallons per day (Mgal/day) in 1985 to 1140 (Mgal/day) in 2010 (USGS, 2010). The surface water was mainly used for public supply (55%), thermoelectric power (18%), irrigation (13%), and livestock (9%). The surface water withdrawal for public supply, irrigation, and livestock increased gradually from 1985 to 2000 and then decreased gradually from 2000 to 2010. In comparison, the surface water withdrawal for thermoelectric power before 2000 was relatively stable but increased rapidly from 143 (Mgal/day) in 2000 to 384 (Mgal/day) in 2010. Generally, these sectors divert water from year-long water bodies (large lakes, reservoirs, and major rivers), thus having more direct effects on the variability of these large year-long water bodies.

The water body area and number of the previous year had statistically significant positive effects on all six analyses, except the seasonal water body area. The water body extent of one year is gradually changed from the water body extent of the previous year. The water bodies that exist through one year will become the water bodies of the next year, positively affecting the water body extents of the subsequent year. As for seasonal

water bodies, they last shorter than 9 months by definition. Thus, seasonal water bodies in one year may dry up some time within that year and have no significant effect on the water bodies of the subsequent year.

Table 2.1 Multiple linear regression analyses of water body area and number with precipitation, temperature and surface water withdrawal in Oklahoma.

Maximum water body area		Maximum water body number		Year-long water body area		Year-long water body number		Seasonal water body area		Annual average water body area	
Variable	Coeff.	Variable	Coeff.	Variable	Coeff.	Variable	Coeff.	Variable	Coeff.	Variable	Coeff.
P	0.76***	P	29.14**	P	0.26*	P	14.65***	P	0.37***	P	0.37**
T		T		T	-57.16*	T	-3038.99**	T		T	-55.87*
SWW		SWW		SWW	-0.23*	SWW	-9.95*	SWW		SWW	-0.30*
$MWBA_p$	0.55***	$MWBN_p$	0.64***	$YWBA_p$	0.48***	$YWBN_p$	0.29*	$SWBA_p$		$AAWBA_p$	0.37*
Constant	644	Constant	286	Constant	2065***	Constant	68535***	Constant	321**	Constant	2396**
Model summary		Model summary		Model summary		Model summary		Model summary		Model summary	
R^2	0.64	R^2	0.59	R^2	0.72	R^2	0.79	R^2	0.36	R^2	0.73
SEE	119	SEE	7160	SEE	81	SEE	2791	SEE	78	SEE	89
F	24.79	F	20.49	F	16.75	F	24.48	F	16.30	F	17.79
Sig.	0.000	Sig.	0.000	Sig.	0.000	Sig.	0.000	Sig.	0.000	Sig.	0.000

* $p < 0.05$; ** $p < 0.01$;*** $P<0.001$

The six dependent variables are maximum water body area (MWBA), maximum water body number (MWBN), year-long water body area (YWBA), year-long water body number (YWBN), seasonal water body area (SWBA) and annual average water body area (AAWBA). P and T are the statewide annual total precipitation and annual average temperature respectively. SWW is the statewide surface water withdrawal in million gallons per day. $MWBA_p$, $MWBN_p$, $YWBA_p$, $YWBN_p$, $SWBA_p$, and $AAWBA_p$ denote the water body status in the previous year. R^2 is the proportion of variance in the dependent variable which can be explained by the selected explanatory variables. SEE is the standard error of the estimate. F and Sig. are the F-statistic and the p-value associated with it.

2.3.4. Variation of open surface water bodies in a dry and wet year

Precipitation is one of the most dominant climate drivers of water availability (Bates et al., 2008). Therefore, precipitation has strong effects on the water body area and number. Statewide annual total precipitation in 2006 and 2007 was 780 mm and 1150 mm, respectively. Compared with the average precipitation over the 32 years (934 mm), 2006 was a dry year while 2007 was a wet year. Figure 2.7 shows the distribution of statewide water body area and number at the maximum water body extent for 2006 and 2007. In the wet year of 2007, the area and number of water bodies were much larger than those

in the dry year of 2006. The number of water bodies in 2006 was about 60,000, which was 27,000 less than that of 2007 (87,000). The additional 27,000 water bodies in 2007 were mainly small water bodies, of which 21,000 were smaller than 0.5 ha, 3,000 were between 0.5 and 1 ha, and 2,000 were between 1 and 5 ha. Accordingly, the changes in total number of water bodies in each year were mainly caused by changes in the number of small water bodies. The existence and detection of these small water bodies were strongly affected by the amount of precipitation. The maximum water body area in 2006 was about 2596 km^2, which was about 550 km^2 less than that of 2007 (3,143 km^2). Of the additional 550 km^2 water area in 2007, 68% was attributed to the increase in area of 148 large water bodies (> 100 ha). Thus, the variability in water body area was mainly caused by variations in the surface area of large water bodies.

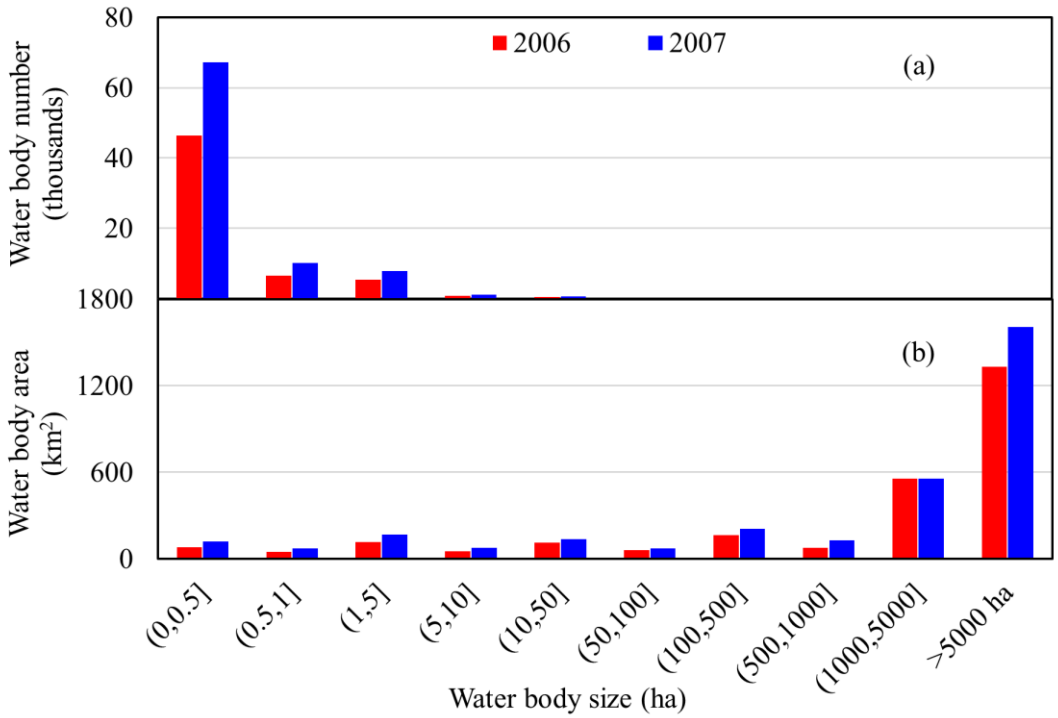


Figure 2.7 Water body number and area distribution of the maximum water body extent in a dry (2006) and wet (2007) year, (a) water body number distribution, (b) water body area distribution.

2.4 Discussion

2.4.1 Advantages and uncertainties of this study

Oklahoma has a large number of small ponds and lakes, and these small water bodies tend to have large temporal variability in extent (i.e., size and water body frequency). In order to characterize the intraannual and interannual variations of the water bodies, we proposed four water body extent related indicators derived from water body frequency maps (maximum = sum area of all effective water body pixels within a year; year-long = pixels covered by water for at least 75% of year, seasonal = pixels covered by water between 25 and 75% of the year, and average = all the effective water body pixels, weighted by the water body frequency). Together, these indicators captured a more complete picture of the variability of surface water bodies. A recent global water body mapping study provided the time of water presence and location of water change in terms of seasonality and persistence (Pekel et al., 2016). However, because of the global scale involved, Pekel et al. (2016) didn't include such detail as the annual change in water body area and number of different water extents, nor any information regarding the annual average water body extent. In addition, our algorithm had a robust performance based on the combined relationships for mNDWI and EVI/NDVI, instead of a certain threshold. The constant thresholds in previous studies could be subjective and time-consuming (Feyisa et al., 2014), and also difficult to extrapolate to other regions due to the difference in different images and locations (Ji et al., 2009).

The classification error of this study was mainly caused by omission error (Table S2.2). Omission error was reported greater than commission error in most water indices (Feyisa et al., 2014; Fisher et al., 2016; Li et al., 2013). Mixed pixels at the edge of water

bodies could be a major reason for water pixel omission (Fisher et al., 2016). Narrow rivers and streams were often not or only partially detected because of the weak water signal in the mixed pixels (Feng et al., 2016). For a single Landsat 5/7/8 image, a river needs to be at least 60-m wide to make sure that it can be detected from that image because smaller rivers may end up in two mixed pixels. However, a river as small as 30-m wide could be captured in the annual or 32-year water body frequency map as seasonal water because it can be captured in some of the images within the study period. The major rivers in Oklahoma often have broad, sand-filled channels with active water courses occupying a small portion of the river bed (Johnson and Luza, 2008). Thus, many rivers and streams had low water body frequency in our study and appeared in the seasonal water body maps rather than year-long water body maps. The low albedo surfaces, including asphalt roads, shadows of mountains, buildings, trees, and clouds are the major source for commission error in water classification (Feng et al., 2016; Feyisa et al., 2014; Verpoorter et al., 2012). Although the cloud mask band was applied in data preprocessing, the undetected residual cloud and cloud shadows would still lead to commission error. The water body frequency threshold (0.25) used here could remove most of the temporal noise out of the water body frequency maps. However, while removing the noise, the frequency threshold also removed some temporary water signals, which may have led to the underestimation of water body area.

2.4.2 Driving factors of water body changes

Before 2011, anytime the maximum water body area had a drop of more than 200 km^2 (1996, 1999 and 2006), it began to recover in the following year (Figure 2.4). However, in 2011, when the maximum water body area dropped 300 km^2, it continued dropping in

2012 and remained low through 2013 and 2014. The shrunken water body area from 2011 to 2014 was very likely caused by the long-lasting drought in Oklahoma from 2011 to 2014 (Hoerling et al., 2014; Kogan and Guo, 2015). In 2012, when the statewide annual precipitation was the second lowest (653 mm) in the last 32 years (Figure 2.4e), the maximum water body area was the smallest (2,548 km^2). The Southern Great Plains of the US are expected to have more frequent and more intense droughts in the future (Shafer et al., 2014). Thus, there is a higher probability for the total water body area to be smaller and a higher chance for the water body area to decline to a new low record – both issues will pose more challenges to the human society and the affected ecosystems. In addition to the climate-based driving factors, anthropogenic activities, including agricultural irrigation, energy production, consumptive water use and water management can also cause changes in open surface water bodies (Liu et al., 2013; Tao et al., 2015). More analysis regarding the influence of human activities on surface water bodies should be considered in future studies. It is also worth to mention that the significant trends of both surface water body area and number are based on current data during 1984–2015. These trends might change in the future if prolonged wet periods occur.

2.4.3 Consequences of water body area shrinkage

Statistically significant downward trends were found in water body area and number over the last 32 years, which indicated the shrinkage of water body area and the gradual vanishing of some water bodies. Open surface water bodies are the major water source for public supply, thermal electric power industry, and livestock production in Oklahoma. The shrinkage of water body area could have a huge influence on Oklahoma's socioeconomic systems. The prolonged drought in 2011 and 2012 reduced the water body

area of Oklahoma to a great degree. For example, the water levels of Oklahoma City's Lake Hefner was at an all-time low and water from other lakes had to be siphoned for public supply (Campfield, 2013). Thermoelectric plants depend on surface water for cooling, fuel processing, and emission control. Water withdrawals for thermoelectric power in Oklahoma increased 170% from 2000 to 2010 (USGS, 2010). The shrinkage of water area could limit the availability of water for withdrawal, expose the water intake structures, and increase water temperatures beyond regulations (Argonne National Laboratory, 2012). Similarly, decreased surface water supplies can threaten Oklahoma's 8 hydroelectric projects, which supplies electricity to about 2 million users across Oklahoma and 5 bordering states (USACE, 2017). The cattle market is the dominant livestock industry in Oklahoma, with approximately 5.5 million cattle and calves on farms and ranches, ranking the state third in the nation for beef cow production (Oklahoma Water Resources Board, 2011). Oklahoma accounts for about 9% of the total freshwater withdrawals for livestock in the US, ranking the state third behind California and Texas (Oklahoma Water Resources Board, 2011). Sixty-five percent of this freshwater in Oklahoma was obtained from open surface water bodies (USGS, 2010). Oklahoma's livestock industry is sensitive to the availability of water resources, as seen after the 2011 drought, when the number of cattle and calves in 2012 decreased about 20% to 4.2 million compared to 2007 (5.4 million) (USDA-NASS, 2014). Thus, the trends in water variability discovered in this study should be considered in Oklahoma water resource planning, especially in the sectors of public water supply, hydroelectric and thermoelectric power, and livestock.

Oklahoma is one of the most ecologically diverse states in the nation. It is one of the four states to have more than ten Level III ecological regions (Woods et al., 2005). The shrinkage of water area could also pose threats to these diverse ecosystems. The 2011 drought decreased the flow of the Kiamichi River, Little River, and Mountain Fork Rivers in southeastern Oklahoma substantially and changed the typical continuous flow to discontinuous flow, resulting in the creation of a series of shallow pools along the river channels (Atkinson et al., 2014). From 1992 to 2011, the drought-induced reductions in stream flow and surface water area of Kiamichi River had led to a > 60% decline in mussel populations (Vaughn et al., 2015). These changes caused the decrease of mussel density and biomass, and a subsequent loss of mussel-provided ecosystem services (Atkinson et al., 2014). There are a large number of small water bodies distributed across the entire state of Oklahoma. A reduction in the size and number of these small water bodies could lead to the loss of wetlands and threaten the aquatic species that depend on these small water bodies, although the disappeared small water bodies may reappear in the future pluvial years. The annual water body frequency map of the last 32 years and the cumulated water body frequency map of 1984 through 2015 could be used to identify vulnerable aquatic ecosystems that may be subject to drying in future drought years. Thus, actions could be taken to protect endangered aquatic species.

2.5 Conclusions

Oklahoma has the largest number of artificial lakes in the United States. Therefore, this water body variation study is helpful to private and public natural resource managers and improves our understanding of water resource vulnerability in the Southern Great Plains, which is experiencing increased variability in climate. In this study, the Landsat 5 and 7

surface reflectance archives from 1984 through 2015 was used to characterize water body variations at 30-mm spatial resolution. Using these data, both the area and number of different water body extent indicators were analyzed to investigate the water body variability and determine trends over the last 32 years. The water body area of the maximum, year-long, and average water extents showed significant downward trends over the last three decades, indicating that open surface water bodies are gradually shrinking in Oklahoma. Statistically significant downward trends were also found in the number of water bodies in the maximum and year-long water extents in the same period, suggesting that water bodies were vanishing gradually. Both the water body area and number underwent obvious variations over the study period. The variability in statewide water body area was mainly influenced by changes in the spatial extent of large water bodies, while the variability in the total number of water bodies was mainly influenced by the small water bodies. Precipitation had statistically significant positive effects on water body area and number while temperature had negative effects. Surface water withdrawals mainly impacted the year-long water bodies. The datasets generated by this study are useful in water resource management, agricultural irrigation, livestock production, and ecological conservation.

Supplementary materials

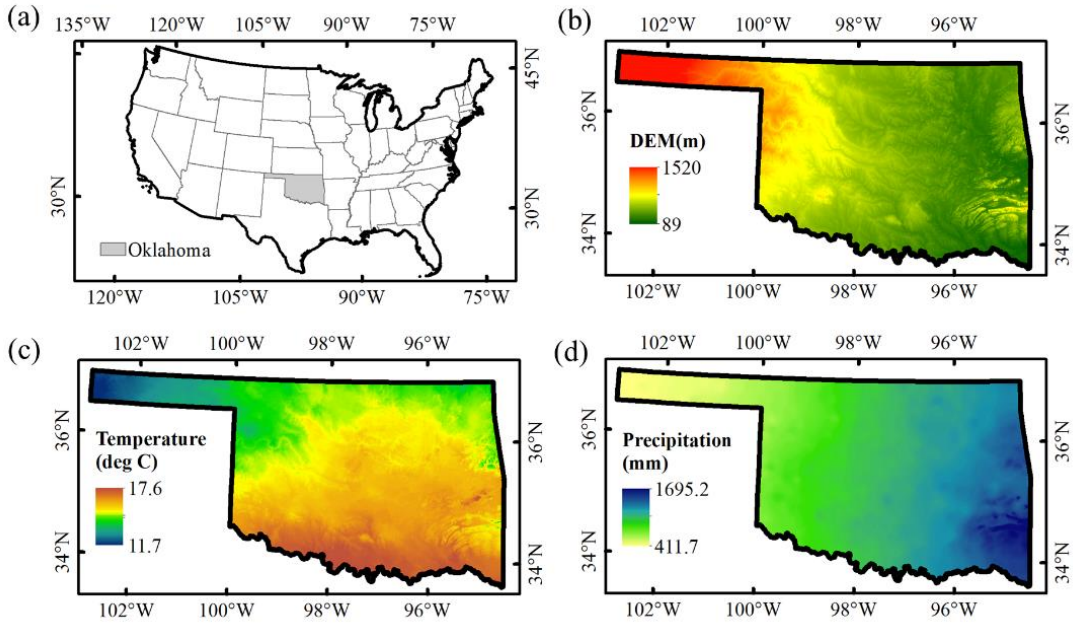


Figure S2.1 Study area. (a) the location of Oklahoma in U.S., (b) digital elevation model (DEM), (c) annual average temperature during 1981-2010, (d) annual total precipitation during 1981-2010. DEM is from USGS EarthExplorer (http://earthexplorer.usgs.gov/) while the temperature and precipitation data are from Oregon State University (http://prism.oregonstate.edu/).

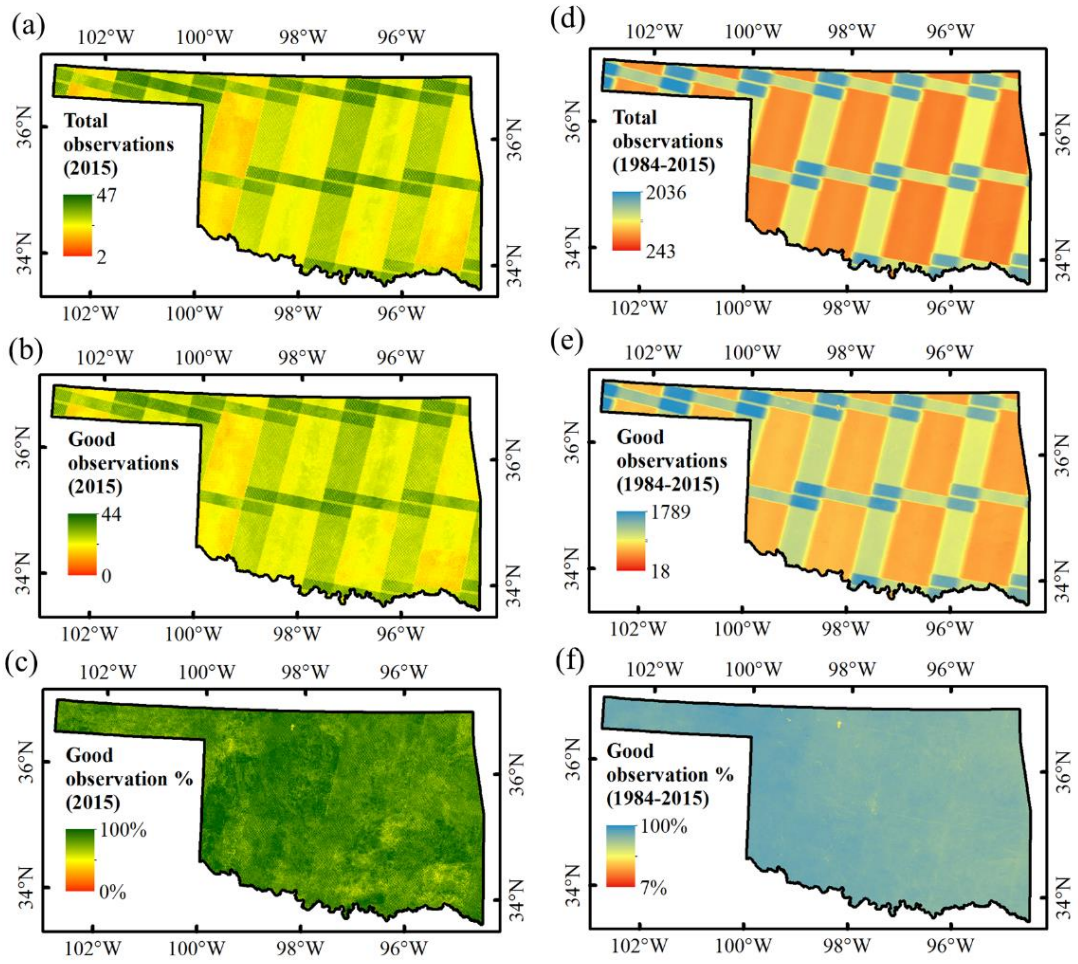


Figure S2.2 Data availability in 2015 (left) and 1984-2015 (right). (a) total observations in 2015, (b) good observations in 2015, (c) good observation percentage in 2015, which is the ratio of good observation to total observation in 2015 (d) total observations in 1984-2015, (e) good observations in 1984-2015, (f) good observation percentage in 1984-2015, which is the ratio of good observation to total observation in 1984-2015.

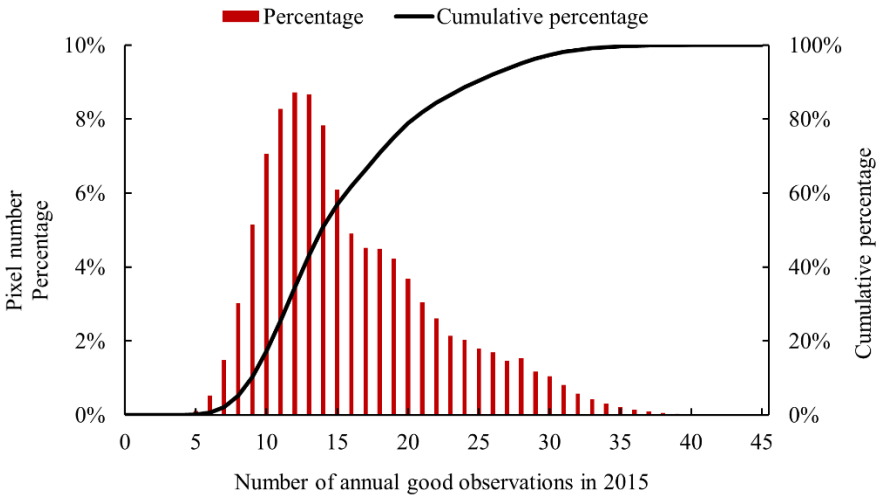


Figure S2.3 Histograms of pixels with various annual good-quality observation numbers in 2015, including (1) frequency distribution (percentage) in red bars and (2) cumulative percentages of pixels with good observations in black curve.

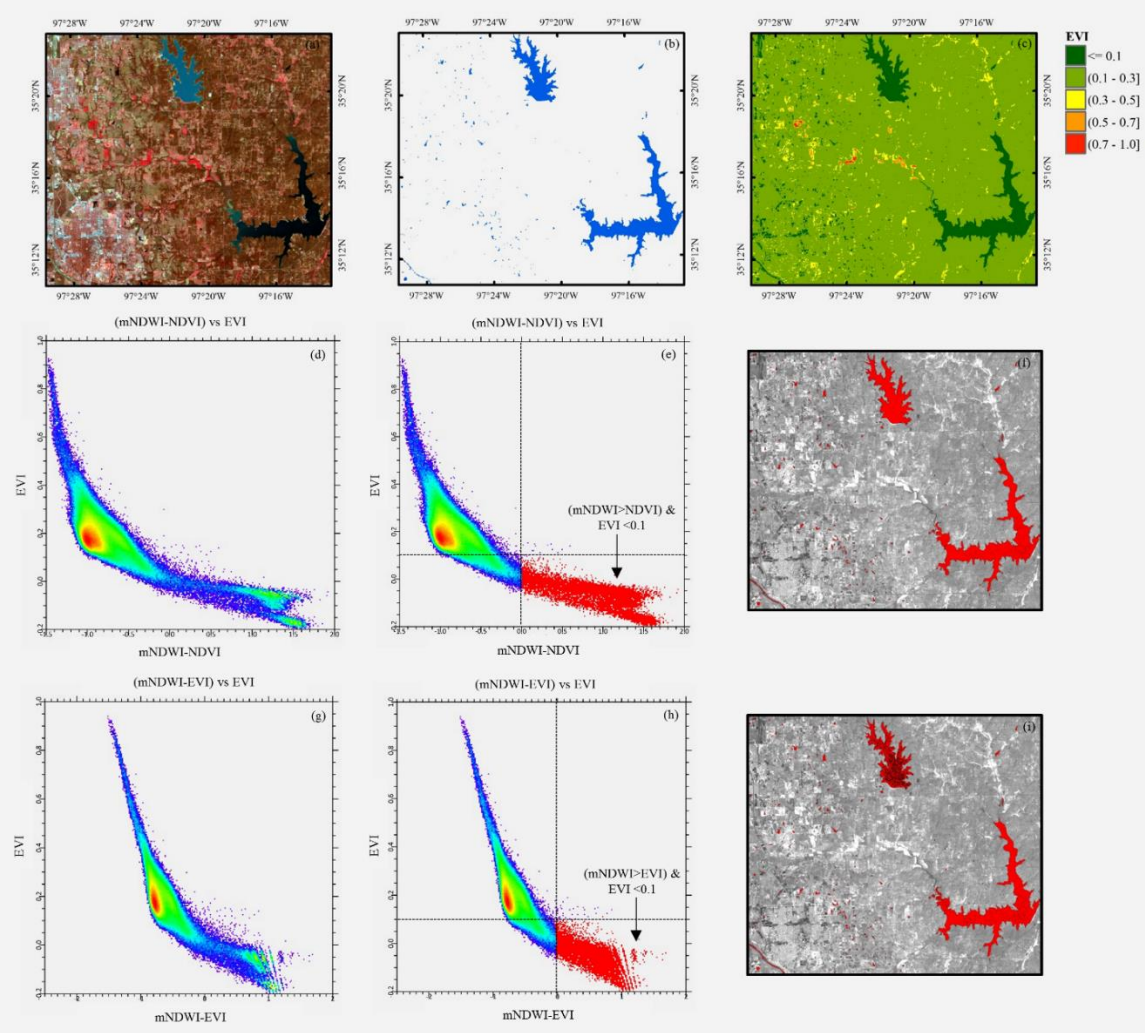


Figure S2.4 Water detection in built-up area. (a) Landsat 7 surface reflectance image (list bands in the false color composite), (b) Final water detection in blue color ((mNDWI>NDVI or mNDWI>EVI) and (EVI<0.1)), (c) EVI, (d) Scatter plot ((mNDWI-NDVI) vs EVI), (e) Scatter plot ((mNDWI-NDVI) vs EVI) with water detection marked red, (f) Surface water in red corresponding to water detection in scatter plot e, (g) Scatter plot ((mNDWI-EVI) vs EVI), (h) Scatter plot ((mNDWI-EVI) vs EVI) with water detection marked red, (i) Surface water in red corresponding to water detection in scatter plot h.

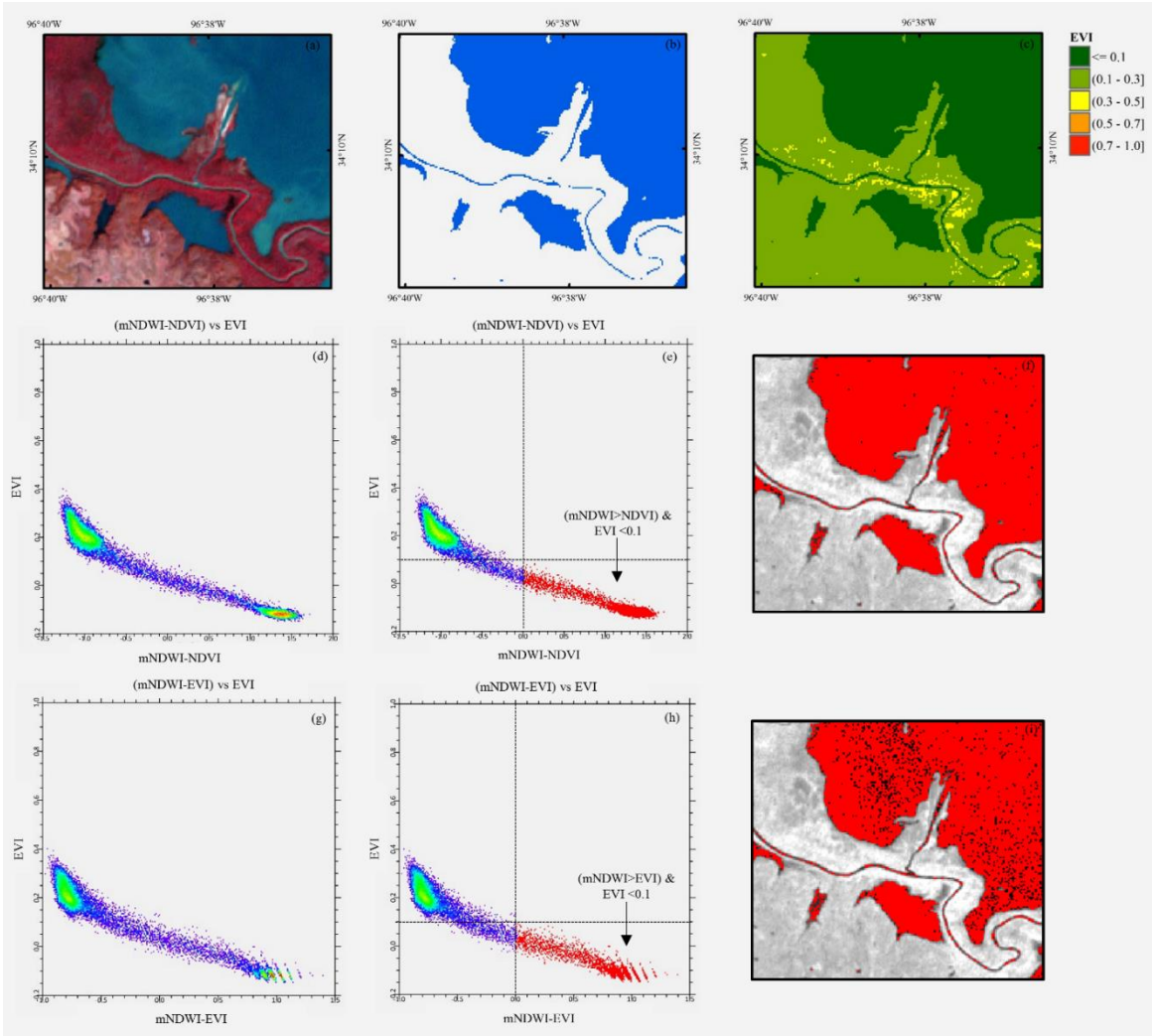


Figure S2.5 Water detection in vegetated area. (a) Landsat 7 surface reflectance image (list bands in the false color composite), (b) Final water detection in blue color ((mNDWI>NDVI or mNDWI>EVI) and (EVI<0.1)), (c) EVI, (d) Scatter plot ((mNDWI-NDVI) vs EVI), (e) Scatter plot ((mNDWI-NDVI) vs EVI) with water detection marked red, (f) Surface water in red corresponding to water detection in scatter plot e, (g) Scatter plot ((mNDWI-EVI) vs EVI), (h) Scatter plot ((mNDWI-EVI) vs EVI) with water detection marked red, (i) Surface water in red corresponding to water detection in scatter plot h.

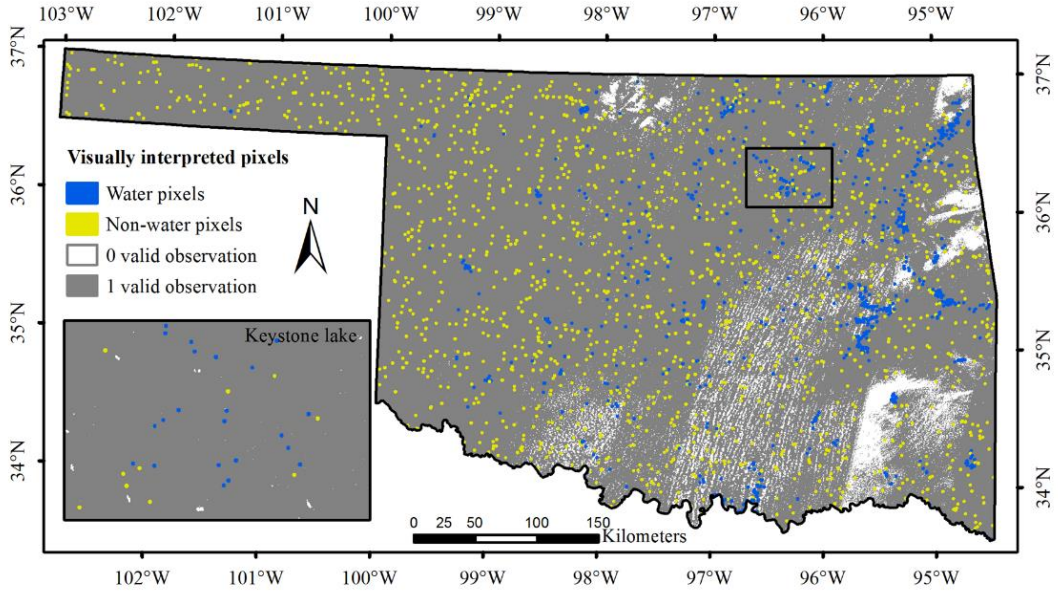


Figure S2.6 Visually interpreted water and non-water pixels.

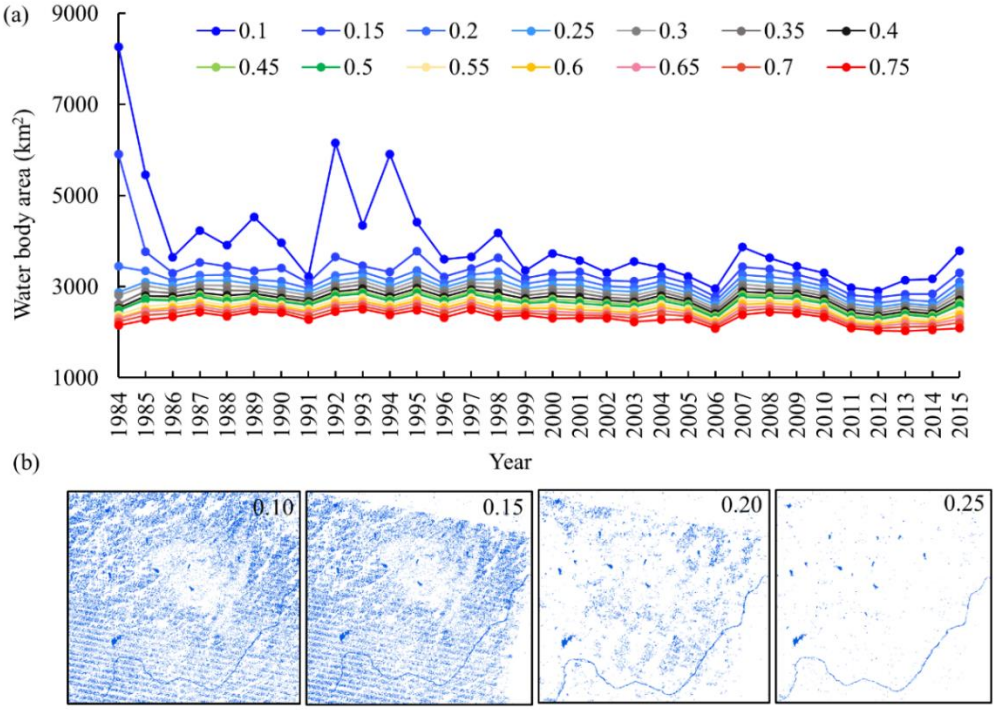


Figure S2.7 Water body frequency threshold selection, (a) maximum water body area using 14 different water body frequency thresholds, (b) Noise conditions of insets in the 1984 maximum water body maps using different water body frequency thresholds: 0.1, 0.15, 0.2, 0.25.

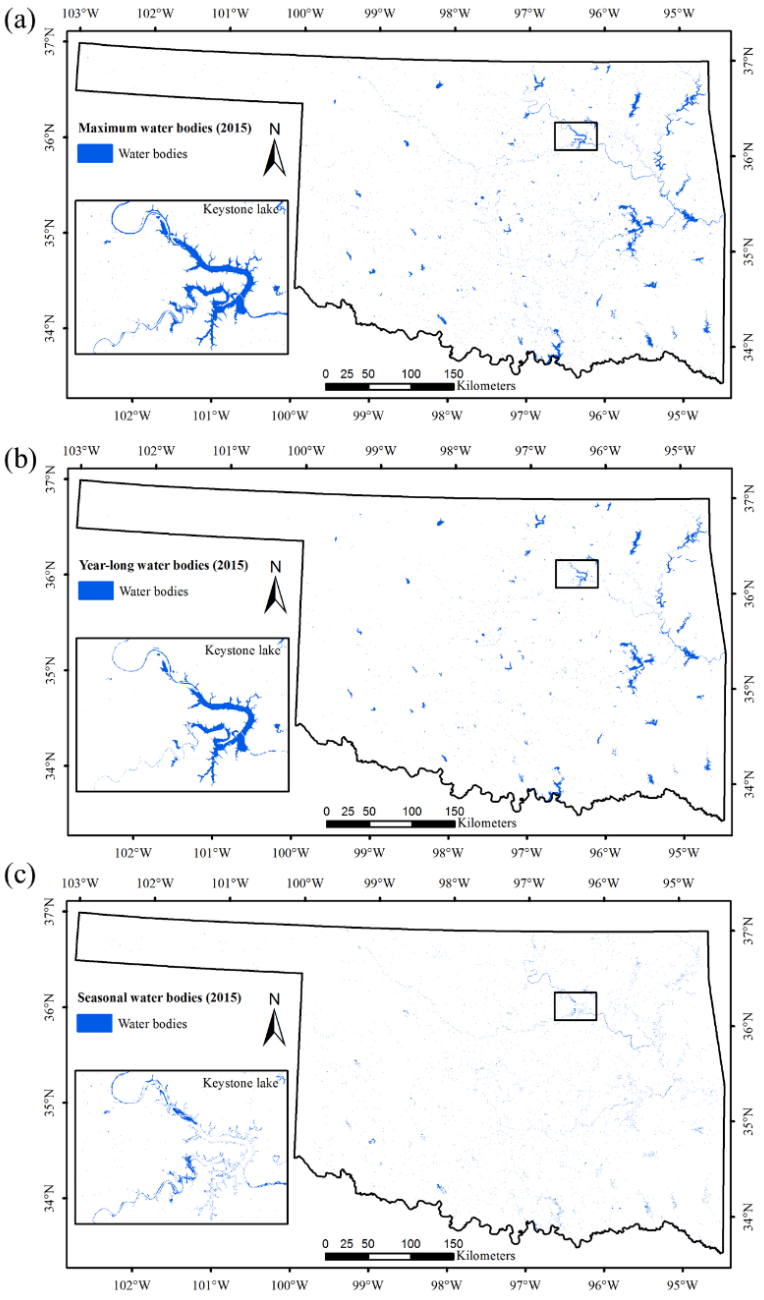


Figure S2.8 Open surface water body maps in 2015: (a) maximum water body map, (b) year-long water body map, (c) seasonal water body map.

Table S2.1 Water indices

Index	Definition	Source
Band-ratio	Band-ratio = Red/NIR [a]	(Boland, 1976)
Normalized difference vegetation index	NDVI = (NIR - Red) / (NIR + Red)	(Rouse Jr et al., 1974)
Normalized difference water index	NDWI = (Green - NIR)/(Green + NIR)	(McFeeters, 1996)
Modified normalized difference water index	mNDWI = (Green - SWIR1)/(Green + SWIR1)	(Xu, 2006)
Land surface water index	LSWI = (NIR - SWIR1)/(NIR + SWIR1)	(Xiao et al., 2002)
Difference between vegetation and water	NDVI - NDWI	(Gond et al., 2004)
Tasselled cap wetness	TCW = 0.0315×Blue+0.2021×Green+0.3102×Red-0.1594×NIR-0.6806×SWIR1-0.6109×SWIR2	(Crist, 1985)
Surface wetness index	SWI = 0.2626×Blue+0.2141×Green+0.0926×Red-0.0656×NIR-0.7629×SWIR1-0.5388×SWIR2	(Bhagat and Sonawane, 2011)
Automated water extraction index without shadows	$AWEI_{nsh}$ = 4×(Green-SWIR1)-(0.25×NIR+2.75×SWIR2)	(Feyisa et al., 2014)
Automated water extraction index with shadows	$AWEI_{sh}$ = Blue + 2.5×Green-1.5×(NIR+SWIR1)-0.25×SWIR2	(Feyisa et al., 2014)
Enhanced Water Index	EWI = (Green-SWIR1+m)/[(Green + SWIR1)×(NDVI + n)]	(Wang et al., 2015)
Water Index $_{2015}$	WI_{2015} = 1.7204 + 171×Green+3×Red-70×NIR-45×SWIR1-71×SWIR2	(Fisher et al., 2016)

[a] Landsat 5/7 bands: Blue (0.45-0.52 μm), Green (0.52-0.60), Red (0.63-0.69), NIR (0.77-0.90), SWIR1 (1.55-1.75), SWIR2 (2.09-2.35)

Table S2.2 The confusion matrix for evaluating the single-temporal water body map of Oklahoma in 2010.

	Ground Reference pixels			
Single-temple water body map (2010)	Water	Non-water	Sum of classified pixels	User accuracy (%)
Water	2680	71	2751	97.42%
Non-water	536	6655	7191	92.55%
Sum of ground reference pixels	3216	6726	9942	OA=94%
Producer accuracy (%)	83.33%	98.94%		Kappa=0.86

Table S2.3 The confusion matrix for evaluating the maximum water body maps using GIW datasets (2000) and NLCD datasets (2001, 2006 and 2011).

		Maximum water body maps (2000, 2001, 2006, 2011)			
Reference datasets		Water (km^2)	Non-water (km^2)	Reference Area (km^2)	Agreement (%)
GIW 2000	Water (km^2)	2569	264	2833	90.68%
	Non-water (km^2)	461	177720	178181	99.74%
	Estimated area (km^2)	3030	177984	181014	
	Agreement (%)	84.78%	99.85%	OA=99.60	Kappa=0.87
NLCD 2001	Water (km^2)	2778	865	3642	76.26%
	Non-water (km^2)	247	177125	177372	99.86%
	Estimated area (km^2)	3024	177990	181014	
	Agreement (%)	91.84%	99.51%	OA=99.3%	Kappa=0.83
NLCD 2006	Water (km^2)	2428	1091	3518	69.00%
	Non-water (km^2)	161	177334	177496	99.91%
	Estimated area (km^2)	2589	178425	181014	
	Agreement (%)	93.77%	99.39%	OA=99.3%	Kappa=0.79
NLCD 2011	Water (km^2)	2418	1094	3512	68.86%
	Non-water (km^2)	201	177302	177502	99.89%
	Estimated area (km^2)	2619	178395	181014	
	Agreement (%)	92.34%	99.39%	OA=99.2%	Kappa=0.79

Table S2.4 Collinearity statistics of all input explanatory variables of each model in SPSS 19 using the "enter" method.

Maximum water body area		Maximum water body number		Year-long water body area		Year-long water body number		Seasonal water body area		Annual average water body area	
Variable	VIF	Variable	VIF	Variable	VIF	Variable	VIF	Variable	VIF	Variable	VIF
P	1.20	P	1.18	P	1.18	P	1.15	P	1.14	P	1.19
T	1.12	T	1.12	T	1.12	T	1.12	T	1.13	T	1.12
SWW	1.89	SWW	2.35	SWW	1.68	SWW	1.72	SWW	1.44	SWW	1.99
$MWBA_p$	1.75	$MWBN_p$	2.19	$YWBA_p$	1.54	$YWBN_p$	1.57	$SWBA_p$	1.34	$AAWBA_p$	1.83

The six dependent variables are maximum water body area (MWBA), maximum water body number (MWBN), year-long water body area (YWBA), year-long water body number (YWBN), seasonal water body area (SWBA) and annual average water body area (AAWBA). P and T are the statewide annual total precipitation and annual average temperature respectively. SWW is the statewide surface water withdrawal in million gallons per day. $MWBA_p$, $MWBN_p$, $YWBA_p$, $YWBN_p$, $SWBA_p$, $AAWBA_p$ are the water body status in the previous year. VIF is short for variance inflation factor, which is the index that measures how much the variance of an estimated regression coefficient is increased because of collinearity.

Chapter 3: Divergent trends of open surface water body area in the contiguous US from 1984 to 2016 from analysis of time series Landsat images

Abstract

The contiguous US (CONUS), especially the West, faces challenges of increasing water stress and uncertain impacts of climate change. The historical information of surface water body distribution, variation, and multi-decadal trends documented in remote sensing images can aid in water resource planning and management, yet it is not well explored. This study detected open surface water bodies in all Landsat 5, 7, and 8 images (~370,000 images, >200 TB) of the CONUS and generated 30-m annual water body frequency maps for 1984–2016. This study also analyzed the interannual variations and trends of year-long water body area, examined the impacts of climatic and anthropogenic drivers on water body area dynamics, and explored the relationships between water body area and land water storage (LWS). Generally, the western half of the US is prone to water stress with small water body area and large interannual variability. During 1984–2016, water-poor regions of Southwest and Northwest US had decreasing trends in water body area, while water-rich regions of Southeast US and far north Great Plains had increasing trends. These divergent trends, mainly driven by climate, enlarged water resource gaps and are likely to continue according to climate projections. Water body area change is a good indicator of LWS dynamics in 58% of the CONUS. Following the 2012 prolonged drought, LWS in California and Southern Great Plains had a larger decrease than surface water body area, likely caused by massive groundwater withdrawals. Our findings provide valuable information for surface water resource planning and management across the CONUS.

3.1 Introduction

Terrestrial open surface water bodies, including lakes, reservoirs, rivers, streams, and ponds, are critically important water resources for agriculture, aquiculture, industrial production, and aquatic and terrestrial ecosystems (Bates et al., 2008; Wood et al., 2011). Numerous open surface water bodies are distributed across the contiguous United States (CONUS), providing 99%, 57%, and 63% of the water used in thermoelectric-power production, agricultural irrigation, and public water supply, respectively (USGS, 2010). According to the water supply stress index (WaSSI) model, surface water stress was found in over 9% of the 2103 CONUS watersheds, mostly distributed in the western half of the US (Averyt et al., 2013). Climate change models predicted a general increase of water stress across the US, with the largest increases in Southwest US through 2050 (Blanc et al., 2014). Southwestern states experienced a spate of dryness in the early 21st century (MacDonald et al., 2008) and are projected to become drier and experience more severe droughts in the latter half of the 21st century by various climate and hydrology models (Cayan et al., 2010; Melillo et al., 2014; Scheff and Frierson, 2012). Water resource managers in the Western US face the challenges of adapting to unprecedented droughts and uncertain impacts of climate change (Miller and Piechota, 2011). The spatial distribution, temporal dynamics, and long-term trends of CONUS surface water bodies, documented in remote sensing images in the last three decades, can provide valuable information for water resource managers in water resource planning and management in coping with drought and climate change, yet the information has not been well explored.

Strong interannual variability of surface water bodies caused by severe drought events have substantially impacted US socioeconomic systems (Hall et al., 2014;

Hoerling et al., 2014). In July 2012, US nuclear-power production hit its lowest seasonal levels in nine years because of a water shortage and high water temperature (Argonne National Laboratory, 2012). The lack of timely rainfall and the scarcity of irrigation water in 2012 caused widespread crop failure across the Great Plains and Midwestern US (Wolf et al., 2016). Corn and soybean yields in 2012 were 26% and 10%, respectively, below the yields forecasted by the USDA at the beginning of the crop growing season (Hoerling et al., 2013). Many reservoirs in the arid and semi-arid western regions were depleted during the 2012 drought (Hoerling et al., 2013), and contingency plans were activated to maintain public water supply (Murti et al., 2016). Reduced water body area due to severe droughts also dramatically impacted ecosystems (Atkinson et al., 2014; Walls et al., 2013). For example, decreased pond water in Southeastern US led to the rapid decline of salamander occupancy from 22.3% in spring 2009 to 9.9% in fall 2012 (Walls et al., 2013). The drought-induced reduction of stream flow and water coverage of the Kiamichi River in Southeastern Oklahoma had substantially reduced the freshwater mussel abundance by over 60% from 1992 to 2011 (Vaughn et al., 2015). Although the consequences of strong water body variations are evident, the interannual variability and trends of open surface water body area across the CONUS in the last three decades have remained unknown.

The spatial distribution and temporal variation of open surface water bodies are affected by both climate and anthropogenic activities (Bates et al., 2008; Pekel et al., 2016). Precipitation and temperature are two dominant climatic factors that affect the changes of open surface water body area (Krueger et al., 2017; Tao et al., 2015). Various anthropogenic activities were also found related to the change of open surface water

bodies, including dam construction (Pekel et al., 2016), water withdrawals for public water supply (Zou et al., 2017), agricultural irrigation (Tao et al., 2015), thermoelectric power production (van Vliet et al., 2012), and coal, oil, and gas mining (Murray, 2013; Tao et al., 2015). Climate change and enhanced demand for public water supply, irrigation, and industrial production in the last three decades have affected US water resources (Bates et al., 2008; Melillo et al., 2014). However, how the climate and human development have affected the variability of surface water body area in individual states across the CONUS have not been examined.

The objective of this study was to fill the above-mentioned knowledge gaps by investigating the interannual variations and trends of surface water body area and how it is affected by climate and anthropogenic factors across the CONUS during 1984–2016. First, we used all available Landsat image archives (~ 370,000 images, >200 terabytes of data) and a spectral index- and pixel-based approach (Zou et al., 2017) to detect water bodies and generate annual frequency maps of surface water bodies of the entire CONUS. Second, using these annual frequency maps, we generated annual maps of year-long, seasonal, and ephemeral water bodies. Third, we analyzed the interannual variability of year-long water body area for each state by calculating their standard deviations and analyzed their multi-decadal trends through linear regressions. Fourth, multiple stepwise regression models were used to assess four primary factors that affect the interannual variability of water body area: precipitation, temperature, surface water withdrawal, and the water body area in the previous year. Fifth, the relationships between open surface water body area and land water storage (LWS), derived from the Gravity Recovery and Climate Experiment (GRACE) liquid water equivalent thickness (LWET) data (Wiese et

al., 2018), were analyzed. Finally, we investigated the impacts of severe drought events on temporal dynamics of surface water body area and LWS in California and the Southern Great Plains.

3.2 Materials and methods

3.2.1 Landsat image

This study used all Landsat 5, 7, and 8 surface reflectance images of the entire CONUS (~ 370,000 images, >200 terabytes of data) in the Google Earth Engine platform (Google Earth Engine, 2017), which were originally from USGS EDC (USGS, 2017a; b). The number of images used in a year ranged from 3501 in 1984 to 17409 in 2014, with more images after the launch of Landsat 7 in 1999 (Figure S3.1a). For each image, the CFmask band was used as a quality control band to remove the cloud, cloud shadow, and snow pixels. The solar azimuth and zenith angles of each image were used along with the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) (Farr et al., 2007) to simulate terrain shadows and remove them. The remaining pixels were considered as good observations that can be used for water body detection. The pixels with 0 good observations in a year account for 0.27% on average during 1984–1998 and 0.04% during 1999–2016 (Figure S3.1b). More than 99.95% of the pixels within the CONUS had a total number of good observations ≥ 33 in the last 33 years while the majority of the pixels have a total number of good observations ranging from 300 to 1200 (Figure S3.1c).

3.2.2 Data on land water storage, water withdrawal, and climate.

The GRACE monthly liquid water equivalent thickness (LWET) products during 2002–2016 were the anomalies relative to the 2004.0–2009.999 time-mean baseline (Wiese et

al., 2018). The mascon-set of 0.5-degree gain factors were applied to the LWET data over land before further analysis. Monthly LWET data were used to calculate annual average LWET data, which were then used as the land water storage (LWS) values to explore its relationships with surface water body area. State-level water withdrawals, gathered every five years by the US Geological Survey (USGS, 2010), were interpolated into annual water withdrawal data (Equation 2.5) and used as a predictor variable in the multiple stepwise regression models for interannual variations of water body area. Statewide annual precipitation and annual average temperature data were gathered from National Centers for Environmental Information (NOAA, 2017) and also used as predictor variables.

3.2.3 Water body detection.

The relationship between water and vegetation indices can be used to detect open surface water bodies (Dong et al., 2015; Xiao et al., 2006), and the water body mapping algorithm with analysis of time series Landsat images was reported in a study for Oklahoma, USA (Zou et al., 2017). Thirty-one out of 459 Landsat tiles that overlap with the CONUS were selected (Figure S3.2). 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 (Figure S3.2a). In terms of time distribution, the sampling plots were selected across the time range of each satellite (Figure S3.2b). 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 (Figure S3.2c). In terms of terrain, 5 sampling plots were in mountainous areas while 27 were in relatively flat areas (Figure S3.2d). 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 (Figure 3.1). 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 (Figure 3.1b). 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 Figure 3.1d and e, respectively.

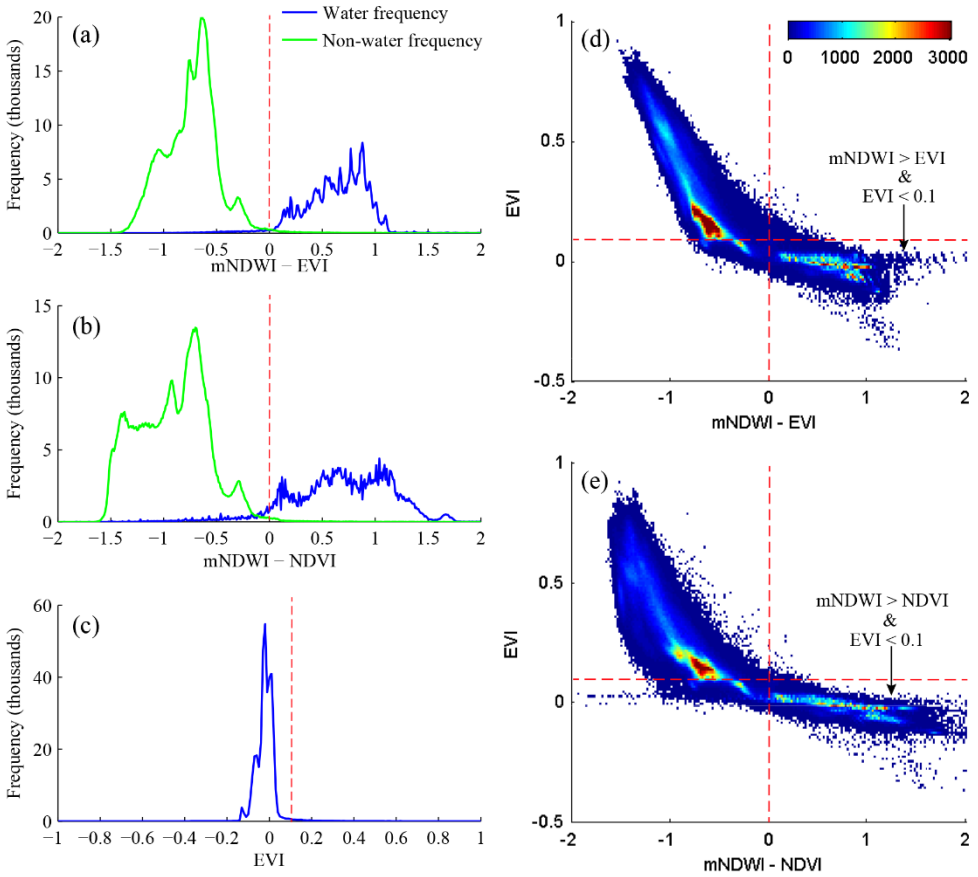


Figure 3.1 Spectral characteristics of 1.26 million sampling pixels. (a) Frequency distribution of (mNDWI-EVI) for water and non-water sampling pixels. (b) Frequency distribution of (mNDWI-NDVI) for water and non-water sampling pixels. (c) Frequency distribution of EVI for water sampling pixels. (d) 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.

3.2.4 Water body verification and application.

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 (Homer et al., 2015), 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 (Figure S3.3). 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 S3.1. 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 S3.1). 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 Figure S3.4. 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.

The water detection algorithms were performed on every good observation pixel in the ~370,000 Landsat images in the platform of Google Earth Engine, a cloud-based geospatial processing platform with large storage and processing power (https://developers.google.com/earth-engine/). For each pixel, its annual and 33-year water body frequency was defined as the ratio of water observations to total good observations (water and non-water observations) in a year and in 1984–2016, respectively. The 33-year water body frequency map was used to generate a non-water mask (33-year water body frequency<0.01) and a permanent water mask (33-year water body frequency ≥ 0.95), which were then applied to the annual water frequency maps to remove low-frequency noise caused by residual cloud, cloud shadow, and to fill the no data values. In each year, the annual water body frequency of pixels masked by the non-water mask and permanent water mask were set to 0 and 1, respectively.

3.2.5 Statistical analyses.

Based on annual water body frequency maps, year-long water body areas were calculated for each of the last 33 years. The interannual variability and trends of water body area during 1984–2016 by individual states and watersheds were calculated and analyzed

through linear regression models with t-test at the 5% significance level. The year-long water body areas within 0.5-degree grid cells were summed in each year of 2002-2016, and their linear relationships with LWS were examined in each of the 2818 0.5-degree grid cells across the CONUS. Multiple stepwise linear regressions were carried out in the platform of MATLAB R2014a to analyze the relationships between statewide year-long water body areas and four predictor variables, including annual precipitation, annual average temperature, annual surface water withdrawal, and the year-long water body area of the previous year. Water body area of the previous year was included as a predictor variable because of the legacy effect. No strong collinearity among the predictor variables was found (Table S3.2).

3.3 Results and Discussion

3.3.1 Water body frequency maps and water body areas

There were ~428 million 30-m pixels with annual water body frequencies > 0 in the CONUS in 2016 (Figure 3.2a), corresponding to ~385,000 km^2 maximum surface water body area. Water pixels with annual water frequencies ≥ 0.75 were defined as year-long water bodies while the other water pixels were classified as seasonal water bodies (≥ 5%) or ephemeral water bodies (< 5%) (Zou et al., 2017) (Text S1). There were about 285 million year-long water pixels (~257,000 km^2) within the CONUS in 2016, comprising the central portions of lakes, reservoirs, and large rivers (Figure 3.2b), which serve as the major sources for surface water withdrawals. The remaining 143 million seasonal and ephemeral water pixels (~128,000 km^2) are the small streams, ponds, and the edges of large surface water bodies (Figure 3.2b). The 33-year frequency map of surface water body over 1984-2016 had very similar spatial patterns to the annual water body frequency

map of 2016 (Figure 3.2b, d). According to the 33-year frequency map (Figure 3.2c), there were 277 million pixels with water frequencies ≥ 0.75, corresponding to ~250,000 km^2, which is close to the year-long surface water body area in 2016.

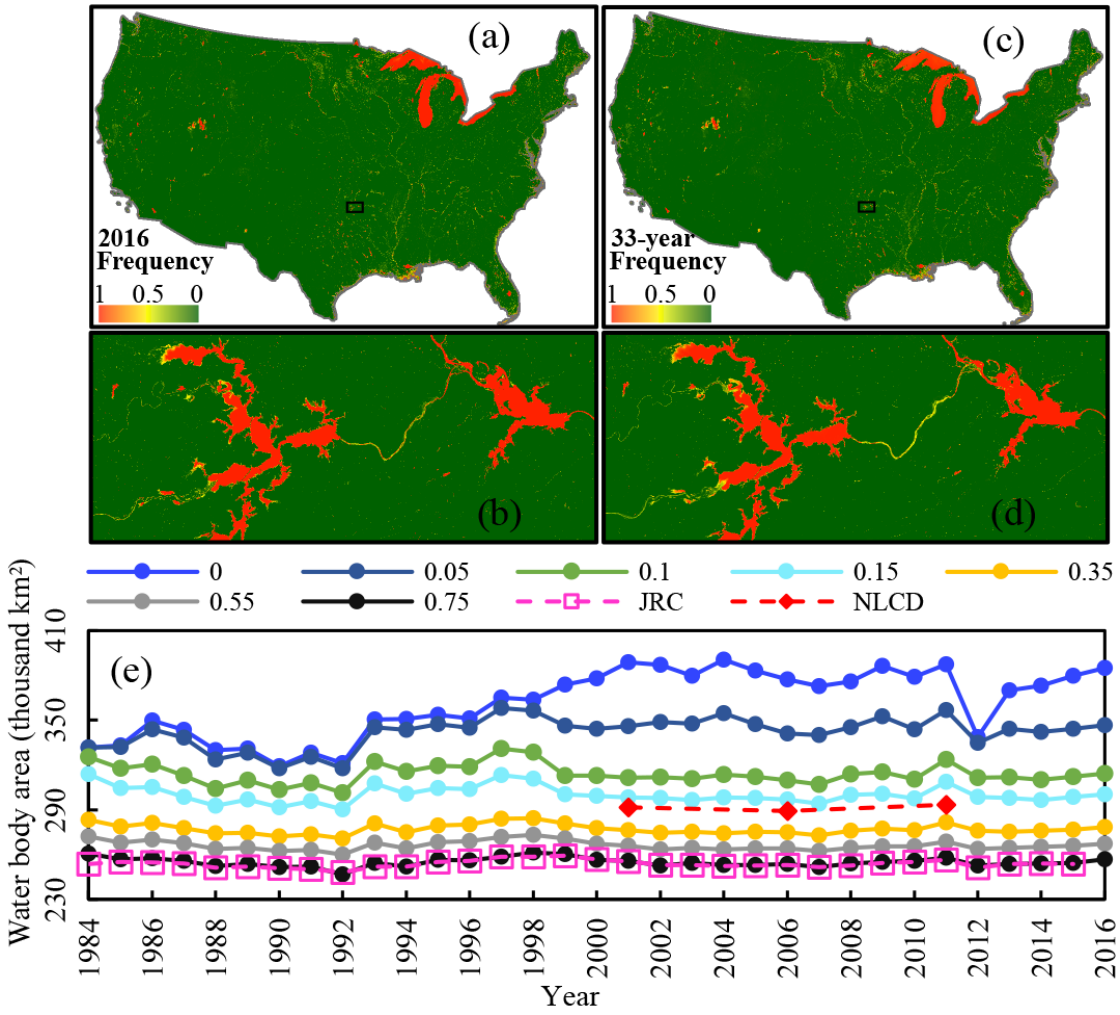


Figure 3.2 Water body frequency maps and water body areas using different frequency thresholds in the CONUS. (a) Annual water body frequency map in 2016 and (b) its zoom-in view of Eastern Oklahoma. (c) 33-year water body frequency map during 1984–2016 and (d) its zoom-in view of Eastern Oklahoma. (e) Total water body area of all pixels within CONUS with water body frequencies >0, ≥ 0.05, ≥ 0.1, ≥ 0.15, ≥ 0.35, ≥ 0.55, and ≥ 0.75, respectively, in our datasets, the permanent water body areas from the Joint Research Centre (JRC) dataset during 1984–2015, and the water body area from National Land Cover Database (NLCD) in 2001, 2006, and 2011. Water body frequency is the ratio of water body observations to total good observations in a year (a) or all 33 years during 1984–2016 (c).

Using different frequency thresholds can yield different estimates of surface water body areas (Figure 3.2e), such as the year-long (≥ 0.75), seasonal (0.05–0.75) and ephemeral (<0.05) water body areas during 1984–2016. At the CONUS scale, the year-long water body area varied from 246,641 km^2 to 261,328 km^2 in the last three decades, with small variability according to its standard deviation (2977 km^2). Year-long water body areas at the CONUS scale have no significant trends during 1984–2016. However, significant increasing trends were found in both seasonal (R^2=0.34, P<0.001) and ephemeral (R^2=0.59, P<0.001) water body areas, which might be related to the increase of large rainfall events and rainfall intensity in the CONUS (Z Feng et al., 2016; Melillo et al., 2014).

We compared the year-long water body areas in the CONUS from our dataset with those from the Joint Research Centre (JRC) (Pekel et al., 2016) and National Land Cover Database (NLCD) (Homer et al., 2015) (Figure 3.2e). The total area and interannual variability of year-long water bodies from our dataset agreed well with those from JRC permanent water bodies (Figure 3.2e), which were derived from analysis of all available Landsat top-of-atmosphere reflectance images during 1984–2015 (Pekel et al., 2016). It is interesting to note that total CONUS water body areas from the NLCD in 2001, 2006 and 2011 (Homer et al., 2015) were much higher than our year-long water body area and JRC permanent water body area (Figure 3.2e). The water body areas of NLCD could include some seasonal or ephemeral water bodies as the NLCD project used only a few good quality Landsat images taken during the plant growing season (Homer et al., 2015), which is often the wet season when more of the land surface is inundated with water. The good agreement between the year-long water body area from our dataset

and the permanent water body area from the JRC dataset clearly demonstrates the value of analyzing all available Landsat images in the study of land and water dynamics.

3.3.2 Water body distribution, variation, and driving factors

Surface water bodies are distributed unevenly across the CONUS with various interannual variabilities (Figure 3.3a). The average water body area (ha) per unit land (km^2) during 1984–2016 ranged from 0.2 ha/km^2 in Arizona and New Mexico to 40.7 ha/km^2 in Michigan, while its standard deviation ranged from 0.0001 ha/km^2 in Washington DC to 0.8ha/km^2 in Utah. Most of the Western half of the US have surface water body areas < 1 ha/km^2. Water resources in these regions have strong interannual variability based on the various standard deviations of water body area during 1984–2016 (Figure 3.3a). The Western, especially Southwestern, US was identified as a hotspot for water shortages in various hydrological model assessments and projections (Averyt et al., 2013; Gaupp et al., 2015; Strzepek and Boehlert, 2010). Water shortages in these areas were aggravated by large water withdrawals for agriculture (Caldwell et al., 2012) and thermal electric power plants (Melillo et al., 2014). Because of the limited water resources, many regions in the Southwest and Northwest US have to import water from beyond their watersheds (Good et al., 2014). The water body area in the Southeastern US is ~1 ha/km^2 higher than those of the Western half of the US and had relatively small interannual variations based on their small standard deviations (Figure 3.3a). Despite the abundance in water resources, the Southeastern US remains vulnerable to changes in water supply and demand (Melillo et al., 2014). Utah has a much larger water body area than its neighboring states because of the Great Salt Lake, and the states in the Great Lakes Region have the highest surface water body areas because of the Great Lakes.

Overall, the eastern half of the US has more water body area and less variability than the western half, which is similar with the annual precipitation pattern (Figure S3.5c).

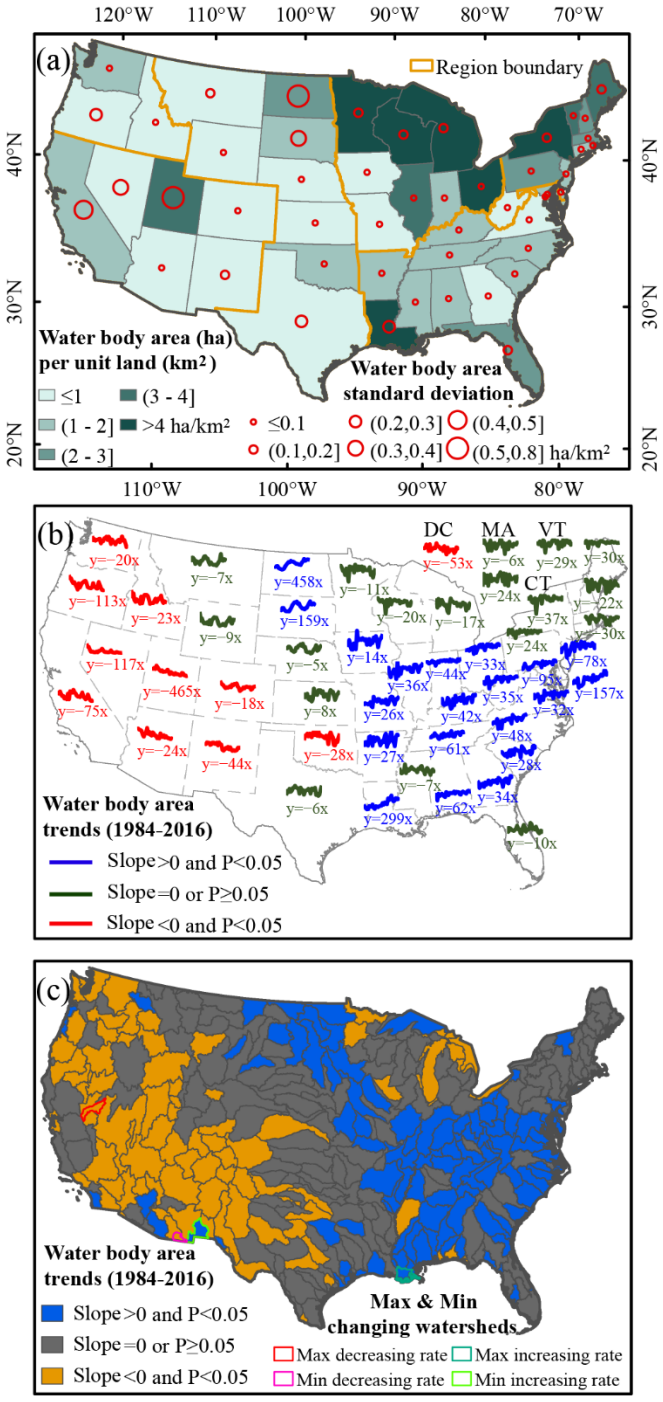


Figure 3.3 Year-long water body area and interannual trends in the CONUS during 1984–2016 by states and watersheds. (a) Average and standard deviation of year-long water body area (ha) per unit land (km^2) during 1984–2016 by states (the region

boundary was based on the Third National Climate Assessment (Melillo et al., 2014) The middle region is the Great Plains, while on its left are the Northwest and Southwest, and on its right are the Midwest, Northeast, and Southeast. The Western half of the US in this study consists of the Northwest, Southwest, and the Great Plains, while the Eastern half of the US consists of the Midwest, Northeast, and Southeast). Interannual trends of year-long water body area (m^2) per unit land (km^2) during 1984–2016 with t-test at the 5% significance level by states (b) and by watersheds (6-digit Hydrologic Units Code (HUC-6)) (c). The intercepts of simple linear regression models in (b) are not shown because of limited space. Slope is the coefficient of independent variable x, which is the year.

The year-long water body areas by individual states during 1984–2016 showed remarkably divergent trends over years (Figure 3.3b). All eight states in the Southwest and Northwest, plus Oklahoma and Washington DC, had significant decreasing trends in their year-long water body areas during 1984–2016. According to the slopes of simple linear regression models, the decreasing rates of year-long water body area (m^2) per unit land area (km^2) ranged from 18 m^2/km^2 per year in Colorado to 465 m^2/km^2 per year in Utah. In contrast, twenty states in the Southeast, far north Great Plains, and Southern Midwest had significant increasing trends in their year-long water body areas during 1984-2016. The increasing rates ranged from 14 m^2/km^2 per year in Iowa to 458 m^2/km^2 per year in North Dakota. The remarkable inter-annual divergent trends of year-long water body area were also found among the 336 watersheds within the CONUS (Figure 3.3c). Eighty-one watersheds, mostly in the Southwest and Northwest, had significant decreasing trends ranging from 3 m^2/km^2 per year in Rio De Bavispe Watershed of Arizona to 1355 m^2/km^2 per year in the Carson Watershed of Nevada (Figure 3.3c). Ninety-seven watersheds, mostly in the Southeast and far north Great Plains, had significant increasing trends, ranging from 3 m^2/km^2 per year in the Mimbres Watershed of New Mexico to 1799 m^2/km^2 per year in Central Louisiana Coastal Watershed of

Louisiana (Figure 3.3c). Most of the significant decreasing trends of surface water body area were found in states and watersheds that have relatively small water body areas and large interannual variabilities, while most of the significant increasing trends were found in states and watersheds that have relatively large water body areas and small variabilities. Thus, in general, the water-poor regions of Southwest and Northwest US were becoming poorer, while the water-rich regions of Southeast US and far north Great Plains were becoming richer over the last three decades.

Climate is the main factor contributing to the interannual variations of surface water body area. Annual precipitation was a significant variable in multiple stepwise regression models for most of states (Figure 3.4c), demonstrating its influence on water body area decrease in the Northwest and Southwest, and the water body area increase in the Southeast and far north Great Plains. Annual average temperature was a significant variable for the water body area decrease in the Northwest, as well as New Mexico, Oklahoma and Mississippi (Figure 3.4d). Surface water withdrawals showed significant influences in only a few states (Figure 3.4e), which could be caused by the water withdrawal regulations associated with water availability, and the infrequent water withdrawal data – reported every five years (USGS, 2010). Water body area in the previous year had significant positive impact in most multiple stepwise regression models, indicating strong legacy effects, especially in the western half of the US (Figure 3.4f). It is also worth noting that about 90% of the ~8000 major dams within the CONUS were constructed before 1984 (National Atlas of the United States, 2006), while the 735 dams constructed during 1984–2003 had an even spatial and temporal distribution (Figure S3.6). Thus, the observed divergent trends of open surface water body area during 1984–

2016 were largely driven by climate factors rather than by human water withdrawals or dam construction. The divergent trends are likely to continue in the future given the strong drying forecast in the Southwest and strong wetting forecast in the Eastern US by climate model simulations (Strzepek et al., 2015).

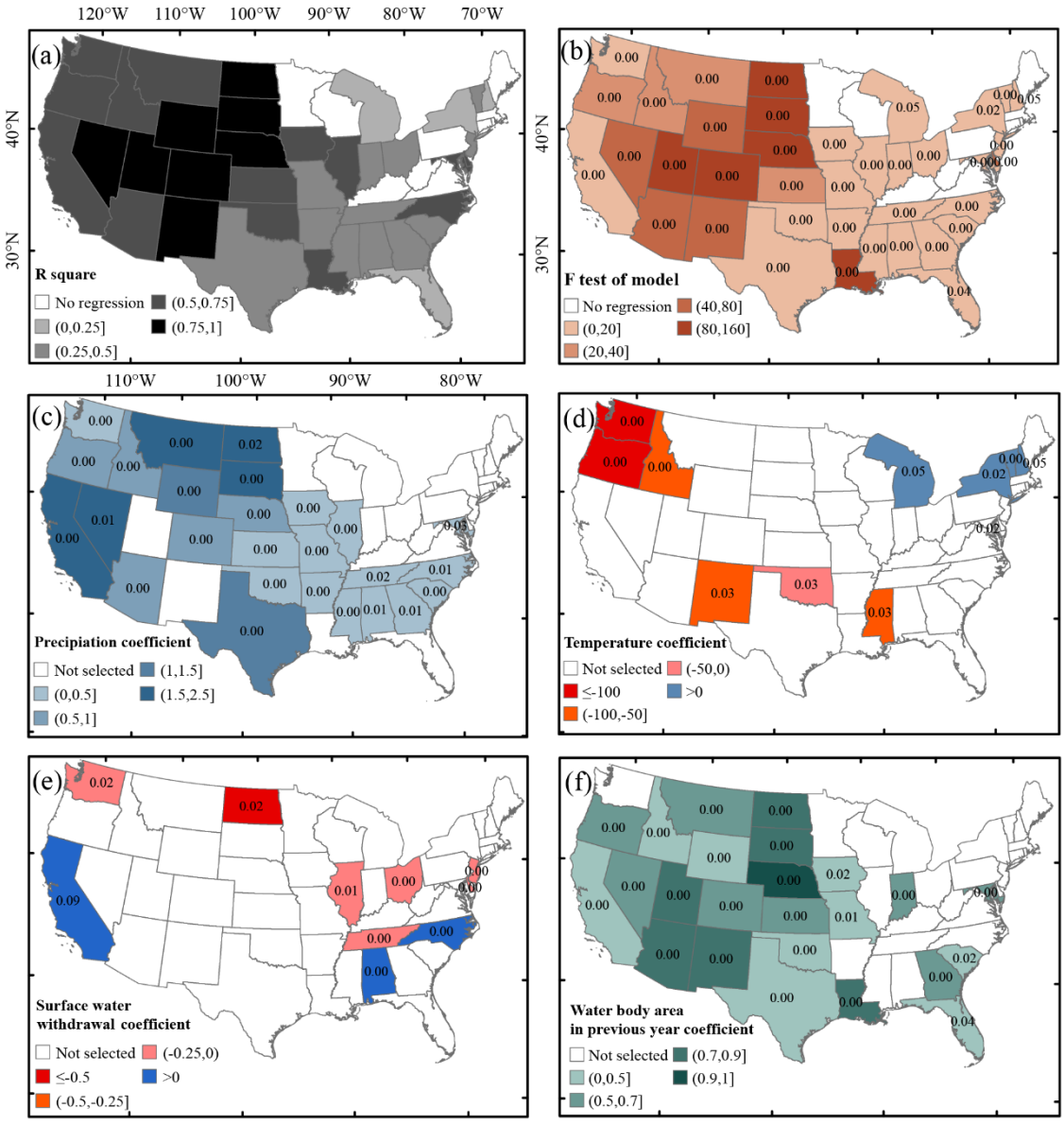


Figure 3.4 Multiple stepwise regression models. (a) R square of regression models. (b) F test and p value of regression models (the filled color is F, while the number is p value). Coefficient (slope) and p value of factor annual precipitation (c), annual average temperature (d), surface water withdrawal (e), and year-long water body area in the previous year (f) (the filled color is coefficient, while the number is p value).

3.3.3 Water body area and land water storage

We investigated spatial-temporal variability of year-long water body area in relation to land water storage (LWS) dynamics as observed by the GRACE satellite in the CONUS. GRACE LWS during 2002–2016 showed that large areas of ten states in the Southwest, Southern Great Plains, and north Midwest had significant decreasing trends (Figure 3.5a). In contrast, significant increasing trends of LWS were found in more than twenty states in the Northwest, Northern Great Plains, Midwest, and Northeast. The changes of LWS in a grid cell is affected by surface water body, soil moisture, groundwater, and water in vegetation (Famiglietti, 2004). We aggregated the annual maps of year-long water bodies at 30-m resolution into 0.5-degree (latitude and longitude) grid cells (Figure 3.5b). The trends of year-long water body area during 2002–2016 were more dispersive, with significant decreasing trends mostly distributed in California and northern Minnesota, and significant increasing trends mostly concentrated in the Northern Great Plains and Southeastern US (Figure 3.5b). The linear regression models between LWS and year-long water body areas during 2002–2016 showed significant positive correlations (Slope>0 and P<0.05) in 58% of the 2818 0.5-degree grid cells within the CONUS (Figure 3.5c, d), mostly in California, the Great Plains, and the Southeast. Open surface water bodies (lakes, reservoirs, rivers, and ponds) were found to be related to the dynamics of groundwater (Brunner et al., 2009) and total land water storage (Proulx et al., 2013). In the water body abundant Prairie Coteau (38,000 km^2), surface water bodies accounted for a significant fraction of GRACE LWS and improved the water budget closure estimation (Proulx et al., 2013). The significant positive correlations between

surface water body area and LWS suggested that the change in year-long water body area is a strong indicator of LWS dynamics.

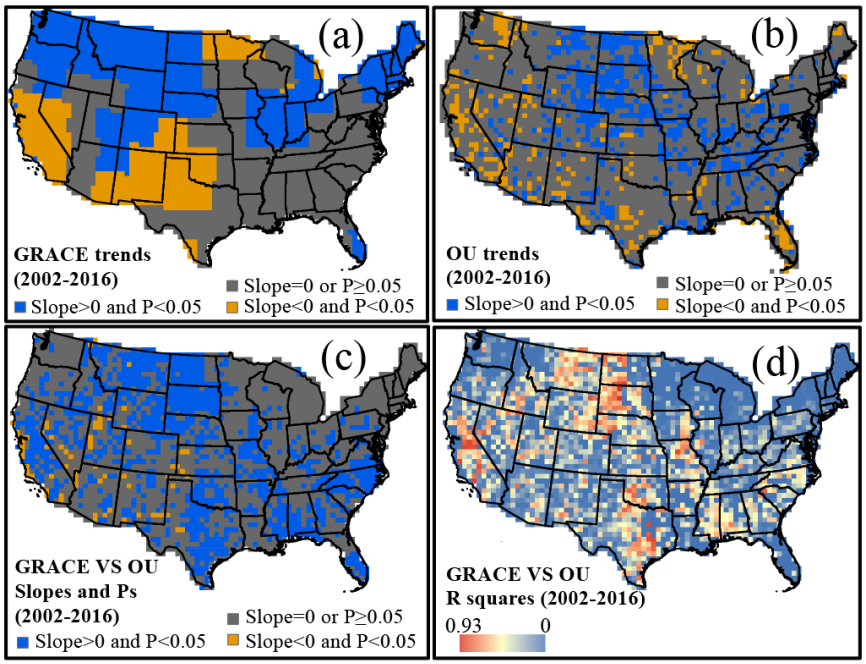


Figure 3.5 The slopes, p-values and r-squared values of simple linear regression models of GRACE land water storage and year-long water body area at 0.5-degree grid cells during 2002–2016 with t-test at the 5% significance level. (a) GRACE land water storage over years. (b) year-long water body area from our datasets (OU datasets) over years. (c) and (d) GRACE land water storage (dependent variable) over OU year-long water body area (independent variable).

3.3.4 The effects of drought on water body area and land water storage

Severe and prolonged droughts can substantially reduce surface water body area. California and the Southern Great Plains are among the top agricultural producing states, where surface water withdrawals play an important role in crop irrigation and livestock production (USGS, 2010). The interannual variation of year-long water body area in California clearly showed four multi-year dry and pluvial rotation events (Figure 3.6a), three of which corresponded to documented drought events in 1986–1992, 2007–2009, and 2012–2015 (California Department of Water Resources, 2015). Surface water body

areas had substantial drops at the beginning of drought events in 1986, 2007, and 2012, reached their lows at the end of these drought events in 1992, 2009, and 2015, and took several more years for the first two drought events to recover fully in 1995, and 2011. As of 2016, surface water body area had not yet recovered from the most recent drought event. The 2012–2015 drought caused the surface water body areas to drop to their lowest levels in 33 years (Figure 3.6a). Precipitation of California in 2016 was above the average value during 1984-2016, but it only resulted in only a minor recovery of surface water body area. In contrast, the Southern Great Plains were dominated by 1- to 2-year drought events. For example, the 2006 drought in Kansas, Oklahoma, and Texas (Dong et al., 2011), the 2011 and 2012 drought in Oklahoma and Texas, and the 2012 drought in Kansas (Hoerling et al., 2014). Surface water body areas in Kansas, Oklahoma, and Texas dropped in the dry year of 2006 and recovered quickly in the subsequent pluvial year of 2007 (Figure 3.6b-d). The Southern Great Plains suffered a prolonged drought that developed in 2011 and reached peak intensity in August 2012 (Hoerling et al., 2014). The year-long water body areas in these three states dropped during 2011–2012, stayed low through 2014, and recovered gradually to their normal condition in 2016. The Southern Great Plains had an extremely pluvial year of 2015 because of El Nino teleconnection (Wang et al., 2015), which aided the recovery of surface water body area.

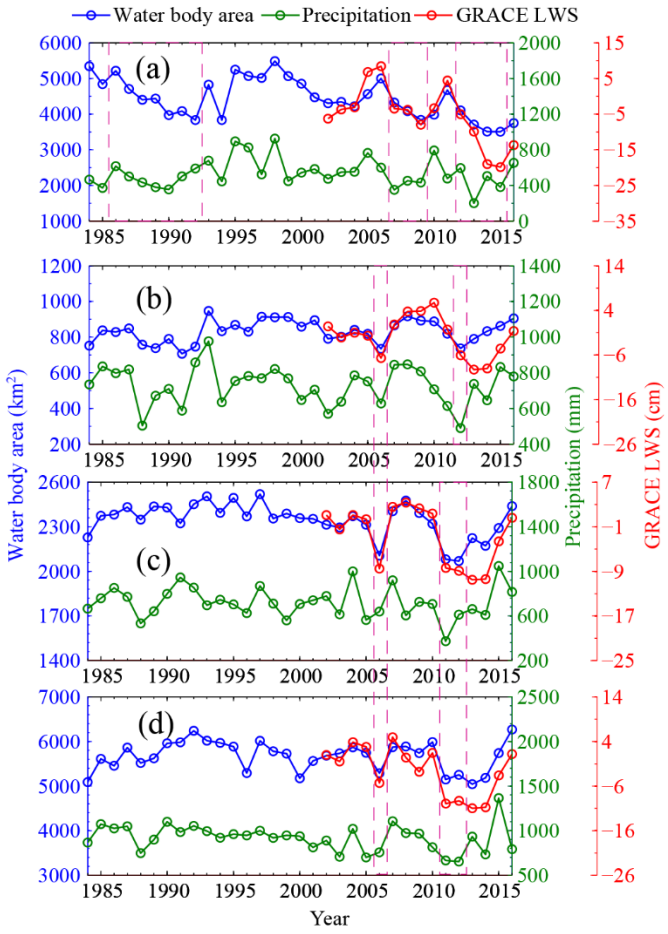


Figure 3.6 Interannual variations of year-long water body area, GRACE land water storage (LWS), and annual precipitation during 1984–2016 in California (a), Kansas (b), Oklahoma (c), and Texas (d). The dashed pink box indicates documented drought events, including 1986–1992, 2007–2009, and 2012–2015 droughts in California, 2006 drought in Kansas, Oklahoma and Texas, 2011–2012 drought in Oklahoma and Texas, and 2012 drought in Kansas.

Prolonged droughts can result in a larger decline of LWS than surface water body area. In California and the Southern Great Plains, the interannual variability of surface water body areas during 2002–2012 agreed well with those of LWS (Figure 3.6a-d). However, because of prolonged droughts, LWS had a much larger decline than surface water body areas in California and Southern Great Plains during 2014–2015 and 2013–2014, respectively. California Central Valley relied heavily on groundwater to mitigate

droughts (Howitt et al., 2014). The 2014-2016 droughts reduced surface water availability by 7.6 km^3/year and increased groundwater use by 6.2 km^3/year compared to an average water year (Howitt et al., 2014). The rate of groundwater decline in California Central Valley predicted by the water balance models (10.0 km^3/yr) that used a large amount of in-situ observations was quite close to that inferred from GRACE (11.2 km^3/yr) during the 2012-2016 droughts (Xiao et al., 2017). Observation data from 497 wells in the California Central Valley showed that droughts played a major role in the depletion of groundwater through increased well drilling and water extraction (Wang et al., 2016). The shrinkage of surface water bodies in drought years had forced water users to drill and mine groundwater (California Department of Water Resources, 2015; Thomas et al., 2017), which could have caused the larger decrease of LWS. Groundwater in California Central Valley was being pumped at far greater rates (20.4 mm yr^{-1}) than it can be naturally replenished, which may raise economic and food security challenges for the US (Famiglietti, 2014). The in-situ observation data from ~10,000 High Plains Aquifer wells also indicated severe drought induced groundwater declines in Southern and Central High Plains Aquifers in 2012 (Brena-Naranjo et al., 2014). Based on the water level data from 7460 wells during 2011 to 2013, the area-weighted, average water-level in the High Plains aquifer declined by 0.64 m, with major declines in the South and Central High Plains, Texas (1.1m), Kansas (0.9m), Colorado (0.7m), and Oklahoma (0.6m) (McGuire, 2014). Groundwater depletion in the irrigated Southern Great Plains and California Central Valley accounted for ~50% of groundwater depletion of the entire US since 1900 (Scanlon et al., 2012). With the low recharge rate in Central and Southern High Plains Aquifer, the current depletion rate would result in 35% of the Southern High Plains

lacking sufficient irrigation water in the next 30 years (Scanlon et al., 2012). The depletion of groundwater could in turn decrease the discharge to surface water bodies (Krueger et al., 2017), aggravating surface water scarcity in these regions. Although the findings in this study and the literatures demonstrated that farmers will switch from using surface water to using groundwater during the prolonged droughts. It is unclear when will the farmers switch. In other words, it is difficult to determine at what degree into the drought will the farmers decide to use groundwater.

3.4 Conclusions and perspective

Consistent with previous studies (Averyt et al., 2013; Blanc et al., 2014), the Western half of the US was identified as a hotspot of water stress with small water body area and large interannual variability in this study. Mainly driven by climate, year-long water body areas were shrinking in water-poor regions of Southwest and Northwest US but expanding in water-rich regions of the Southeast and far north Great Plains. These divergent changes have enlarged the water resource gaps across the CONUS in the last three decades. Thus, water resource management is becoming more and more challenging in Western US, especially during the 2012 prolonged droughts (Hoerling et al., 2013). Various climate and hydrological models have predicted the Southwest to be drier and face more severe droughts in the second half of the 21th century (Melillo et al., 2014; Scheff and Frierson, 2012; Strzepek et al., 2015), aggravating the challenges in water resource planning and management. The results from the analyses of historical Landsat images during 1984–2016 clearly shed new insight on the spatial distribution, temporal dynamics, and long-term trends of open surface water bodies in the CONUS and highlight the unoptimistic surface water body conditions in the Southwest and Northwest. These findings can be

used to assist decision makers and stakeholders across the CONUS, especially in the West, to develop and implement water resource planning and management in coping with the increasing water stress, unprecedented droughts, and uncertain impacts of climate change.

Supplementary materials

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 (USGS, 2017b). 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 (Google Earth Engine, 2017; USGS, 2017a; b). 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 (Pekel et al., 2016). 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 Figure 3.2e, 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 (Zou et al., 2017), 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).

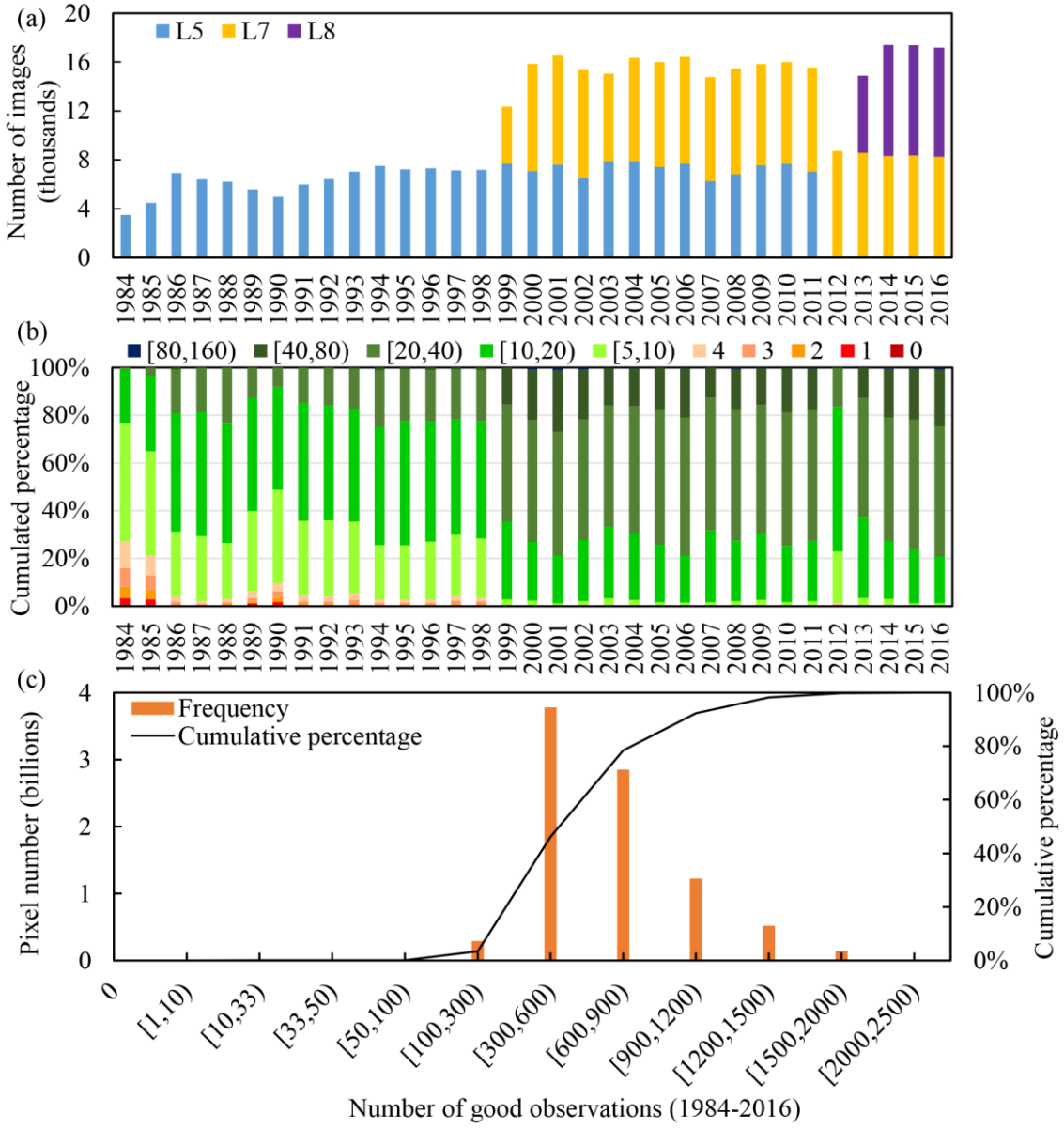


Figure S3.1 Landsat data availability within the CONUS during 1984-2016. (a) Number of Landsat 5, 7 and 8 images in each year. (b) Cumulative percentage of pixels with good observations of 0, 1, 2, 3, 4, [5, 10), [10, 20), [20, 40), [40, 80), and [80, 160), respectively. (c) Number of pixels with good 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.

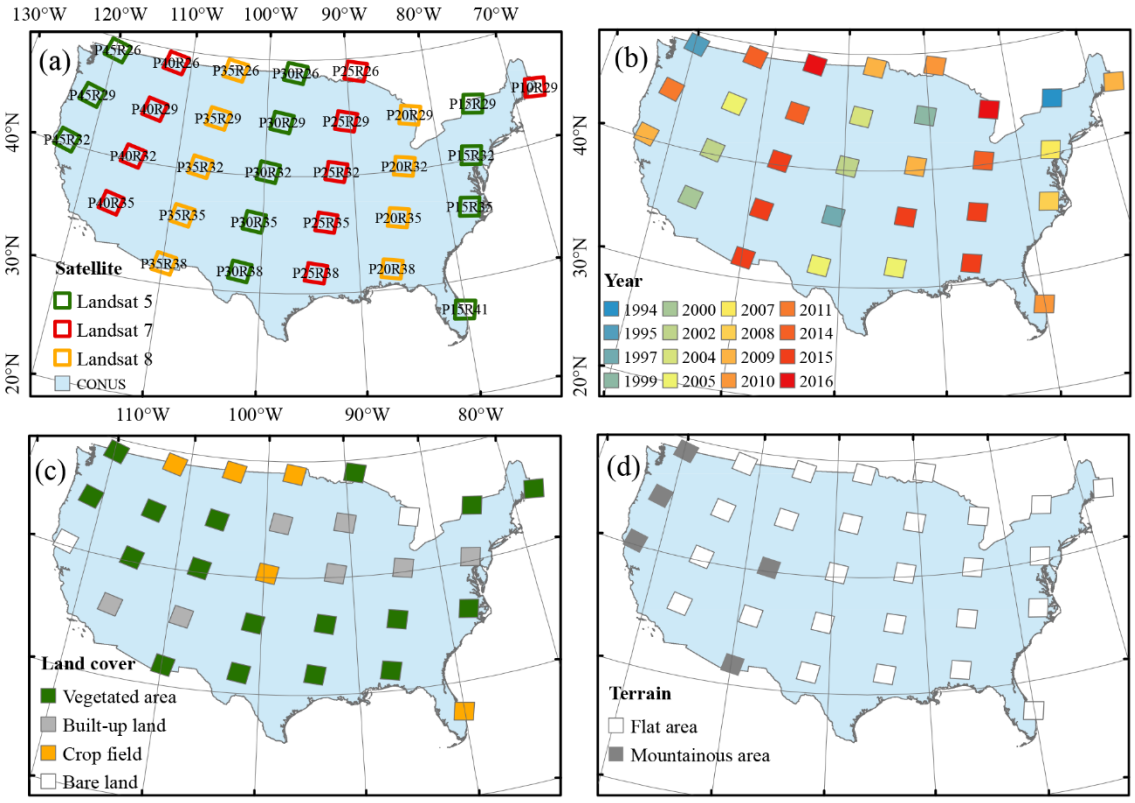


Figure S3.2 Landsat tile (path/row) and sampling plot distribution. (a) Distribution of selected Landsat tiles (path/row) by satellites. (b) Distribution of years of sampling plots. (c) Distribution of land cover types of sampling plots. (d) Distribution of terrain types of sampling plots. These sampling plots were used for visual delineation of water body and non-water features, which were further used to analyze their spectral characteristics.

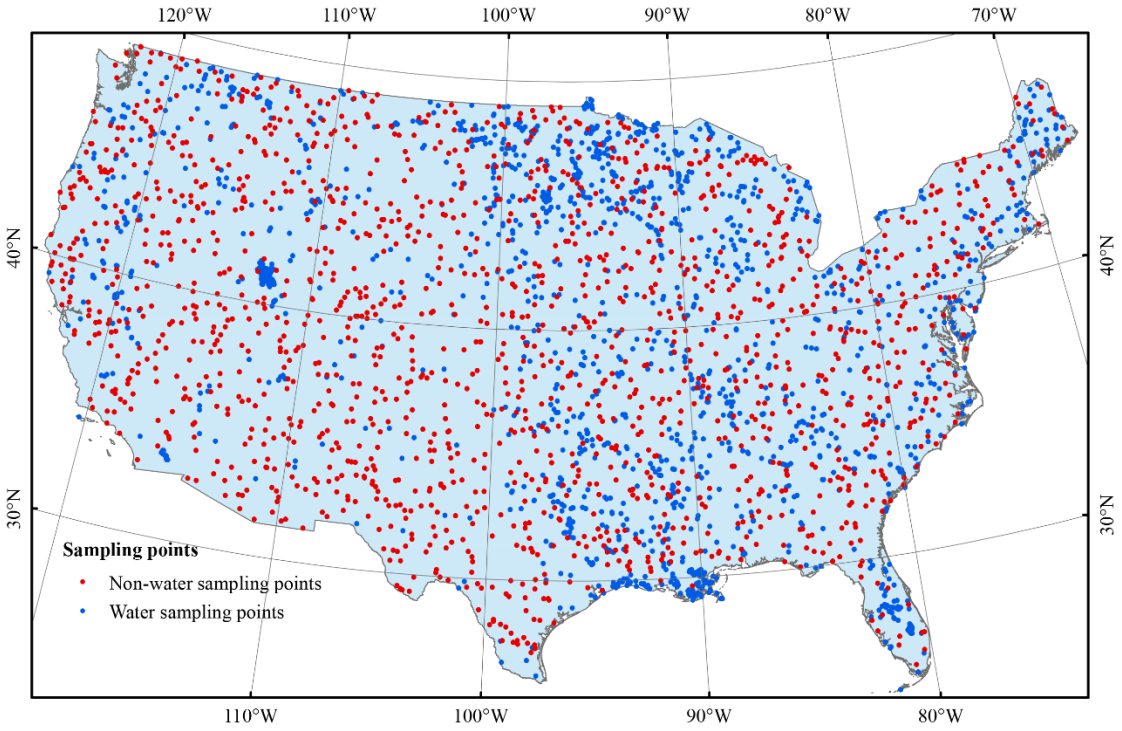


Figure S3.3 Spatial distribution of 3200 verification sampling points that were used to select ~12,000 sampling pixels for accuracy assessment.

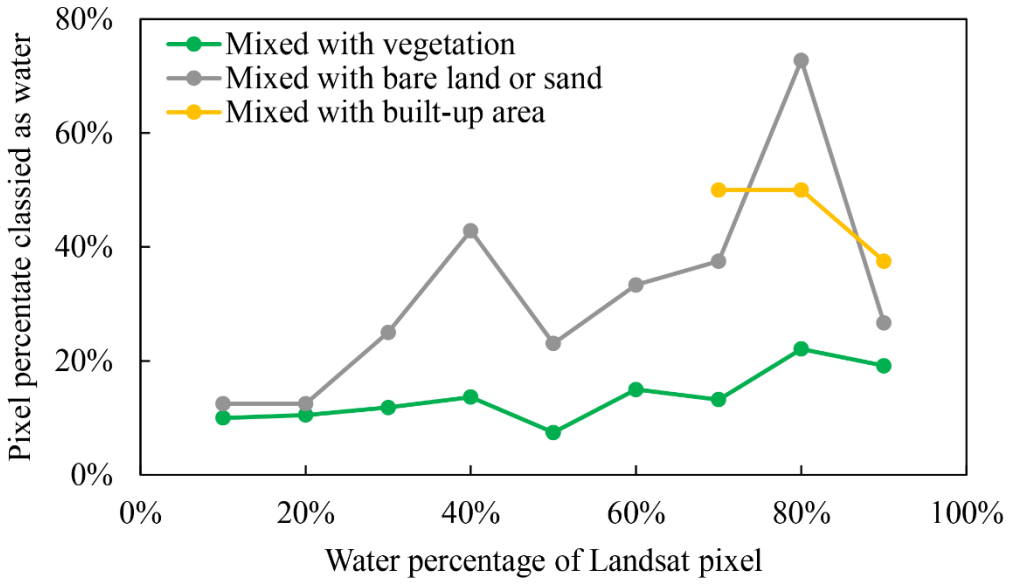


Figure S3.4 Mixed pixel water detection rate.

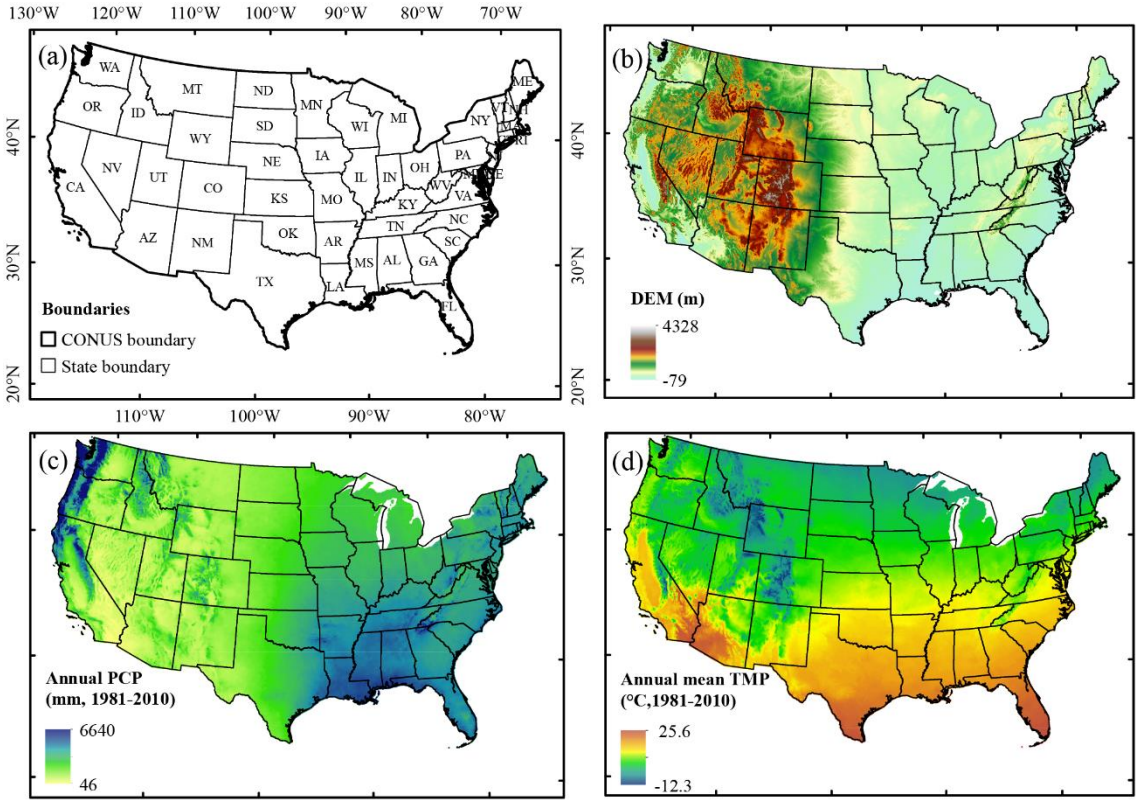


Figure S3.5 Study area. (a) State and CONUS boundary. (b) Digital elevation model (DEM). (c) Annual precipitation (PCP) during 1981–2010. (d) Annual mean temperature (TMP) during 1981–2010. DEM is from USGS EarthExplorer (http://earthexplorer.usgs.gov/), while the precipitation and temperature are from the Parameter-elevation Relationships on Independent Slopes Model (PRISM) Climate Group, Oregon State University (http://prism.oregonstate.edu/).

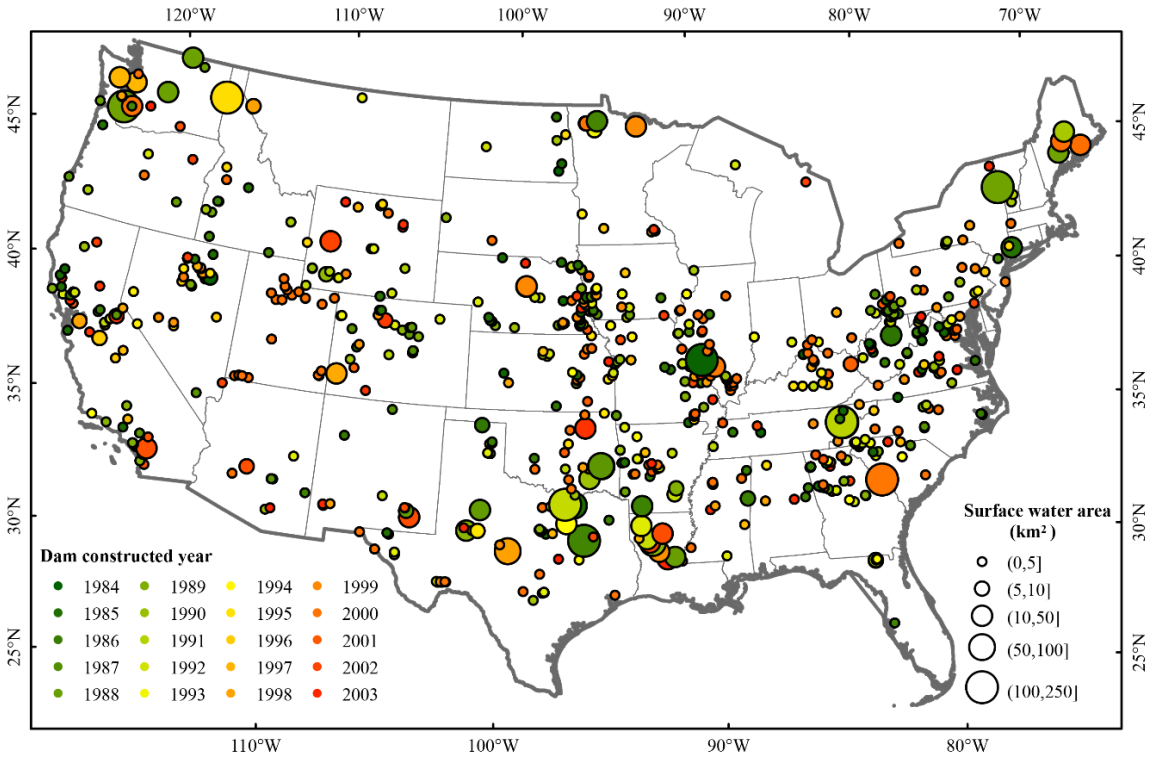


Figure S3.6 Spatial and temporal distribution of major dams constructed during 1984–2003 within the CONUS (National Atlas of the United States, 2006) (the circle filled color represents dam constructed year, while the circle size represents the surface water body area of the impoundment at its normal retention level).

Table S3.1 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 S3.2 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).

Chapter 4: Variations and trends of global surface water body area and land water storage in the past decades from analysis of time series Landsat images

Abstract

Changes in terrestrial water resources affect social, economic, and environmental systems, yet these changes have not been analyzed thoroughly across the globe using both surface water area and land water storage data. We generated a new global multi-decadal surface water body dataset at 30-m resolution using 3.8 million Landsat images during 1984–2017. About 8.5 million 0.01° grid cells had significant increasing or decreasing trends in surface water area over the past decades, forming interesting spatial patterns in northern Greenland, Tibetan Plateau, western US, the Great Lakes, Gulf of Bothnia, central South America, etc. The regional curves of surface water area at 5° tiles revealed its interannual variations, the magnitude of variability, and the multi-decadal trends. There were 189 and 170 tiles with significant increasing and decreasing trends, respectively. Divergent trends between land water storage and regional surface water area occurred in Greenland, China, the Indus Basin, and central Africa, mainly driven by climate and anthropogenic activities. Our dataset and findings uncovered unprecedented spatial and temporal details of global water resource dynamics, critical for water resource research, planning, and management.

4.1 Introduction

Terrestrial water resources are critical for human society, natural environment, and global biodiversity (Pekel et al., 2016; Rodell et al., 2018; Vorosmarty et al., 2010). Global population increased from about 5 billion in 1984 to 7.6 billion in 2017 (DESA, 2017). Population growth and increased food demand have stressed and modified global water systems through water diversion for irrigation, dam construction, land reclamation, and

groundwater mining (Rodell et al., 2018; Vorosmarty et al., 2010). Under the impacts of climate variability and anthropogenic activities, strong variations and significant changes of water resources, such as the dry-up of surface water bodies and the loss of groundwater, has threatened municipal water supplies, reduced crop production, and decreased aquatic biodiversity (Zou et al., 2018).

Global surface water bodies were recently mapped using Landsat images in a single year (M Feng et al., 2016; Verpoorter et al., 2014), multiple years (Liao et al., 2014; Yamazaki et al., 2015), and multi-decadal period (Pekel et al., 2016). Currently, there has been only one global multi-decadal surface water body dataset (1984-2015), which was released by the Joint Research Center (JRC) in 2016 (Pekel et al., 2016). The variations and trends of surface water area in a few regions were reported in some selected years where the unobserved area was less than 5% (Pekel et al., 2016). To date, the time-series variations and trends of surface water area in the past decades were not investigated thoroughly at local and regional scales across the globe. Land water storage, derived from the Gravity Recovery and Climate Experiment (GRACE) satellite mission, can be used to reveal the integrated change of land water resources, including surface water, groundwater, soil moisture, snow, and ice (Rodell et al., 2018). A previous study used GRACE land water storage data during 2002–2016 to illustrate emerging trends in freshwater availability across the globe (Rodell et al., 2018). To date, no studies have investigated surface water area and land water storage jointly to better characterize the spatial distribution, temporal variability, and multi-decadal trends of global water resources.

We used 3.8 million Landsat images during 1984–2017 to generate a new global multi-decadal surface water body dataset at 30-m resolution, with a slightly longer time series, a slightly larger spatial extent, and more Landsat data input compared to the JRC dataset (Pekel et al., 2016). Global surface water body frequency and area were derived from the new dataset and compared with those from JRC. Globally, we detected locations where surface water area had significantly changed during 1984–2017 in 0.01°grid cells. We analyzed and visualized the interannual variations and multi-decadal trends of regional surface water area in 5° tiles. Multiple stepwise regression models were used to examine how varying precipitation and temperature impact the multi-decadal dynamics of regional surface water area. Finally, we assessed the interannual variations and trends of both regional surface water area and land water storage during 2002–2016 to shed new light on the consistency and divergence between surface and land water resource dynamics.

4.2 Materials and methods

4.2.1 Data.

This study used all Landsat 5, 7, and 8 calibrated top-of-atmosphere (TOA) reflectance images from Collection 1 Tier 1 in Google Earth Engine (GEE) (Chander et al., 2009) to identify and map terrestrial open surface water bodies within 60°S–85°N. A total of ~3.8 million Landsat images (~ 2.2 petabytes of data) during 1984–2017 were used in this study (Figure S4.1), which is ~24% more than the images (3.1 million) used in the JRC study (Pekel et al., 2016), or 7% (~215,000 images) more in the overlapping study period 1984–2015. Some of the data gaps in the JRC dataset were filled with data brought from the international image receiving stations by the United State Geological Survey Landsat

Global Archive Consolidation (Wulder et al., 2016) (Figure S4.2). The Quality Assessment 16-bit Band (US Geological Survey, 2018) in each image was used to remove cloud, cloud shadow, snow/ice, designated fill, and radiometric saturation. The solar azimuth and zenith angles in each image were also used along with the digital elevation models (Farr et al., 2007; Hulley et al., 2015) to simulate terrain shadows and remove them. The remaining pixels were considered as good observations suitable for surface water body detection.

Given that the revisit cycle of the Landsat satellites is 16 days and there are two satellites in most of the study period, we divided a year into forty-six 8-day intervals (day 1-8, day 9-16…). For each 8-day interval, a binary quality mask was set as good if there are any good observations within the interval. Forty-six binary quality masks in a year were converted into one 64-bit quality band at 30-m resolution and output from GEE to local high-performance computers. In the last 34 years, central Greenland, mountainous permafrost, and tropical cloudy regions had no or small number of good observations (Figure S4.1). Approximately 93% of the total ~212 billion 30-m pixels across the global land surface had at least fifty 8-day intervals with good observations during 1984-2017 (Figure S4.1).

The Gravity Recovery and Climate Experiment (GRACE) monthly liquid water-equivalent thickness (LWET) product at 0.5° resolution were anomalies relative to the time-mean baseline of 2004.0–2009.999 (Watkins et al., 2015). Monthly LWET from April 2002 to December 2016 were used to calculate annual average LWET, which were used as annual land water storage values to analyze interannual dynamics of total land water resources. The year 2017 was not included because of the failure of GRACE sensor.

In calculation of annual average land water storage, the data gaps in the first three months of 2002 were filled by the data in the same months of 2003.

The Moderate Resolution Imaging Spectroradiometer (MODIS) monthly average land surface temperature product (MOD11C3) during 2000–2017 in the spatial resolution of 0.05° (Wan et al., 2015) were resampled into 0.01°, in an effort to estimate the number of months in a year surface water was in a liquid state at high-latitude or high-altitude regions. Monthly mean precipitation rate at surface and monthly mean air temperature at 2-m in Gaussian grid of 192 by 94 from 1984 to 2017 were obtained from NCEP-DOE Atmospheric Model Inter-comparison Project (AMIP-II) reanalysis (R-2) (Kanamitsu et al., 2002). These datasets were interpolated into a geographic grid of 0.5°. The annual average values within 5° x 5° (latitude/longitude) tiles were calculated and used as climatic factors in the multiple stepwise regression models to explain the interannual variations of regional surface water area.

4.2.2 Algorithm development.

We developed the surface water body mapping algorithms based on the relationships between water index and vegetation indices [((mNDWI-NDVI) > 0 or (mNDWI-EVI) > 0) and (EVI < 0.1)] using surface reflectance images. The algorithms were applied in Oklahoma, US (Zou et al., 2017), and the contiguous United States (Zou et al., 2018). In this study, we modified the algorithms for Landsat TOA reflectance data because the surface reflectance data did not have the global coverage when we carried out the study. Systematic sampling method was used and 162 Landsat tiles (path/row) were selected over the global land area (Figure S4.3). Within each Landsat tile boundary, a very high-resolution image in Google Earth was randomly selected as reference image, and then a

Landsat TOA image, which has an acquisition date close to the date of very high-resolution image, was also selected. One or two sampling blocks, covering half water and half non-water features, were extracted from the selected Landsat image. Approximately 10,000 sampling pixels were extracted from one Landsat tile. In eight Landsat tiles, sampling blocks cannot be obtained for reasons such as no Landsat images in some ocean islands, no liquid water in icecaps, and no water in deserts. A total of 157 sampling blocks from 154 Landsat tiles were obtained, and they were distributed across different terrains, land cover types, altitudes, and satellites (Landsat 5, 7, and 8) (Table S4.1).

All water and non-water features in the sampling blocks were manually delineated in reference to the very high-resolution image in Google Earth. There were 1,641,461 sampling pixels delineated from the 157 sampling blocks, with 709,912 water pixels, and 931,549 non-water pixels. According to the frequency distribution curves of water and non-water sampling pixels (Figure S4.3), our previous algorithms were modified into [((mNDWI-EVI) > 0.25 or (mNDWI-NDVI) > 0.25) and (EVI < 0.1 or NDVI < 0.1)]. Compared with the previous algorithms, the thresholds were changed because of the use of TOA data instead of surface reflectance data. Also, NDVI < 0.1 was added to the algorithms because some EVI values of surface water bodies, derived from a few Landsat 8 TOA images, showed anomalies while the corresponding NDVI values were normal (Figure S4.3). Besides the changes of some thresholds, the essence of the algorithms remains the same with the Oklahoma and CONUS studies, which is that only pixels with stronger water signal than vegetation signal were classified as water pixels and the vegetation noise was further removed. Running these algorithms on the above visually

interpreted 1,641,461 pixels resulted in a water detection producer accuracy of 98.3%, a water detection user accuracy of 97.33%, and an overall accuracy of 98.1% (Table S4.2).

The world was divided into 2,592 tiles (5° x 5°) and 1,076 tiles were included in the study because other tiles had no land area. The water detection algorithms were run tile by tile in GEE. For each 8-day interval, a binary water mask was set as water if there are any water detections from the good observations within this interval. Forty-six binary water masks in a year were converted into one 64-bit water band at 30-m resolution and output from GEE to local high-performance computers. The entire output dataset included 36,584 water images and 36,584 quality images, covering the globe during 1984–2017.

4.2.3 Accuracy assessment.

The strategies of accuracy assessment were similar with those used in the JRC study (Pekel et al., 2016). We divided the entire Earth into 28,800 grid cells (1.5° x 1.5°, latitude/longitude) and removed the grid cells in ocean and central Greenland ice sheet. There were 8,904 grid cells overlapping with landmass, from which 8,904 sampling points were randomly selected, with an even distribution in non-water, seasonal water, and permanent water in reference of the JRC water occurrence map (Pekel et al., 2016). For each sampling point, a very high-resolution image from Google Earth or ESRI World Imagery was selected as a reference image. Then a Landsat image, whose acquisition time was close to the date of the very high-resolution image, was also selected. Four Landsat pixels around the sampling point were selected and water detection algorithms were performed to classify them into water or non-water pixels. The water coverage percentage within each Landsat pixel boundary was recorded in reference to very high-

resolution image, with 100% as pure water pixel, 0% as pure non-water pixel, and the rest (10%, 20%...90%) as mixed pixels. The validation sampling pixels distributed across the last three decades and across different satellites (Table S4.3). A total of 35,616 Landsat pixels were validated, out of which 15,255 were pure non-water pixels, 15,098 were pure water pixels, 3,506 were mixed pixels of water and non-water features, 381 were masked out by quality band, and 1,376 had no Landsat images (Figure S4.4). According to the confusion matrix for 30,353 pure water and non-water pixels (Figure S4.4), the algorithms had a water detection producer accuracy of 98.63%, a water detection user accuracy of 99.10%, and an overall accuracy of 98.86%. Including the 3,506 mixed pixels (mixed pixels with water percentage > 50% were classified as water, while ≤ 50% were classified as non-water), the producer accuracy, user accuracy, and overall accuracy were 93.44%, 97.36%, and 95.49%, respectively (Figure S4.4). Mixed pixels are one of the biggest challenges in water detection using remote sensing data. Specifically, mixed pixels are one of the major sources of omission error in water detection.

4.2.4 Surface water body maps and analysis.

The 8-day global surface water body map and good observation (quality) map at 30-m resolution can be derived by converting the annual 64-bit binary outputs into decimal format. Monthly water body frequency map was generated by dividing the number of 8-days with water observations to the number of 8-days with good observations in a month. Annual water body frequency map was generated by averaging monthly water body frequency in a year. To get the 34-year water body frequency map, first, 8-day water body frequency was calculated as the ratio of water observations to good observations of the same 8-day in the last 34 years. Then, only those 8-day water body frequencies with good

observations ≥ 5 in the last 34 years were averaged to generate the 34-year water body frequency map.

To exclude non-water features and potential commission errors caused by omitted cloud shadows and terrain shadows, pixels that meet three conditions (34-year water body frequency ≥ 0.05, 34-year water observations ≥ 2, and 34-year good observations ≥ 5) formed the maximum water body extent and only these pixels were included in the further analysis (Figure S4.5). In this way, some of the extremely ephemeral water bodies captured by water body maps, such as flash floods, were not included in variation and trend analysis across time. The interannual dynamics of surface water area were analyzed in 0.01° grid cells, 0.5° grid cells, and 5° tiles, respectively. Data gaps caused by bad weather and the incomplete Landsat data archive are the biggest challenges for the change analysis of surface water area across time. To avoid bias and make full use of the data, change analysis of surface water area was based on the common regions with valid observations across all selected years (Figure S4.5).

4.2.5 Interannual variations and trends.

To study interannual dynamics of surface water area in the past 34 years, we first selected annual valid observation pixels, those with enough good observations to represent water body conditions in a year (Figure S4.5). Water bodies at high-latitude and high-altitude regions freeze in cold seasons, reducing the number of months with good observations (Figure S4.1, S4.6). In these regions, annual water body conditions were represented by good observations in the non-frozen months. Good observations were also reduced across the globe because of cloud coverage, especially the regions with long rainy seasons, such as the tropics (Figure S4.1, S4.6). Good observations in mountainous regions, especially

in high latitudes, were reduced in the months and seasons with low sun elevation angles because of terrain shadows (Figure S4.1). Thus, one condition to select the annual valid observations is that they should have good observations in at least half number of the non-frozen months (Figure S4.5). An alternative condition is that they should have good observations in multiple seasons (Figure S4.5). For each year, annual valid observations will be classified as annual water body pixels if their annual water body frequency ≥ 0.5. The threshold of 0.5 was selected because it can best mitigate the impacts from both commission and omission errors in regions with limited good observations, such as the tropical, high-latitude, and high-altitude regions. It is worth to mention that the interannual variation patterns of surface water areas in large regions don't change across different frequency thresholds (Zou et al., 2018). The maximum annual water body extent was the union of annual water body pixels in each of the past 34 years. Surface water area dynamics at various scales were analyzed within the maximum annual water body extent.

For 0.01^{o} grid cells, the years with common annual valid observation pixels covering 100% of the maximum annual water body extent were selected (Figure S4.5). Annual surface water area in each selected year was derived from the common valid observation pixels. If the selected number of years ≥ 10 (Figure S4.7), trend analysis was carried out and this grid cell was marked as a location whether surface water area had significantly changed or not. For 5^{o} tiles, select the years with common annual valid observation pixels covering at least 25% of the maximum annual water body extent (Figure S4.7). Regional surface water area in each selected year was also derived within the common valid observation pixels. Interannual variation of regional surface water area was analyzed and the long-term trend analysis was carried out if the selected number of

years ≥ 10 (Figure S4.5). For regional surface water area in the past decades within a 5° tile, its interannual variability, defined as the ratio of its range to its mean, was displayed as the amplitude of y axis in its curve. The regional water area coverage, defined as the ratio of surface water area to the total land area within a 5° tile, was displayed as linewidth. Long-term trend analysis of surface water area at 0.01° grid cells and 5° tiles were carried out through simple linear regression models (annual surface water area as dependent variable while year as independent variable) with t-test at the 5% significance level using Python module Statsmodels.

Interannual trends of land water storage from 2002 to 2016 were analyzed at 0.5° grid cells using simple linear regression models (annual land water storage as dependent variable while year as independent variable). Surface water areas during 2002-2016 at 0.5° grid cells were derived using the same methods as those used in 5° tiles. Relationships between annual land water storage and surface water area were also explored at 0.5° grid cells using simple linear regression models (annual land water storage as dependent variable while surface water area as independent variable). Annual land water storage data at 5° tiles were averaged from those at 0.5° grid cells. Interannual variations and linear regression trends of both surface water area and land water storage were shown at 5° tiles across the globe to indicate the consistent and divergent changes of surface and land water resources during 2002–2016.

Ocean water, outside the boundaries of world countries (ESRI, 2018), was not included in the statistics or change analysis of surface water area and land water storage. Multiple stepwise linear regression models were used in the platform of MATLAB R2014a to examine the impact of annual precipitation and annual temperature on the

interannual variations of regional surface water areas during 1984–2017 at 5^o tiles (Figure S4.8). The statistics of maximum water body extent and surface water areas in 2017 by country (ESRI, 2018) were provided in Table S4.4.

4.3 Results and discussion

4.3.1 Surface water body frequency and area

Our global 34-year surface water body frequency map provides the location and extent of surface water bodies and illustrates the stability of surface water resources over the past three decades (Figure 4.1a). High frequency values represent consistent multi-decadal water bodies, while low frequency values represent discontinuous inundation, such as seasonal water bodies and newly constructed reservoirs (Figure 4.1b). According to this 34-year frequency map, there are 5.6, 3.5, 2.9, and 2.6 million km^2 global surface water areas with water body frequencies ≥ 0.05, ≥ 0.25, ≥ 0.5, and ≥ 0.75, respectively, which were larger than those derived from the JRC global water occurrence map (4.1, 3.3, 2.8, and 2.5 million km^2) (Pekel et al., 2016). We compared surface water areas from these two datasets over 1,042 overlapping 5° tiles (Figure 4.1c), and the results showed that our dataset (University of Oklahoma, OU) captured more surface water bodies, especially those with low water body frequencies. Global maximum surface water body extent from the OU dataset (5.6 million km^2) was ~30% bigger than that of the JRC dataset (4.3 million km^2) (Figure 4.1d), attributed in part to a slightly longer study period (1984–2017 vs 1984–2015), a slightly larger spatial extent (60°S–85°N vs 60°S–78°N), more Landsat data input (3.8 vs 3.1 million images) (Figure S4.1, Figure S4.2), and the difference in water detection and post-processing algorithms.

Annual total global surface water area from the OU and JRC datasets were sensitive to the total area with valid observations each year (Figure 4.1d), which was related to the number of available Landsat images (Figure S4.1a). Thus, to avoid the bias of data gaps, interannual change analysis of surface water area in this study was based on the common regions with valid observations across all selected years. Three global annual surface water area estimates, derived from the OU dataset using annual water body frequency ≥ 0.25, 0.5, 0.75, respectively, had very similar temporal variation patterns in the past decades (Figure 4.1d). The annual surface water area with annual water body frequency ≥ 0.5 was chosen for the interannual change analysis in this study because it can best mitigate the impacts of both omission and commission errors in regions with limited good observations, such as tropical, high-latitude, and high-altitude regions.

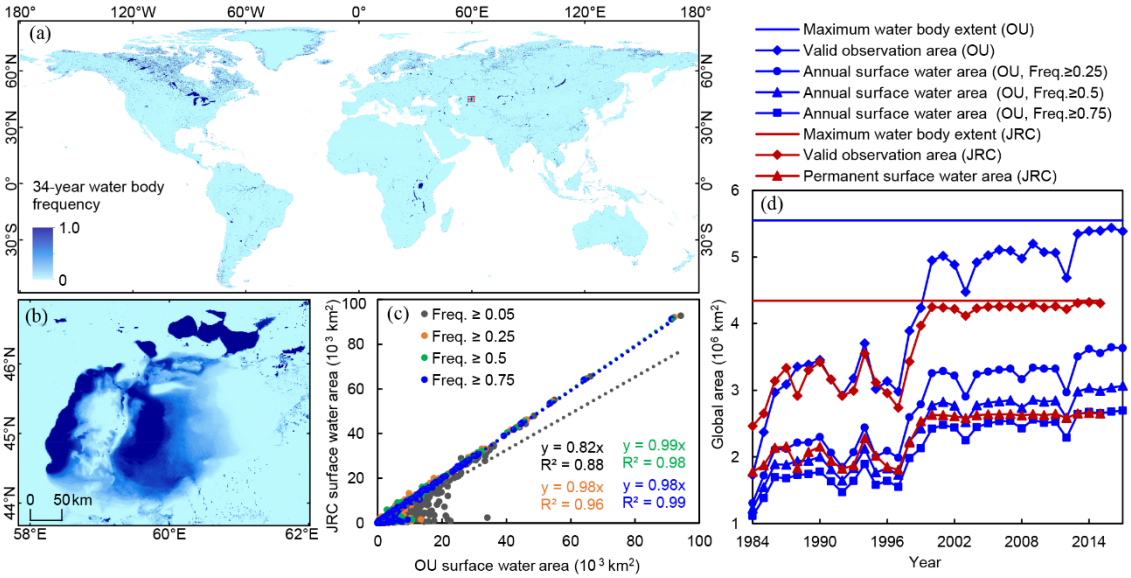


Figure 4.1 Global water body frequency and area. Global 34-year water body frequency map in 1984–2017 (a), and its zoom-in of Aral Sea (b). (c), Comparison of surface water areas from the OU 34-year water body frequency map and JRC water occurrence map in 1042 overlapping 5° tiles across the globe. (d), Global maximum surface water body extent, valid observation area, and surface water area from the OU and JRC datasets over the past decades.

4.3.2 Variations and trends of surface water area

Multi-decadal trends of surface water area within 0.01° grid cells provide the locations where surface water area had significantly increased or decreased during 1984–2017 (Figure 4.2a and its zoom-ins in Figures 4.3,4.4). Among the 36.7 million 0.01° grid cells across the globe that had surface water bodies and enough valid observations for the trend analysis, 4.0 million had significant increasing trends while 4.5 million had significant decreasing trends. The interannual variations, amplitude of variability, and long-term trends of regional surface water area were shown in 5° tiles (Figure 4.2b). Among the 880 five-degree tiles with enough valid observations for trend analysis, 189 and 170 had significant increasing and decreasing trends, respectively.

In Asia, significant increasing trends of surface water area occurred in the Tibetan Plateau, eastern China, and parts of India, Myanmar, Thailand, and Laos, while significant decreasing trends were found in the northern Caspian Sea and Aral Sea region, Lake Baikal region, and southern Japan (Figure 4.2a, Figures 4.3). Most lakes on the Tibetan Plateau, especially in its northern region, expanded substantially (Figure 4.3c), which increased the regional surface water area (Figure 4.2b). Glacial meltwater has contributed to the increase in surface water area (Song et al., 2013), and our regression models showed that precipitation also contributed significantly to the increase (Figure S4.8). In eastern China, grid cells with increasing trends of surface water area were clustered in river channels and new reservoirs (Figure 4.3d). This could be explained by the construction of more than 200 large reservoirs in China between 1984 and 2011, with a total capacity of ~250 km^3 (Ministry of Water Resources, 2013; Yang and Lu, 2014). The Three Gorges Dam alone created a surface water area of ~1,000 km^2 in this region

(Zhao et al., 2000). While man-made reservoirs increased, many natural lakes in eastern China shrank and disappeared because of land reclamation, lake isolation, and dam construction (Du et al., 2011; Mei et al., 2015). Thus, a large number of grid cells in this region had decreasing trends in surface water area. In mid-western India and parts of Myanmar, Thailand, and Laos (Figure 4.3e, f), many large reservoirs were created by high dams, such as the Indira Sagar Dam constructed in 2005, the Yeywa Dam constructed in 2010, and the Nam Ngum 2 Dam constructed in 2011. These new reservoirs had remarkably increased the regional surface water area (Figure 4.2b).

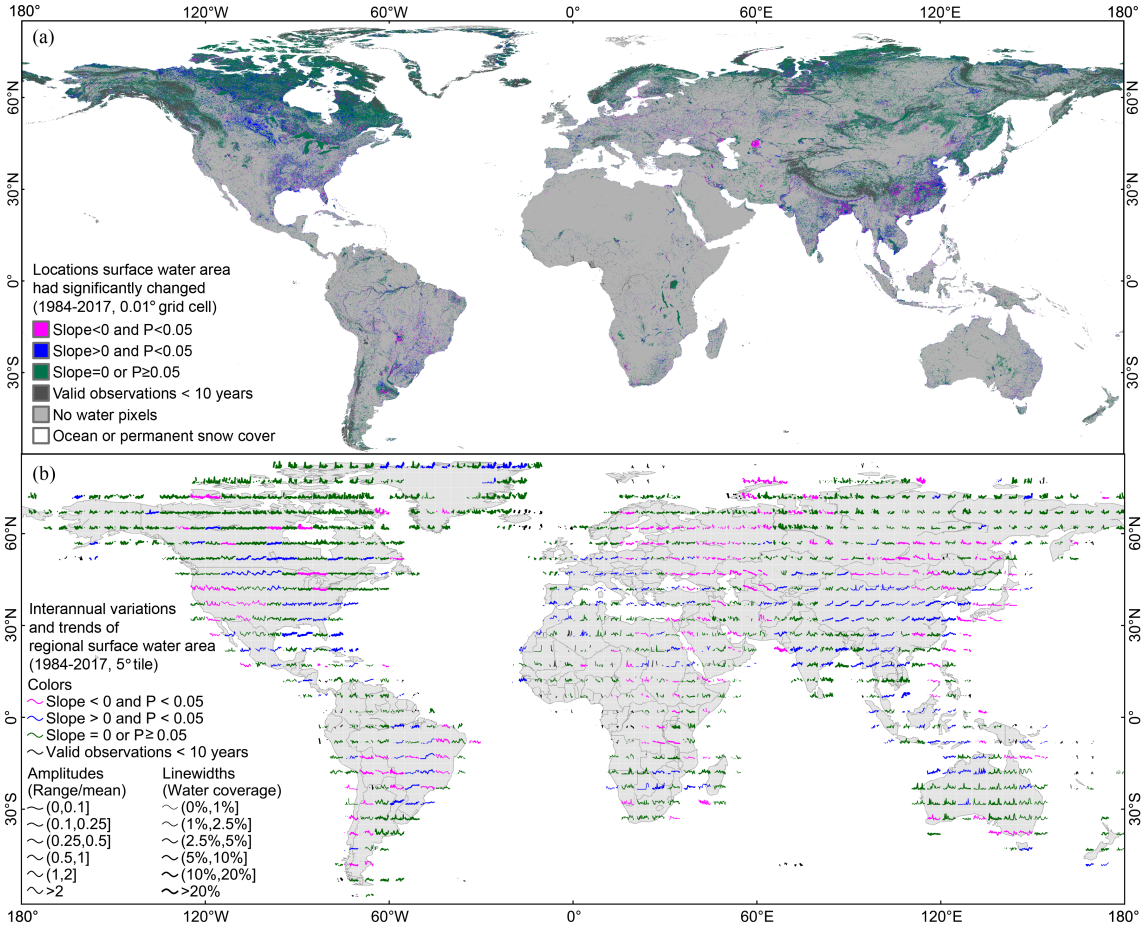


Figure 4.2 Changes of surface water area across the globe. (a), Locations where surface water area had significantly changed during 1984–2017 at 0.01° grid cells. (b), Interannual variations and multi-decadal trends of regional surface water area during 1984-2017 at 5° tiles.

Regional surface water area had a small decrease in northern Caspian Sea region and a large decrease in the Aral Sea region (Figure 4.2b). The Aral Sea shrank from 39,734 km^2 in 1989 to 17,382 km^2 in 2006 due to upstream water diversions for irrigation (Micklin, 2007) (Figure 4.3i). In Lake Baikal region, surface water area in all river channels and the Selenga River Delta had decreasing trends (Figure 4.3g). The flow of the Selenga River near its delta was greatly diminished during 1996–2015 because of reduced precipitation, especially during the summer (Frolova et al., 2017). Regional surface water area had a large decrease in southern Japan with decreasing grid cells clustering along the coastal areas (Figure 4.2b, Figure 4.3h). Over the past decades, urbanization in Japan has expanded into deltas, flood plains, and coastal plains (Yoshimura et al., 2005). Sea enclosing and land reclamation were the major reasons for surface water area decrease in Japanese coastal regions, such as Tokyo Bay, Osaka Bay, Ise Bay, and Hakata Bay (Figure 4.3h).

In Europe, surface water area had significant decreasing trends in Novaya Zemlya Archipelago and Gulf of Bothnia, but significant increasing trends in lower Europe (Figure 4.2b, Figure 4.3). Grid cells with decreasing trends were clustered in the outlet glaciers in Novaya Zemlya Archipelago (Figure 4.3b), while the regional surface water area had a weak decreasing trend with strong temporal variations (Figure 4.2b). Most of the outlet glaciers were found to be retreating in recent decades due to global warming (Carr et al., 2014). Grid cells with decreasing trends were clustered at the shorelines of the Gulf of Bothnia (Figure 4.3j), resulting in a slight decrease in regional surface water area (Figure 4.2b). The decrease of coastal surface water area was likely caused by the Fennoscandian land uplift with a speed of up to 10 mm/year in some areas (Muller et al.,

2012). Based on 13 Finnish tide gauges, the Finnish coast had an average rise of 5.3 mm/year after subtracting sea level rise (Johansson et al., 2004), which cumulated in a total rise of ~180 mm over the last 34 years. Regional surface water areas in southern Europe showed a small increase along with some variations and data gaps (Figure 4.2b). The increasing grid cells had a relatively dense distribution in central France and where some new reservoirs have been constructed in the southwestern region of Spain and Portugal (Figure 4.3k), such as Alqueva Reservoir constructed in 2002 and Alange Reservoir constructed in 1990.

In North America, surface water area significantly increased in northern Greenland, central North America, and the southeastern US, but significantly decreased in the western US and Great Lakes (Figure 4.2b, Figure 4.4). Grid cells with increasing trends were clustered in the outlet glaciers and the edges of ice sheet in northern Greenland while the regional surface water area increased dramatically in the last decade (Figure 4.2b, Figure 4.4b). Global warming has led to the melting of glaciers and an increase of surface meltwater (Fettweis, 2007). In the record melt years of 2010, 2012, and 2016 (van As et al., 2018), the surface area of supraglacial lakes expanded and new lakes formed at higher elevations (Fitzpatrick et al., 2014), which could have caused the peaks in regional surface water area (Figure 4.2b). A large number of grid cells with increasing trends were concentrated in central North America (Figure 4.4c). The increase in regional surface water area on Canada side was slow and steady while the increase on the US side was accompanied by multi-year, fluctuating cycles (Figure 4.2b). Regression models showed that precipitation was the major factor behind the water area increase in this region (Figure S4.8). Grid cells with increasing trends were scattered across the entire

eastern US and clustered in the Mississippi River Delta and the East Coast (Figure 4.4d). The increase of surface water area in the interior region was likely caused by precipitation according to regression models (Figure S4.8), while the increase in the Mississippi River Delta and the East Coast was caused by sea level rise (Ezer, 2013) and the reduction of sediment load due to dam construction (Blum and Roberts, 2009).

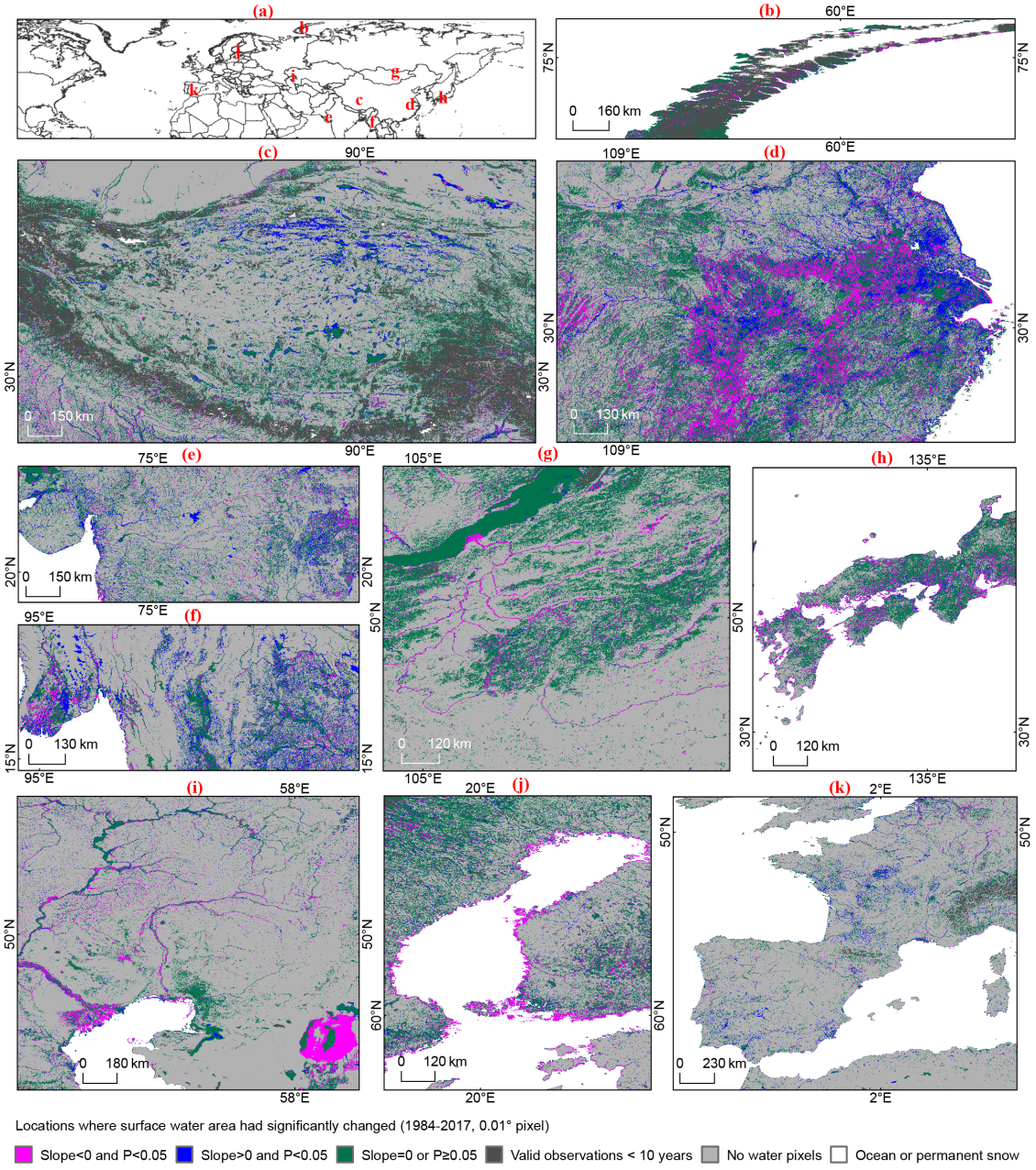


Figure 4.3 Zoom-ins of Figure 4.2a in Asia and Europe. (a), Locations of the zoom-ins. (b), Novaya Zemlya Archipelago. (c), Tibetan Plateau. (d), Eastern China. (e), Middle western India. (f), Myanmar, Thailand, and Laos. (g), Lake Baikal region. (h), southern Japan. (i), Northern Caspian Sea and Aral Sea region. (j), Gulf of Bothnia. (k), Southern Europe.

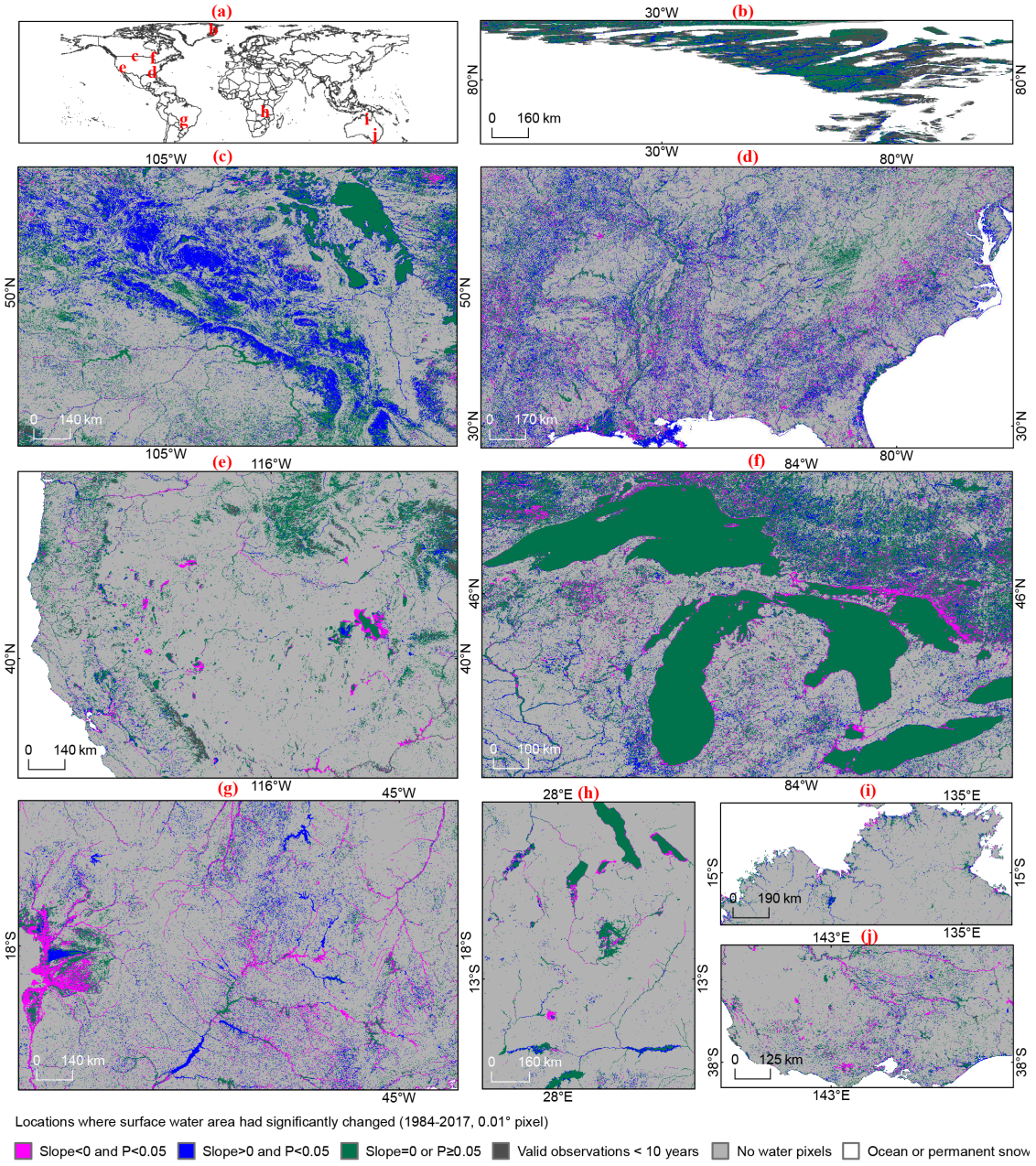


Figure 4.4 Zoom-ins of Figure 4.2a in North America, South America, Africa, and Australia. (a), Locations of the zoom-ins. (b), Northern Greenland. (c), Central North America. (d), Southeastern US. (e), Western US. (f), Great Lakes. (g), Central South America. (h), Lower Africa. (i), Northern Australia. (j), Southeastern Australia.

Most of the large water bodies in western US shrank remarkably in the past decades (Figure 4.4e), forming decreasing trends in nine 5° tiles (Figure 4.2b). The

Western US has been a hotspot of water stress (Melillo et al., 2014). Drought and increased warming has caused water scarcity in the southwestern US, and changes in the timing of streamflow reduced summer water supply in the northwest (Melillo et al., 2014). Grid cells with decreasing trends were clustered at the edges of Great Lakes, especially the northern shores of Lake Huron, leading to a minor decrease in regional surface water area (Figure 4.2b, Figure 4.4f). Two major water-level lowering episodes occurred in Lake Michigan-Huron, Superior, and Erie in late 1980s and in 1997–2000, with low inflow, low precipitation, and high lake evaporation as the major reasons for the latter episode (Assel et al., 2004). The water level of Lake Superior and Erie recovered around 2002 and 2005, respectively, while Lake Michigan-Huron did not recover by the end of 2012 (Gronewold et al., 2013).

In South America, we found variable increasing and decreasing trends of regional surface water area in the central region of this continent (Figure 4.2b). Grid cells with increasing trends were scattered across the region and were clustered in the newly constructed reservoirs, while grid cells with decreasing trends were concentrated in the Pantanal wetlands and most river channels (Figure 4.4g). While new reservoirs had created large surface water area in this region, they had also decreased surface water area in downstream rivers and reservoirs. The shrinkage of three large reservoirs in the lower reaches of the Paranaíba River (São Simão Reservoir, Itumbiara Reservoir, and Emborcação Reservoir, constructed before 1984) were likely related to the construction of nine large upstream reservoirs after 1993 (Figure 4.4g). Reservoir construction and canalization are among the top threats to Pantanal wetland (Junk and de Cunha, 2005).

In Africa, many grid cells with significant trends were concentrated at the edges of large lakes in the lower region of the continent, forming steady decreasing trends in the north and variable increasing trends in the south (Figure 4.2b, Figure 4.4h). Lake Rukwa had the largest decrease in surface water area in this region, which was mainly caused by water withdrawal from its tributary, the Katuma River, for rice irrigation (Elisa et al., 2010). The area of numerous water bodies in Lukanga Swamp Ramsar site had shrunk, which could be related to low rainfall, deforestation, and the expansion of agricultural activities (Chabwela et al., 2017).

In Australia, grid cells in river channels and shorelines of northern Australia had increasing trends while grid cells in lakes, reservoirs, and river channels of southeastern Australia had decreasing trends (Figure 4.4i, j). In southeastern Australia, we found smaller regional surface water area during the Millennium Drought (2001–2009) (Figure 4.2b), a period with many years of below median rainfall (van Dijk et al., 2013). The peak of the regional surface water area in 2011 was related to a strong La Niña event during 2010–2011 (Beard et al., 2011).

4.3.3 Land water storage and surface water area

Trend analysis of land water storage during 2002–2016 at 0.5° grid cells showed that, among the 71,000 grid cells across the globe, 18,000 had significant decreasing trends while 21,000 had significant increasing trends (Figure 4.5a). Surface water area had a significant positive relationship with land water storage in ~28% of the 52,000 half-degree grid cells that have enough data for regression analysis (Figure 4.5b). Change of surface water area in these regions is a good indicator of land water storage dynamic.

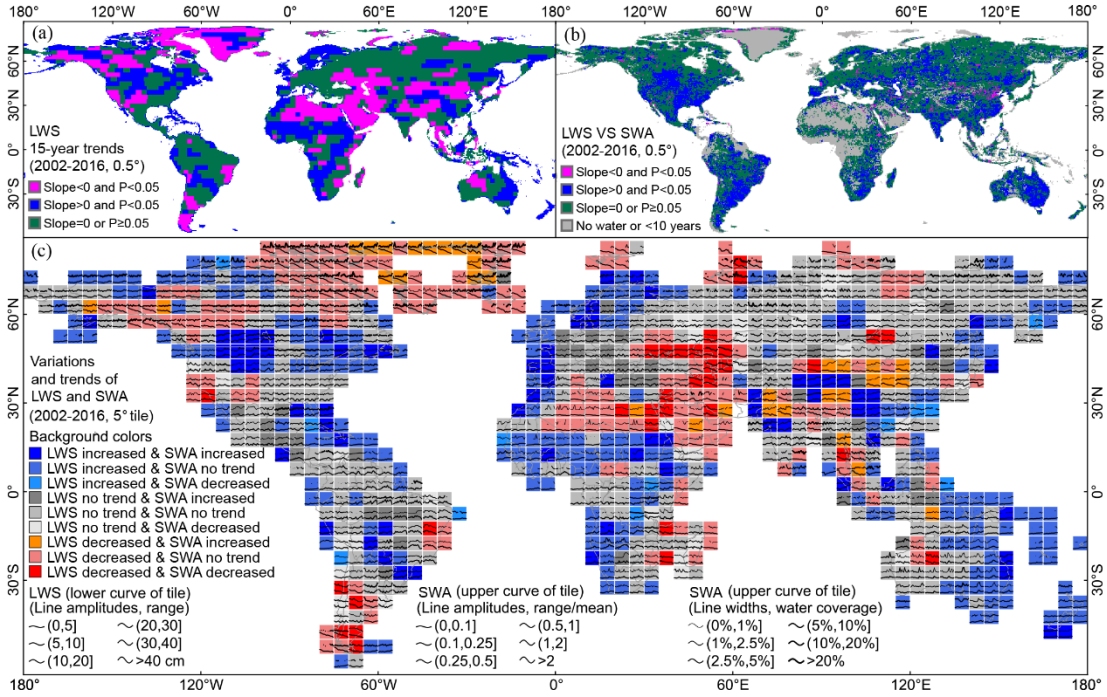


Figure 4.5 Changes of land water storage and surface water area. (a), 15-year trends of land water storage (LWS) at 0.5° grid cells during 2002–2016. (b), Regression analysis with land water storage as dependent variable and surface water area (SWA) as independent variable at 0.5° grid cells during 2002–2016 (regression was not analyzed when data < 10 years). (c), Interannual variations and 15-year trends of land water storage and surface water area at 5° tiles during 2002–2016 (land water storage or surface water area was considered as no trend when data < 10 years).

The interannual variations and trends of both regional surface water area and land water storage at 5° tile from 2002 to 2016 revealed the consistent and divergent trends of surface and land water resources (Figure 4.5c). Consistent increasing trends of both regional surface water area and land water storage indicated a general increase of surface and land water resources in central North America, Scandinavia, central and northern Tibetan Plateau, central China, central South America, and eastern Australia (Figure 4.5c). In these regions, surface water area and land water storage had similar interannual variation patterns, suggesting that they might be driven by some common factors. For example, a progression from a dry to a wet period in central North America and eastern

Australia (Rodell et al., 2018), water level increase in surface water bodies and groundwater in Scandinavia (Wang et al., 2013), temperature and precipitation increase in Tibetan Plateau (Kuang and Jiao, 2016), reservoir construction in central China (Ministry of Water Resources, 2013), and the recovery from early period drought in central South America (Rodell et al., 2018).

Consistent decreasing trends of both regional surface water area and land water storage indicated a general decrease of surface and land water resources in lower South America, the upper Red Sea surrounding region, and the Caspian Sea surrounding region (Figure 4.5c). Similar interannual variation patterns also existed in these consistent trends and could be driven by some common factors, such as ice-field melt and a progression from a wet to a dry period in lower South America (Rodell et al., 2018), a decline in groundwater and regional rainfall in upper Red Sea surrounding region (Fallatah et al., 2017), and groundwater depletion, precipitation decline, and Caspian Sea decline in Caspian Sea surrounding region (Chen et al., 2017; Deng and Chen, 2017).

Divergent trends with decreased land water storage but increased regional surface water area occurred in the edges of Greenland, a vast region of northern China, and the Indus Basin (Figure 4.5c). The Greenland ice sheet had a big decreasing rate in mass (-279.0±23.2 Gt yr^{-1}) because of melting during 2002–2016 (Rodell et al., 2018), and its meltwater expanded the supraglacial lakes (Fitzpatrick et al., 2014). In the vast arid and semiarid region of northern China, ground water pumping for food production was one of the major reasons of land water storage decrease in North China Plain (Rodell et al., 2018), the Loess Plateau (Zhao et al., 2013), and the lower Mongolia Plateau (Zhang et al., 2018). Groundwater pumping for afforestation projects in the lower Mongolia Plateau

might have also contributed to the local decrease in land water storage (Zhang et al., 2018). The increased surface water area in northern China could be related to dam construction, water regulation, and water diversion projects (Ministry of Water Resources, 2013). The annual average runoff at lower reaches of Yellow River increased during 2003–2011 because of water regulation and sediment flushing operations (Kong et al., 2015). In the lower Mongolia Plateau, a government ecological water diversion project was carried out to recover rivers and lakes in the lower reaches of Heihe River since 2000 (Cheng et al., 2014). In the Indus Basin region, land water storage decrease was caused by groundwater depletion (Rodell et al., 2018), while surface water area increase was caused by an increase in upstream rainfall (Laghari et al., 2012).

Divergent trends with increasing land water storage but decreasing regional surface water area occurred in a few 5° tiles of eastern China and central Africa (Figure 4.5c). The increase of land water resource in these regions could have been driven by changes in climate (Rodell et al., 2018) and anthropogenic activities, such as reservoir construction (Ministry of Water Resources, 2013). The decrease of surface water area in the coastal region of Zhejiang and Guangdong Provinces in eastern China was mainly caused by large-scale sea enclosing and land reclamation for the expansion of cities, ports, and industries (Wang et al., 2014). According to the reclamation plans approved by China's State Council, Zhejiang and Guangdong Provinces will reclaim 506 and 230 km^2 of coastal area during 2011-2020, respectively (Wang et al., 2014). In central Africa, the decrease of regional surface water area in the southeastern Democratic Republic of the Congo was mainly caused by the shrinkage of Lake Mweru, Mweru-Wantipa, Upemba,

and Tanganyika (Figure 4.4h), while the decrease in Uganda was caused by the water-level decline of Lake Kyoga, Victoria, and Albert (Moore and Williams, 2014).

4.4 Conclusions and perspective

Data gaps due to the incomplete Landsat data archive and bad weather remain the biggest challenge in the multi-decadal remote sensing analysis of surface water bodies (Pekel et al., 2016). Here we generated a new multi-decadal global surface water body dataset with more Landsat data input and relatively fewer data gaps compared with the JRC dataset (Figure S4.2). Compared with the JRC study, our algorithms are simple and easy to be sued in large scale studies. Moreover, our dataset has a longer study period (1984–2017 vs 1984–2015) and a bigger spatial extent (60°S–85°N vs 60°S–78°N), which is more suitable for time series trend analysis. To avoid the bias of data gaps, change analysis in this study was based on the common regions with valid observations across all selected years. Through this strategy, locations where surface water area had significantly changed in the past decades were provided in 0.01° grid cells, while the interannual variations and multi-decadal trends of regional surface water area were shown in 5° tiles globally. This study carried out a joint analysis of surface water area and land water storage at 5° tiles across the globe for the first time. The consistent and divergent trends between surface and land water resources, driven by climate and anthropogenic activities, were revealed. Global population is projected to increase from 7.6 billion in 2017 to ~10 billion in 2050 (DESA, 2017), which will further aggravate water scarcity, stress the environment, and threaten global biodiversity. The unprecedented spatial and temporal details of global water resource dynamics, uncovered by this study, are useful in water resource research,

planning, and management in coping with water scarcity and food security associated with population growth.

Supplementary materials

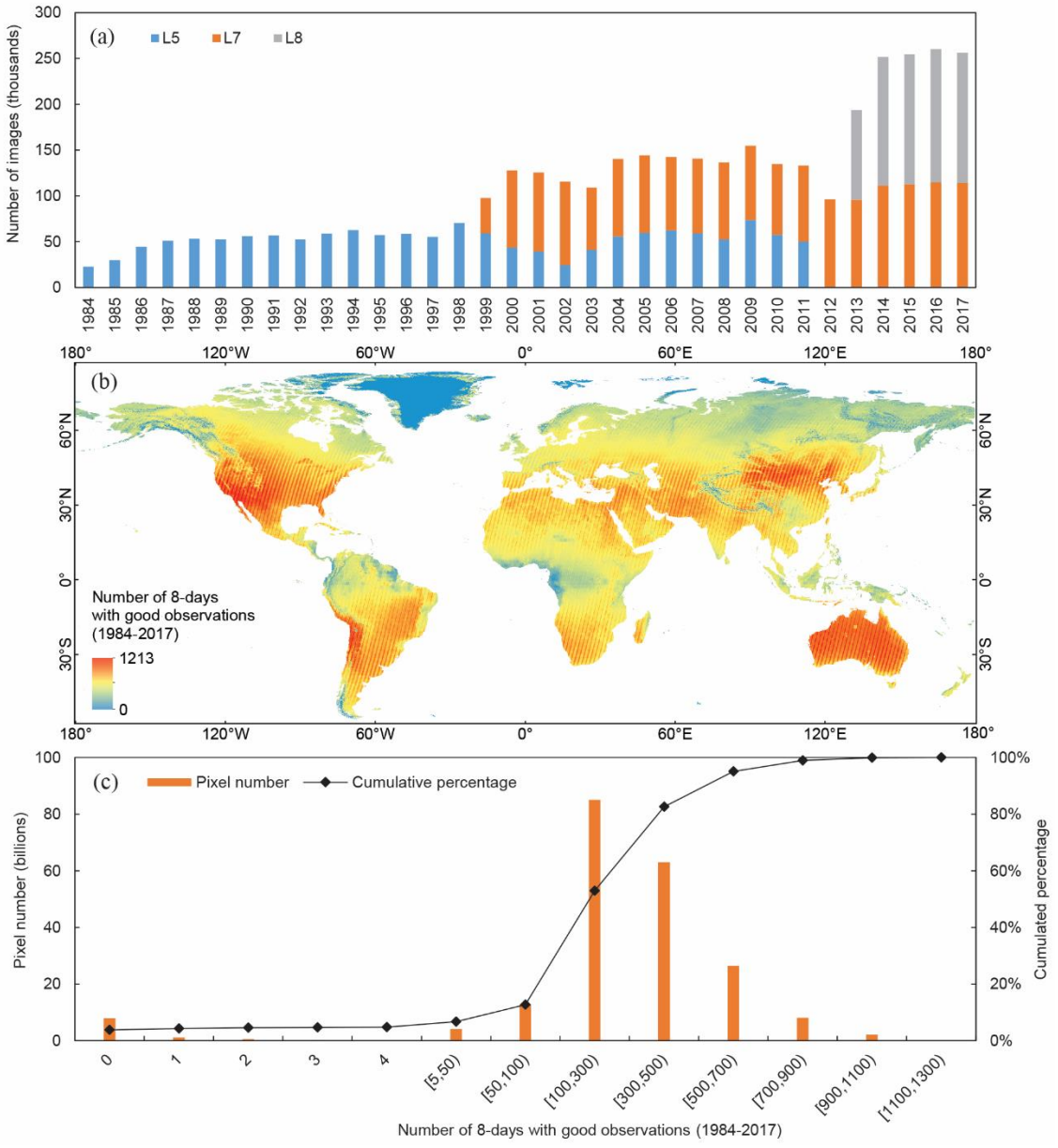


Figure S4.1 Landsat data used for water mapping. (a), Number of Landsat images from different satellites (Landsat 5, 7, and 8) in each year. (b), Number of 8-days with good observations at 30-m pixel scale. (c), Number of 30-m pixels with different good observations.

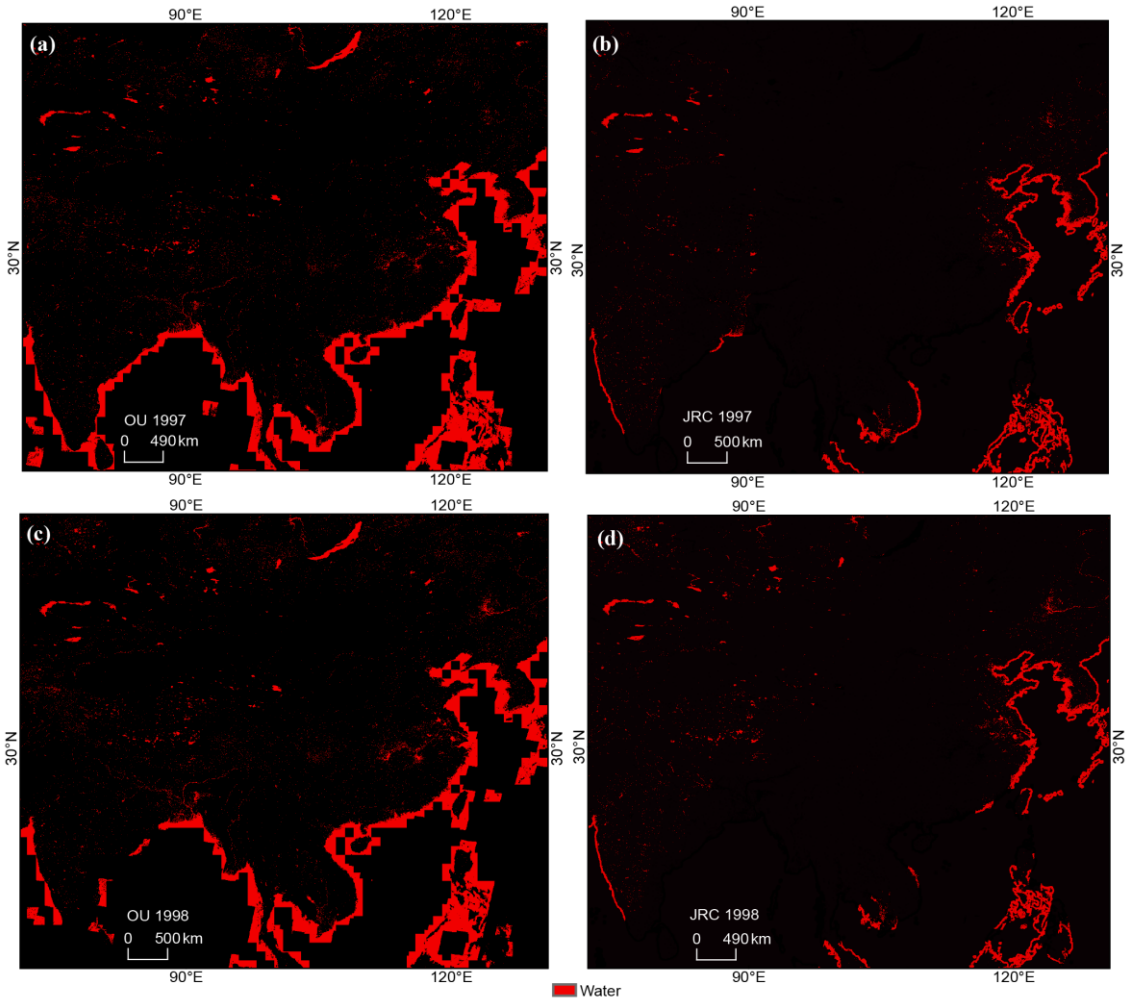


Figure S4.2 Water body maps of eastern Asia from OU and JRC datasets in 1997 and 1998. (a), OU water body map of 1997. (b), JRC water body map of 1997. (c), OU water body map of 1998. (d), JRC water body map of 1998. JRC water body maps in 1997 and 1998 did not captured water bodies in large regions of eastern Asia because no Landsat data in these regions when they carried out the research.

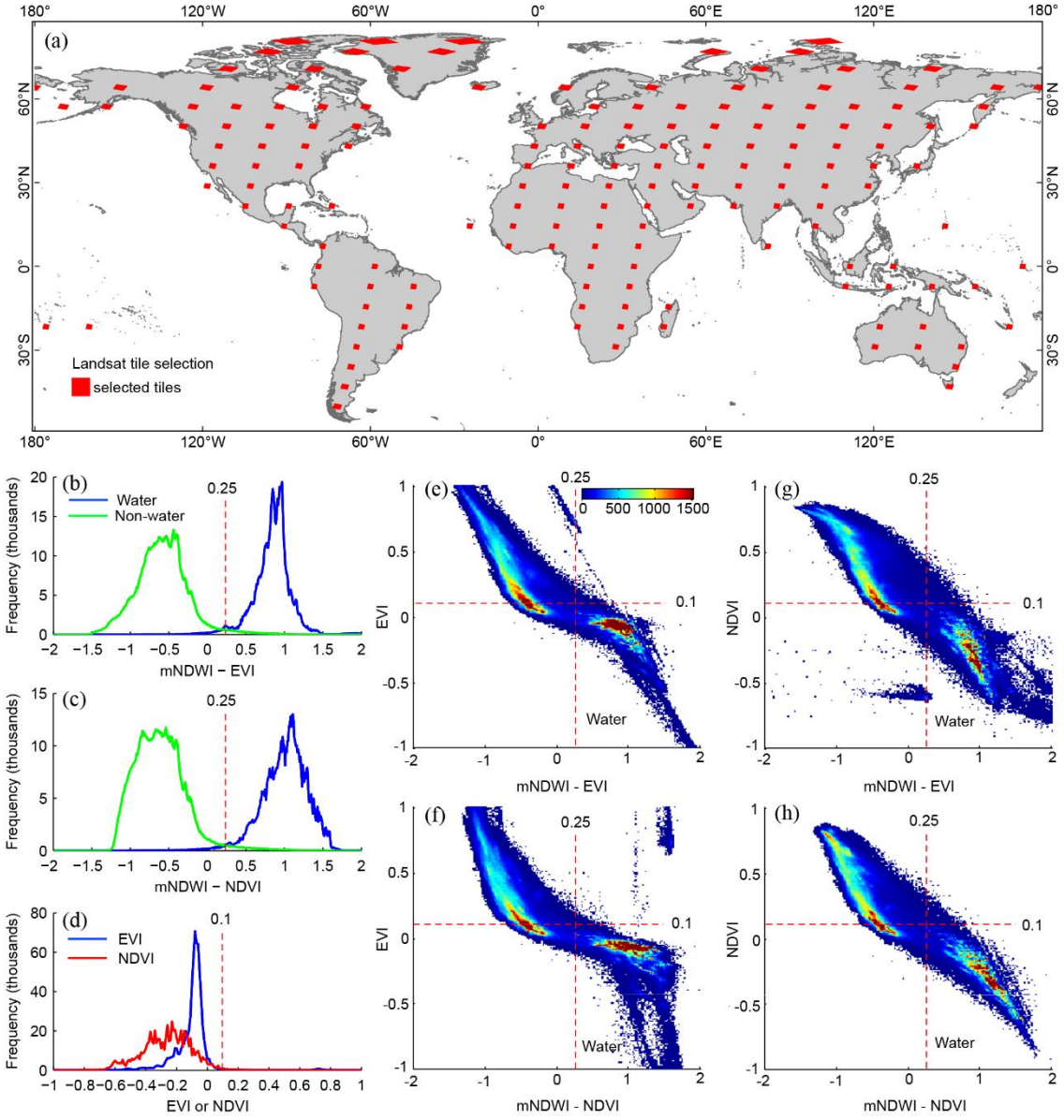
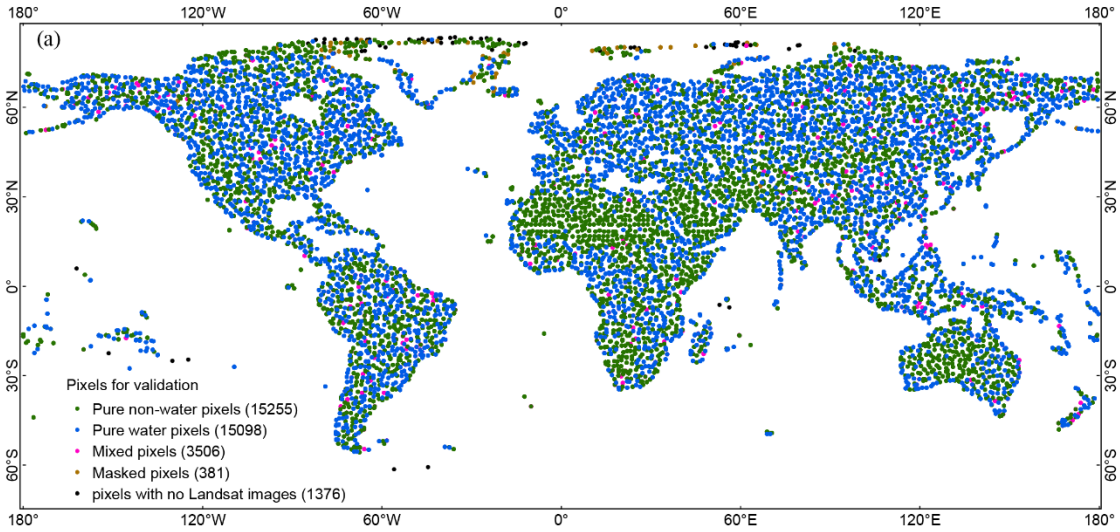


Figure S4.3 Algorithm development. (a), Landsat tiles selected. (b), Frequency distribution of (mNDWI-EVI) for water and non-water sampling pixels. (c), Frequency distribution of (mNDWI-NDVI) for water and non-water sampling pixels. (d), Frequency distribution of EVI and NDVI for water sampling pixels. Scatter density plots for all sampling pixels of EVI vs (mNDWI-EVI) (e), EVI vs (mNDWI-NDVI) (f), NDVI vs (mNDWI-EVI) (g), and NDVI vs (mNDWI-NDVI) (h).



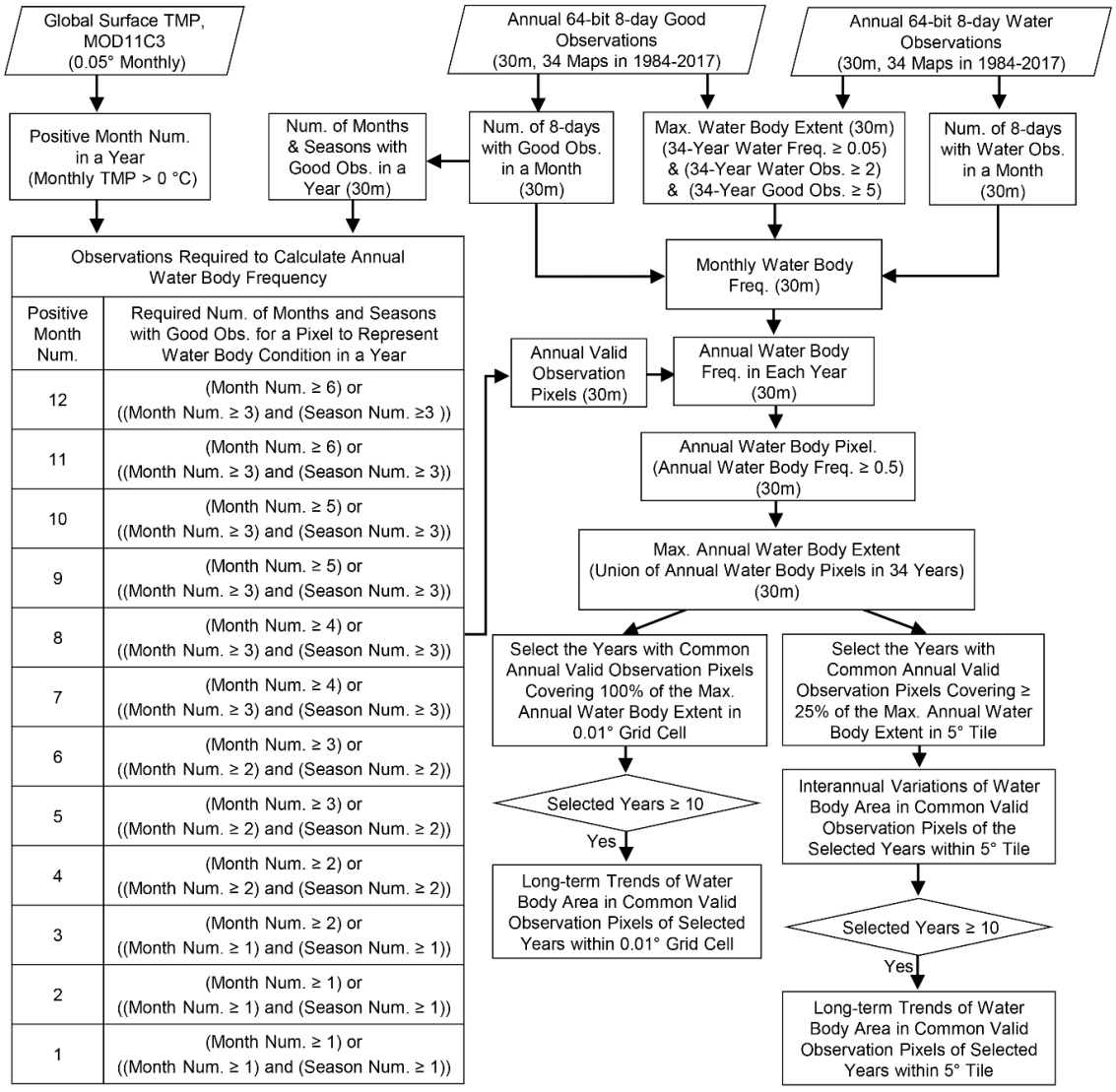
(b)

	Pure Reference pixels		Sum	User accuracy
	Water	Non-water		
Water	14891	138	15029	99.10%
Non-water	207	15117	15324	98.65%
Sum	15098	15255	30353	
Producer accuracy	98.63%	99.10%		OA=98.86%

(c)

	All Reference pixels		Sum	User accuracy
	Water	Non-water		
Water	15685	426	16111	97.36%
Non-water	1101	16647	17748	93.80%
Sum	16786	17073	33859	
Producer accuracy	93.44%	97.51%		OA=95.49%

Figure S4.4 Validation. (a), Validation pixel distribution. (b), Validation against pure water and non-water reference pixels only. (c), Validation against all reference pixels (mixed pixels with water percentage > 50% were classified as water pixels, while ≤ 50% were classified as non-water pixels).



Observations Required to Calculate Annual Water Body Frequency	
Positive Month Num.	Required Num. of Months and Seasons with Good Obs. for a Pixel to Represent Water Body Condition in a Year
12	(Month Num. ≥ 6) or ((Month Num. ≥ 3) and (Season Num. ≥3))
11	(Month Num. ≥ 6) or ((Month Num. ≥ 3) and (Season Num. ≥ 3))
10	(Month Num. ≥ 5) or ((Month Num. ≥ 3) and (Season Num. ≥ 3))
9	(Month Num. ≥ 5) or ((Month Num. ≥ 3) and (Season Num. ≥ 3))
8	(Month Num. ≥ 4) or ((Month Num. ≥ 3) and (Season Num. ≥ 3))
7	(Month Num. ≥ 4) or ((Month Num. ≥ 3) and (Season Num. ≥ 3))
6	(Month Num. ≥ 3) or ((Month Num. ≥ 2) and (Season Num. ≥ 2))
5	(Month Num. ≥ 3) or ((Month Num. ≥ 2) and (Season Num. ≥ 2))
4	(Month Num. ≥ 2) or ((Month Num. ≥ 2) and (Season Num. ≥ 2))
3	(Month Num. ≥ 2) or ((Month Num. ≥ 1) and (Season Num. ≥ 1))
2	(Month Num. ≥ 1) or ((Month Num. ≥ 1) and (Season Num. ≥ 1))
1	(Month Num. ≥ 1) or ((Month Num. ≥ 1) and (Season Num. ≥ 1))

Figure S4.5 Diagram of interannual variation and multi-decadal trend analysis of surface water area. TMP, Num., Max., Obs., and Freq. are short for temperature, number, maximum, observation, and frequency, respectively.

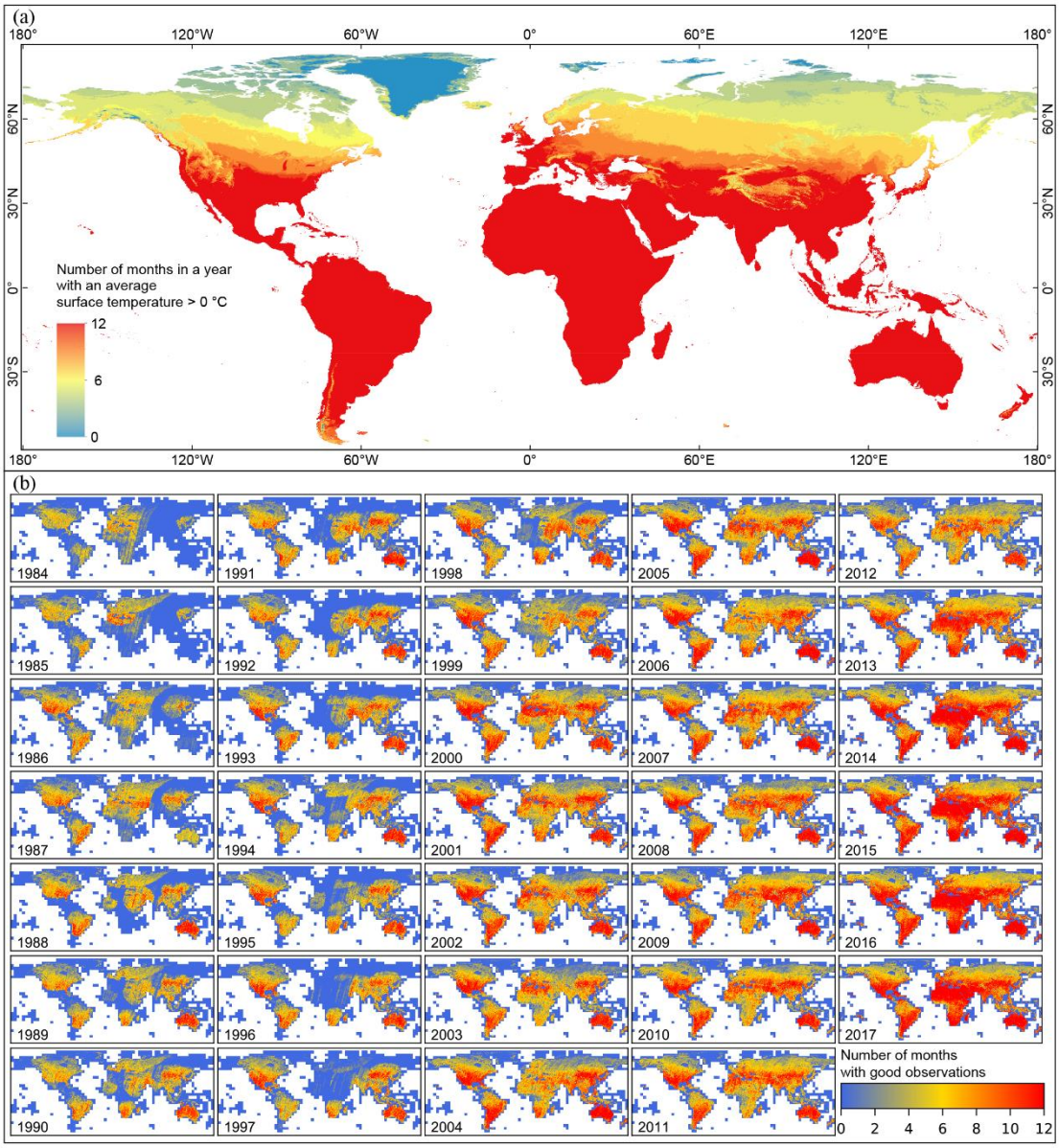


Figure S4.6 Non-frozen months and months with good observations. (a), Number of months in a year with an average surface temperature > 0 °C . (b), Number of months in a year with good observations. Monthly average surface temperature was defined as the mean of the same month across 2000–2017.

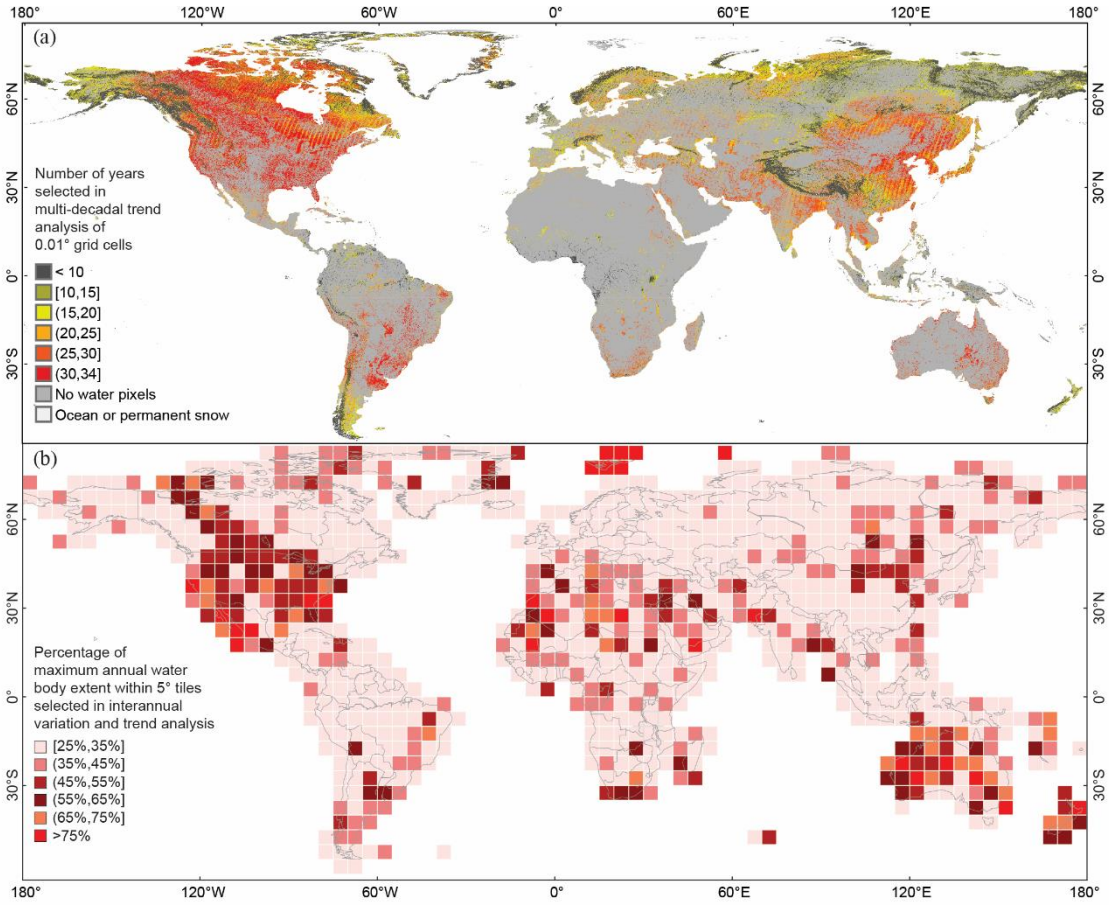


Figure S4.7 Data used in the variation and trend analysis. (a), Number of years selected in multi-decadal trend analysis of 0.01° grid cells. (b), Percentage of maximum annual water body extent within 5° tile selected in interannual variation and trend analysis (The number of years selected in variation and trend analysis was the same as the number of years included in regression models, which was shown in Figure S4.8a).

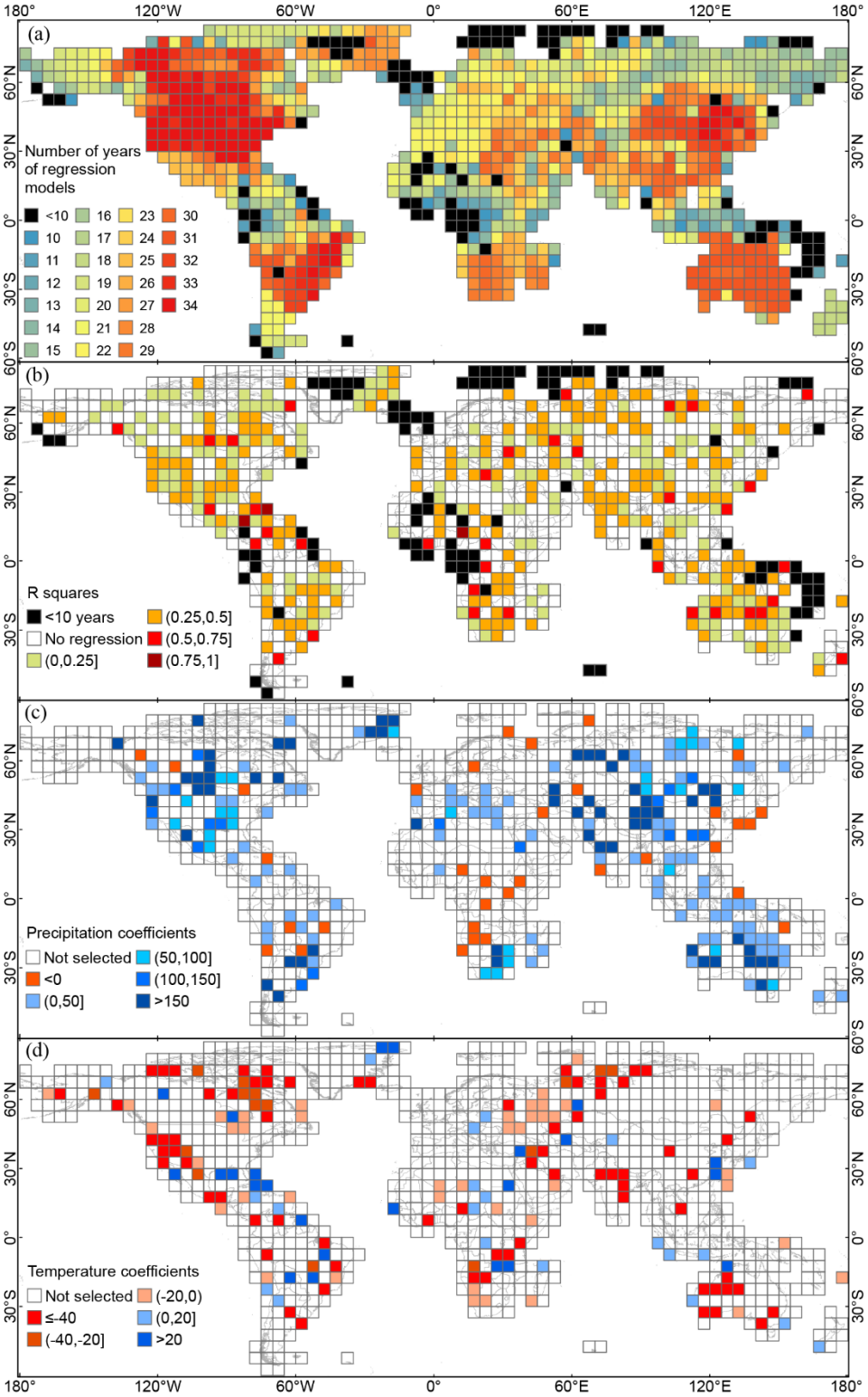


Figure S4.8 Multiple stepwise regression models. (a), Number of years with regional surface water area data to build regression models. (b), R squares of regression models. (c), Coefficients of factor annual average precipitation. (d), Coefficients of factor annual average temperature.

Table S4.1 Distribution of 157 sampling blocks

Terrain		Major land cover		DEM range		Satellite		Day range with very high-resolution image	
Type	Num.	Type	Num.	Range	Num.	Type	Num.	Range	Num.
Rolling	60	Vegetation	121	<100	61	L5	50	0	7
Flat	97	Crop	10	[100,500)	56	L7	33	[1-5)	40
		Urban	10	[500,1000)	18	L8	74	[5-10)	40
		Bare land	16	[1000,2000)	17			[10,15)	43
				[2000,3000)	3			[15,27]	20
				[3000,4000)	0				
				[4000,5000)	2				
Sum	157		157		157		157		150

DEM is short for digital elevation model. Seven sampling blocks close to North Pole have no very-high resolution image as reference, visual delineation in these sampling blocks was based on Landsat image itself.

Table S4.2 Confusion Matrix of algorithm development

	Ground Reference pixels		Sum	User accuracy (%)
	Water	Non-water		
Water	697,812	19,158	716,970	97.33%
Non-water	12,100	912,391	924,491	98.69%
Sum	709,912	931,549	1,641,461	OA=98.10%
Producer accuracy (%)	98.30%	97.94%		Kappa = 0.9613

Table S4.3 Distribution of Landsat validation pixels

Satellite		Year	
Satellite	Num.	Year range	Num.
Landsat 5	15444	<1990	3708
Landsat 7	13436	[1990-1995)	3908
Landsat 8	5360	[1995-2000)	3820
		[2000-2005)	4824
		[2005-2010)	4912
		[2010-2015)	8648
		≥2015	4420
Sum	34240	Sum	34240

Table S4.4 Maximum water body extent and water body areas by country

Country	Maximum water body extent/area (km^2)	2017 valid observation area percentage	2017 annual water body area (Freq. ≥0.25, km^2)	2017 annual water body area (Freq. ≥0.5, km^2)	2017 annual water body area (Freq. ≥0.75, km^2)
Canada	1489811	0.98	1172263	1049350	966443
Russian Federation	939405	0.96	574406	455270	387537
United States	508551	0.98	376405	336108	310690
China	380532	0.95	185602	147731	124694
Brazil	178116	0.96	120449	108150	90020
Greenland	174830	0.98	124038	85954	60091
India	132180	0.97	63100	41652	26946
Kazakhstan	129393	1.00	76107	63880	58957
Argentina	127725	0.98	65996	50114	40061
Australia	108340	1.00	43034	30341	22742
Tanzania	64150	1.00	58812	57729	56526
Chile	49420	0.86	20510	14706	12855
Congo DRC	45514	1.00	41519	40450	38596
Sweden	44524	0.97	38562	35831	33368
Indonesia	44334	0.92	22938	18145	13776
Mongolia	43990	0.98	18868	14432	13833
Pakistan	43054	0.92	14225	8220	4647
Bolivia	41785	0.84	18902	16924	14991
Iran	41042	1.00	13393	8443	5954
Uganda	37661	1.00	37089	36893	36560
Vietnam	36613	0.99	16143	8568	5466
Uzbekistan	36532	1.00	16088	13570	11962
Finland	36445	0.97	31775	30150	28654

Bangladesh	35343	1.00	15996	10539	5076
Iceland	33407	0.90	23046	15235	8505
Norway	31940	0.90	21726	17627	15106
Mexico	29288	1.00	17956	14609	11630
Peru	28290	0.94	18230	16421	13820
Malawi	24947	1.00	24249	24167	23805
Cambodia	24498	1.00	10248	5537	4102
Thailand	23824	1.00	12164	7430	5679
Zambia	23754	1.00	14206	12884	12375
Myanmar	22955	0.99	11424	8163	6270
Venezuela	22600	0.99	15500	12671	10361
Colombia	21645	0.97	13859	11641	9117
Turkey	19801	0.97	13346	12117	11414
Afghanistan	17019	0.94	4071	1552	854
Iraq	17000	1.00	6814	5220	4385
Tajikistan	16370	0.81	6587	3422	2271
Ukraine	15917	0.99	11790	10986	10276
Kyrgyzstan	15764	0.87	10554	8319	7496
Mozambique	15705	1.00	12465	11687	11116
Kenya	13340	1.00	12278	12015	11806
Ethiopia	12282	1.00	8865	8000	7336
Japan	12023	0.98	4567	3712	3191
Egypt	12022	1.00	7921	7236	6408
Nigeria	11559	0.92	5533	4355	3476
Philippines	11021	0.97	5616	4775	3973
Turkmenistan	10852	1.00	7046	6314	5552
North Korea	10311	1.00	2689	1551	1249
Mali	10293	1.00	3868	2550	1341
Nicaragua	10292	0.99	9598	9415	9239
New Zealand	9434	0.90	5634	4823	4344
Madagascar	9261	1.00	4057	2917	2223
Paraguay	8712	1.00	5389	4705	4424
South Africa	8675	1.00	4698	3949	3343
Ghana	8285	1.00	6189	5857	5407
Namibia	8093	1.00	1984	816	485
Angola	7721	0.97	1953	1400	1069
Papua New Guinea	7471	0.87	4391	3732	2839
Sudan	7354	1.00	5256	4353	3307
Chad	6734	1.00	3112	2216	1691

Ecuador	6618	0.89	3988	3276	2346
France	6503	0.95	4063	3271	2614
South Korea	6367	1.00	1993	1511	1220
Botswana	6284	1.00	4120	2909	1496
Italy	6090	0.92	3136	2797	2532
Uruguay	6033	1.00	4535	4180	3970
Germany	5936	0.97	4383	3921	3493
Spain	5875	0.99	3235	2423	1972
Romania	5619	0.96	3207	2927	2706
Malaysia	5609	0.99	3523	3014	2599
United Kingdom	5432	0.96	3415	2828	2324
Poland	5312	0.97	4168	3664	3233
Algeria	5036	1.00	1699	1084	651
Zimbabwe	5016	1.00	4398	4152	3912
Cuba	4856	1.00	3428	2827	2345
Nepal	4615	0.84	1229	552	296
Belarus	4589	0.99	2608	2053	1812
Cameroon	4485	0.96	3080	2561	2050
Bahamas	4381	1.00	3404	2835	2345
Congo	3766	0.97	2851	2615	2276
Laos	3762	1.00	3141	2838	2541
Senegal	3741	1.00	1664	962	473
South Sudan	3356	1.00	1366	946	497
Azerbaijan	3180	0.96	1684	1218	942
Saudi Arabia	2726	1.00	1028	584	367
Suriname	2715	0.92	2099	1944	1756
Tunisia	2684	1.00	1535	980	715
Greece	2542	0.98	1830	1683	1547
Estonia	2492	1.00	2262	2186	2132
Sri Lanka	2439	1.00	1305	1022	741
Mauritania	2377	1.00	862	459	230
Switzerland	2279	0.82	1598	1515	1460
Ireland	2231	0.95	1483	1300	1123
Guyana	2220	0.83	1067	892	744
Morocco	2212	1.00	947	726	578
Niger	2204	1.00	741	467	213
Hungary	2121	0.99	1491	1371	1256
C??te d'Ivoire	2110	0.82	1328	1152	886
Burundi	2090	1.00	2058	2053	2045

Austria	1973	0.78	857	721	655
Syria	1899	1.00	1286	1124	1025
Armenia	1692	0.98	1409	1358	1330
Rwanda	1631	1.00	1577	1561	1532
Georgia	1592	0.72	545	372	267
Bhutan	1587	0.69	262	111	64
Honduras	1568	1.00	1037	917	768
Bulgaria	1508	0.97	1001	929	855
Latvia	1482	0.99	1209	1053	905
Burkina Faso	1466	1.00	918	720	450
Lithuania	1425	0.96	1177	1033	897
Netherlands	1420	0.98	1125	973	832
Portugal	1389	1.00	873	699	579
Central African Republic	1361	1.00	1104	945	703
Guinea	1260	1.00	606	350	154
Denmark	1195	0.99	1062	954	869
Serbia	1189	0.97	838	766	695
Oman	1157	1.00	353	192	138
Guinea-Bissau	1122	1.00	507	276	143
Libya	1116	1.00	598	354	253
Panama	1045	0.98	783	664	504
Guatemala	1040	1.00	680	582	494
Croatia	1009	0.99	670	557	488
Falkland Islands	984	0.99	578	486	374
Albania	934	0.98	704	655	608
Somalia	926	1.00	302	224	159
Czech Republic	900	0.96	653	582	511
Eritrea	852	1.00	391	297	207
Gabon	849	0.85	639	606	536
Dominican Republic	809	0.99	622	590	555
Belize	808	1.00	565	496	451
Sierra Leone	752	0.94	411	280	160
United Arab Emirates	750	1.00	321	207	138
Jordan	711	1.00	557	499	485
French Guiana	706	0.42	249	229	201
Moldova	671	1.00	425	379	330
The Former Yugoslav Republic of Macedonia	608	0.99	540	530	519
Yemen	608	1.00	333	269	214
Israel	606	1.00	516	502	485

Bosnia and Herzegovina	601	0.97	322	276	231
Slovakia	598	0.88	346	302	274
French Southern Territories	502	0.97	192	153	133
Costa Rica	501	0.97	319	269	210
Haiti	488	1.00	336	306	275
Turks and Caicos Islands	476	0.92	330	267	202
Montenegro	471	0.98	352	296	276
El Salvador	467	1.00	386	340	267
Djibouti	461	1.00	225	190	153
Gambia	426	1.00	204	135	89
Canarias	399	0.99	206	151	116
Benin	389	0.93	196	168	134
Belgium	359	0.99	197	148	114
Solomon Islands	320	0.95	259	243	207
Togo	320	0.96	206	160	110
New Caledonia	301	1.00	183	152	125
Svalbard	300	1.00	146	93	75
Kuwait	278	1.00	129	76	41
Palestinian Territory	247	1.00	221	220	219
Liberia	219	0.62	82	50	20
Lesotho	208	0.99	89	59	44
Fiji	199	0.99	158	141	110
Slovenia	198	0.89	88	58	45
Qatar	175	1.00	76	56	45
Timor-Leste	157	1.00	62	39	29
Puerto Rico	145	0.99	110	97	83
Singapore	120	0.95	27	24	20
Swaziland	108	1.00	58	51	43
Kiribati	104	0.96	84	81	77
Jamaica	103	1.00	72	63	54
Cabo Verde	102	0.99	69	60	47
Cyprus	96	0.97	57	47	39
Bahrain	77	1.00	17	13	10
South Georgia and South Sandwich Islands	74	0.65	28	24	21
Trinidad and Tobago	61	0.89	34	30	23
Brunei Darussalam	58	1.00	36	30	22
Lebanon	50	0.98	27	21	18
Vanuatu	49	0.98	42	38	30
Bonaire	47	1.00	42	40	38

French Polynesia	43	0.92	28	25	21
Antigua and Barbuda	36	0.97	26	24	21
Comoros	34	0.99	22	19	15
R??union	31	0.99	19	16	14
Guadeloupe	31	0.99	21	19	15
Equatorial Guinea	30	0.62	12	9	4
British Virgin Islands	30	0.99	24	22	19
US Virgin Islands	27	0.98	19	17	14
Seychelles	24	1.00	17	16	13
Samoa	23	0.98	19	17	15
Curacao	22	1.00	16	14	11
Martinique	21	1.00	14	13	8
Mauritius	19	0.89	15	14	13
Azores	19	0.97	15	14	12
Saint Lucia	17	1.00	14	12	9
Saint Pierre and Miquelon	17	0.99	14	13	11
Grenada	15	0.99	9	8	6
Bermuda	14	0.99	12	11	10
Cayman Islands	14	1.00	10	8	6
Isle of Man	14	0.95	9	7	4
Saint Martin	13	1.00	11	11	9
Malta	13	0.99	10	9	8
Andorra	13	0.89	4	1	1
Luxembourg	13	0.98	9	7	5
Mayotte	12	1.00	11	10	8
Guernsey	11	0.98	9	8	5
Saint Vincent and the Grenadines	11	1.00	8	8	5
Saint Kitts and Nevis	11	0.98	9	8	7
Glorioso Islands	10	0.98	6	4	3
Sint Maarten	10	1.00	8	8	7
Anguilla	9	1.00	8	7	7
Dominica	9	0.90	6	6	4
Palau	8	1.00	7	6	5
Aruba	6	1.00	5	5	4
Saint Barthelemy	6	1.00	5	5	5
Jersey	6	0.88	4	3	2
Tonga	5	1.00	5	5	4
Liechtenstein	4	0.76	2	2	1
Montserrat	4	1.00	1	1	1

Barbados	4	0.84	3	2	2
Guam	4	0.96	3	3	2
Northern Mariana Islands	4	0.99	3	3	2
Saint Eustatius	3	1.00	3	3	3
Maldives	2	1.00	2	2	2
Martinique	2	1.00	1	1	1
Madeira	1	0.98	1	1	1
Saba	1	1.00	1	1	0
Gibraltar	1	1.00	0	0	0
Tuvalu	1	1.00	1	0	0
United States Minor Outlying Islands	1	1.00	1	1	0
Monaco	1	1.00	0	0	0
San Marino	0	0.95	0	0	0
British Indian Ocean Territory	0	0.91	0	0	0
Vatican City	0	1.00	0	0	0
American Samoa	0	0.00	0	0	0
Bouvet Island	0	0.00	0	0	0
Heard Island and McDonald Islands	0	0.00	0	0	0
Nauru	0	0.00	0	0	0
Norfolk Island	0	0.00	0	0	0
Christmas Island	0	0.00	0	0	0
Cocos Islands	0	0.00	0	0	0
Saint Helena	0	0.00	0	0	0
Sao Tome and Principe	0	0.00	0	0	0
Pitcairn	0	0.00	0	0	0
Clipperton	0	0.00	0	0	0
Cook Islands	0	0.00	0	0	0
Niue	0	0.00	0	0	0
Tokelau	0	0.00	0	0	0
Wallis and Futuna	0	0.00	0	0	0
Asian very small islands	0	0.00	0	0	0
Marshall Islands	0	0.00	0	0	0
Micronesia	0	0.00	0	0	0
Faroe Islands	0	0.00	0	0	0
Jan Mayen	0	0.00	0	0	0
Juan De Nova Island	0	0.00	0	0	0

Chapter 5: The potential of time series Landsat and Sentinel-2 images for estimating chlorophyll-a content of open surface water bodies

Abstract

Algal blooms and the associated algal toxins, fish death and decrease of water quality have become a problem in Oklahoma, affecting the tourism industry, drinking water safety, and aquatic ecosystems. Chlorophyll-a concentration is an important indicator of algal blooms and water quality. However, field sampling and lab measurement of chlorophyll-a content is time-consuming and costly. Here, we evaluate the potential of time series Landsat and Sentinel 2 images for chlorophyll-a concentration estimation. Multiple stepwise regression analysis was used to explore the relationships between chlorophyll-a measurement and surface reflectance of satellite images. Regression models of Landsat data showed various performance in different water bodies across Oklahoma, with relatively good performance in Eufaula Lake, Keystone Lake, Copan Lake, Hugo Lake, Foss Reservoir, and Atoka Reservoir. The brightness temperature band of Landsat images was selected in one third of chlorophyll-a estimation models, indicating that temperature is among the most important factors of algal bloom in Oklahoma. The Red Edge 2 band of Sentinel 2 image showed great potential in chlorophyll-a estimation among different water sampling sites and water bodies across Oklahoma. Chlorophyll-a estimation using Landsat 5/7/8 and Sentinel 2 satellite images can serve as a supplement to the expensive in-situ field measurement in some lakes and reservoirs of Oklahoma.

5.1 Introduction

Chlorophyll-a concentration is one of the most important indicators of water quality. Oklahoma Water Resource Board uses the Carlson's Trophic State Index (TSI), derived from Chlorophyll-a concentration (Carlson, 1977), to represent the trophic status in lakes, reservoirs, and major rivers. According to the 2015 Oklahoma Lake Report of the Beneficial Use Monitoring Program (OWRB, 2015), during 2014–2015, six lakes (including Lake Thunderbird) were hypereutrophic (TSI $\geq$ 61) and 22 lakes were eutrophic (60 $\geq$ TSI $\geq$ 51), covering 7% and 88% of the total water bodies sampled, respectively. Harmful algal blooms are likely to occur in eutrophic and hypereutrophic lakes, which might cause fish mortality, jeopardize drinking water quality, and cause health problems. The invasive, toxigenic golden alga (Prymnesium parvum) is dispersed throughout the Lake Texoma with algal blooms and fish kills happening in specific parts of the lake (Hambright et al., 2010). The United States Army Corps of Engineers (USACE) had issued a warning of exposure to blue-green algal toxins (blue-green algae > 100,000 cells mL^{-1}) in many areas of Lake Texoma during Aug.25-31, 2011. According to the Grand River Dam Authority, Grand Lake experienced a harmful algal bloom event in the summer of 2010 and a lake-wide cyanobacteria monitoring program has been established since then.

While lab measurement of Chlorophyll-a concentration using water samples collected from the field can get the accurate results, this approach is time consuming and expensive. Thus, the Beneficial Use Monitoring Program collects water samples from the same sampling site in every year or every two to three years. This strategy might miss the algal blooms and fail to capture the seasonality of Chlorophyll-a concentration. Landsat

satellites scan the entire earth in every 16 days since 1980s while Sentinel 2 satellites scan the entire earth in every 10 days since Jun. 2015. If robust relationships can be found between chlorophyll-a field measurements and satellite remote sensing data, we can estimate chlorophyll-a concentration directly using Landsat and Sentinel 2 images. Satellite-image based water quality detection has the potential to provide Chlorophyll-a concentration information of the entire state in high temporal frequencies and spatial resolution, which could provide an alternative and supplement to the expensive in-situ water sample collection and lab measurement.

Landsat images have been used to quantify water chlorophyll-a concentration in previous studies (Dall'Olmo et al., 2003; O'Reilly et al., 1998). Compared with Landsat images, Sentinel 2 images have four additional red edge bands, designed to capture the reflectance of green vegetation and chlorophyll. Therefore, sentinel 2 images have the potential of chlorophyll-a estimation. Chlorophyll-a estimation algorithms can be different across different lakes, regulated by different optical properties, such as turbidity, suspended solids, and colored dissolved organic matter (Matsushita et al., 2015). For clear water, remote sensing reflectance at blue and green bands were used to estimate chlorophyll-a because they have a high signal-to-noise ratio at the first absorption peak of chlorophyll-a (O'Reilly et al., 1998). For turbid water, the red and near-infrared bands can better estimate chlorophyll-a concentration because they can provide a sufficient signal-to-noise ratio at the second absorption peak of chlorophyll-a (Dall'Olmo et al., 2003). However, if the water body contains significant amount of suspended solids and colored dissolved organic matters, chlorophyll-a estimation algorithms could be very different (Matsushita et al., 2015). This study aims to explore the potential of Landsat

5/7/8 and Sentinel 2 surface reflectance images in chlorophyll-a estimation at various Oklahoma lakes.

5.2 Materials and methods

5.2.1 Data.

We received 11,851 chlorophyll-a concentration field measurement records from Oklahoma Water Resource Board. These measurements were taken from the major lakes, reservoirs, and rivers during 2002–2016 by the Beneficial Use Monitoring Program in Oklahoma Water Resource Board. Most of the water samples were collected from ~0.5m depth. There were 908 water sample collection sites, distributing across the entire Oklahoma.

For each chlorophyll-a measurement record, its geographic coordinates of water sampling site were added to google earth engine. The pixel at the water sampling site was extracted from a Landsat image (Table 5.1), which has the closest acquisition date to the water sampling date compared to other Landsat images. In similar approach, a pixel was also extracted from a Sentinel 2 image (Table 5.1). Out of the 11,851 chlorophyll-a measurement records, we successfully extracted 11,369 Landsat pixels, of which 4938 were from Landsat 5, 5507 were from Landsat 7, and 924 were from Landsat 8. About 7% (816) of the extracted Landsat pixels were acquired by the satellites within 1 day of the field water sample collection, while ~80% (9082) of the extracted Landsat pixels were acquired by the satellites within 10 days of the field water sample collection (Table 5.2). In similar approach, we successfully extracted 176 Sentinel 2 pixels, of which 19% (34) were acquired by the satellites within 10 days of the field water sample collection. There were much less Sentinel 2 pixels because Sentinel 2A satellite was launched in Jun. 2015

and only a small portion of the chlorophyll-a measurements were acquired after that. In comparison, there are two Landsat satellites at work in most of the chlorophyll-a measurement period (2002–2016).

Table 5.1 Band name and wavelength of Landsat 5/7/8 and Sentinel 2 images.

Landsat 5/7		Landsat8		Sentinel 2A	
Band name	Wavelength (μm)	Band name	Wavelength (μm)	Band name	Wavelength (μm)
B1 (blue)	0.45-0.52	B1 (ultra-blue)	0.435-0.451	B1 (Aerosols)	0.4292-0.4562
B2 (green)	0.52-0.60	B2 (blue)	0.452-0.512	B2 (blue)	0.4434-0.5414
B3 (red)	0.63-0.69	B3 (green)	0.533-0.590	B3 (green)	0.5373-0.5823
B4 (near infrared)	0.77-0.90	B4 (red)	0.636-0.673	B4 (red)	0.6456-0.6836
B5 (shortwave infrared 1)	1.55-1.75	B5 (near infrared)	0.851-0.879	B5 (red edge 1)	0.6946-0.7136
B6 (brightness temperature)	10.40-12.50	B6 (shortwave infrared 1)	1.566-1.651	B6 (red edge 2)	0.7315-0.7495
B7 (shortwave infrared 2)	2.08-2.35	B7 (shortwave infrared 2)	2.107-2.294	B7 (red edge 3)	0.7688-0.7968
		B10 (brightness temperature 1)	10.60-11.19	B8 (near infrared)	0.7603-0.9053
		B11 (brightness temperature 2)	11.50-12.51	B8A (red edge 4)	0.8482-0.8812
				B11 (shortwave infrared 1)	1.5422-1.6852
				B12 (shortwave infrared 2)	2.0814-2.3234

Table 5.2 Satellite and day range of the extracted Landsat pixels

Satellite		Day range between satellite image and chlorophyll-a measurement	
Type	Num.	Range (day)	Num.
L5	4,938	0	816
L7	5,507	[1-5]	5,837
L8	924	[6-10]	2,429
		[11,15]	1,107
		[16,20]	614
		[21, 25]	233
		[25, 30]	143
		[31,94]	190
Sum	11,369	Sum	11,369

5.2.2 Regression analysis with Landsat data

Chlorophyll-a concentration associated with algal blooms changes across the seasons, affected by nutrients, temperature, sunlight, water chemistry, etc. Thus, satellite pixels acquired closer with the chlorophyll-a field measurement date could better represent the actual chlorophyll-a concentration. This study classified the chlorophyll-a measurements with a corresponding satellite pixel within 10 days as measurements qualified for further analysis. Out of the 908 water sampling sites, 768 have at least one qualified measurement. There were 420 water sampling sites that have at least 10 qualified measurements and only these sites were included into the stepwise multiple regression analysis in MATLAB R2014 a.

5.2.3 Regression analysis with Sentinel 2 data

Considering there were only 34 chlorophyll-a field measurements with corresponding satellite pixels within 10 days (Figure 5.1). We included these 34 measurements, from different water sampling sites across Oklahoma, into the multiple regression models with all the reflectance bands of Sentinel 2 pixels as potential factors (Table 5.1).

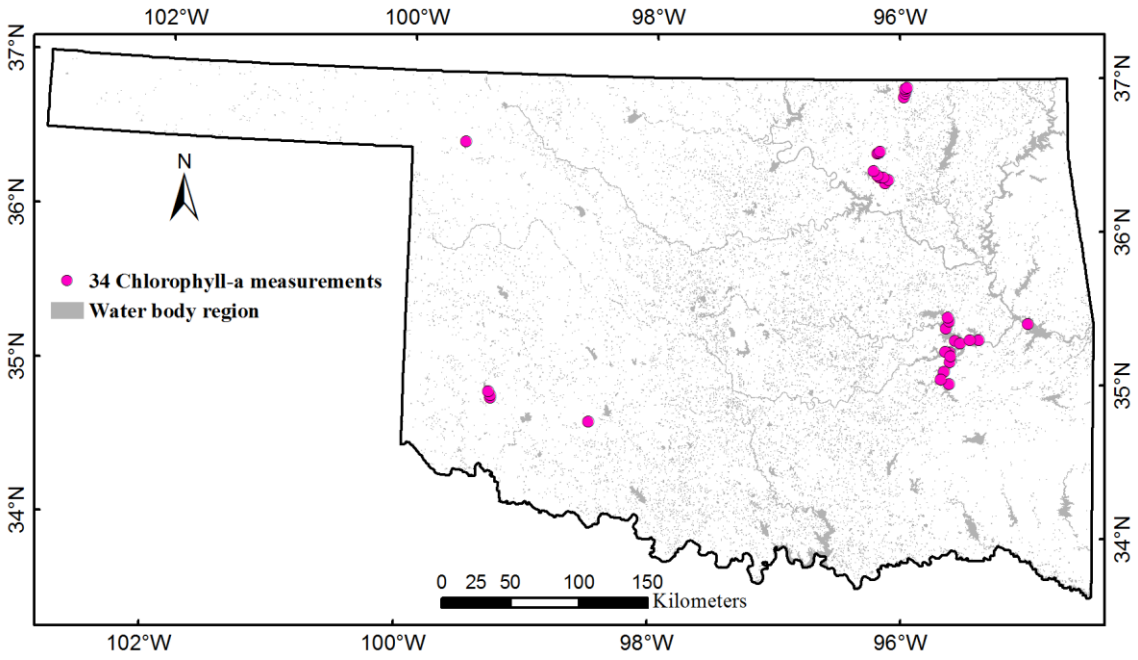


Figure 5.1 Locations of 34 chlorophyll-a measurements. Each measurement has a corresponding Sentinel 2 image pixel within 10 days.

5.3 Results and discussion

5.3.1 Chlorophyll-a estimation using Landsat data

Among the 420 water sampling sites qualified for stepwise multiple regression analysis, only 165 successfully built the regression models. Nine (5%) regression models had R squares <0.25, 77 (47%) had R squares between 0.25 and 0.5, 53 (32%) had R squares between 0.5 and 0.75, and 26 (16%) had R squares >0.75 (Figure 5.2, Table S5.1). The performance of regression models varies substantially across Oklahoma, with relatively good performance in Eufaula Lake, Keystone Lake, Copan Lake, Hugo Lake, Foss Reservoir, and Atoka Reservoir. The performance even varies across different water sampling sites within a lake. For example, model performance in northern Grand Lake and southern Oologah Lake was much better than the other portions of these two lakes. For different water sampling sites, the significant influencing bands selected by the

multiple stepwise regression models were also different (Table S5.1). Among the 165 multiple regression models, blue band was selected by 24 models, green band 31 models, red band 36 models, Near infrared band 34 models, Shortwave infrared-1 band 38 models, brightness temperature band 51 models, and shortwave infrared-2 band 21 models. These phenomena were likely caused by the difference in nutrients, temperature, suspended matters, and water chemistry (Matsushita et al., 2015) among different water bodies across Oklahoma. However, the specific influencing factors of chlorophyll-a concentration in each water body of Oklahoma remains unknown. Further studies will be required to solve this problem. It is worth to mention that brightness temperature band was included in about one third of all the regression models, more than any other bands, indicating that water temperature is one of the most important factors of algal bloom in Oklahoma and that this band has great potential in water chlorophyll-a estimation.

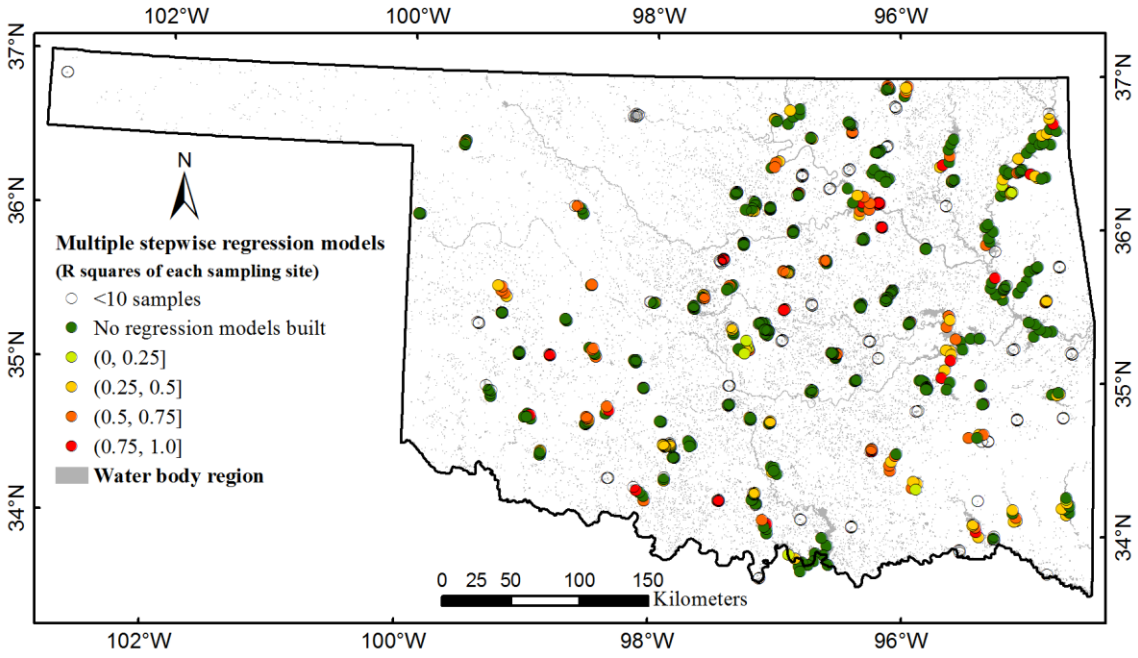


Figure 5.2 R squares of multiple stepwise regression models. 768 water sampling sites across the entire Oklahoma.

5.3.2 Chlorophyll-a estimation using Sentinel 2 images

A regression model was built using chlorophyll-a measurement and sentinel 2 data (Equation 5.1). Only the reflectance of Red Edge-2 band (Table 5.1) was selected by stepwise regression model. The R square of the regression model is 0.313, standard error of the estimate is 5.85, and the F value is 14.61 (P=0.001).

$$\text{Chol. A} = 372.392 \times \rho_{Red\ Edge\ 2} - 3.617 \qquad \text{(Equation 5.1)}$$

Where Chol. A is Chlorophyll-a concentration (mg/m^3), $\rho_{Red\ Edge\ 2}$ is surface reflectance of Red Edge 2 band (0.7315-0.7495 μm).

The model performance is not very good (R^2 =0.313), which is likely affected by the various water optical properties among the different water sampling sites and different lakes across Oklahoma (Figure 5.1). On the other hand, the Red Edge-2 Band showed significant linear relationships with chlorophyll-a concentration across different water bodies, indicating its great potential in chlorophyll-a content estimation.

5.4 Conclusions and perspective

Thousands of chlorophyll-a field measurements were used to evaluate the potential of Landsat 5/7/8 and Sentinel 2 images in chlorophyll-a estimation. Regression models were built in 165 water sampling sites with at least 10 measurements that have corresponding satellite data within 10 days. Regression models have various performance among different water sampling sites and water bodies across Oklahoma, with relatively good performance in Eufaula Lake, Keystone Lake, Copan Lake, Hugo Lake, Foss Reservoir, and Atoka Reservoir. The brightness temperature band of Landsat images was selected by one third of the regression models, indicating its great potential in chlorophyll-a estimation. It also indicates that temperature is one of the most important factors of algal

bloom in Oklahoma. The Red Edge-2 band of Sentinel 2 was selected as the only significant factor in the regression model built using 43 chlorophyll-a measurements from different water bodies across Oklahoma. This indicated the great potential of Red Edge 2 band in chlorophyll-a estimation. Sentinel 2A was launched in Jun. 2015 and Sentinel 2B was launched in Mar. 2017. There were only a few Sentinel 2A images overlapped with the chlorophyll-a field measurement data, collected during 2002-2016. The potential of Sentinel 2 data in chlorophyll-a estimation was not well evaluated because of the limited satellite images in our study period.

Supplementary materials

Table S5.1 Multiple stepwise regression models of 165 sampling sites.

ID	Num.	R^2	P	Inter.	Blue	Green	Red	NearIn	ShortIn1	TemC	ShortIn2
1	51	0.08	0.040	30.54	-298.93						
2	81	0.09	0.007	4.64						0.94	
3	85	0.22	0.000	10.66		-187.55				1.12	
4	20	0.22	0.039	2.82					60.61		
5	19	0.22	0.042	10.25					436.35		
6	35	0.23	0.004	-1.04			469.85				
7	46	0.23	0.003	5.78					-216.88	1.58	
8	20	0.24	0.030	3.00					99.93		
9	17	0.24	0.048	36.92							-840.22
10	16	0.25	0.048	25.06			-267.89				
11	16	0.25	0.048	7.83						-0.04	
12	17	0.25	0.039	8.45							663.43
13	18	0.25	0.033	12.84					614.41		
14	16	0.26	0.045	37.72	-507.59						
15	17	0.26	0.037	2.01					75.60		
16	19	0.26	0.024	10.46						-0.04	
17	17	0.26	0.035	0.87						0.15	
18	19	0.27	0.024	59.37			-480.68				
19	16	0.27	0.038	23.88			-162.89				
20	15	0.27	0.045	14.82	-172.77						

21	84	0.28	0.000	54.39	-751.93	-347.60		735.15		
22	17	0.28	0.030	3.95					0.01	
23	86	0.28	0.000	45.53		-501.63			457.89	
24	16	0.29	0.030	15.06			-80.80			
25	16	0.30	0.028	73.52			-575.11			
26	17	0.30	0.022	15.00	-169.46					
27	14	0.31	0.038	6.18					500.79	
28	14	0.32	0.036	21.38					-0.06	
29	13	0.33	0.039	3.56				113.52		
30	32	0.34	0.003	33.92		-428.91		315.67		
31	16	0.34	0.018	5.81					-0.04	
32	12	0.35	0.044	24.14	-223.72					
33	18	0.35	0.010	35.48	-414.20					
34	15	0.36	0.019	66.68				-862.70		
35	16	0.36	0.014	7.59					516.27	
36	34	0.36	0.001	27.85			-369.86		292.04	
37	17	0.37	0.010	26.73					-0.19	
38	13	0.37	0.028	26.32				-382.82		
39	18	0.37	0.007	-1.40				346.81		
40	16	0.37	0.012	63.72	-940.70					
41	19	0.37	0.005	0.58			70.59			
42	13	0.38	0.026	-0.06					0.65	
43	13	0.38	0.026	13.88					-0.07	
44	14	0.38	0.018	12.06						549.30
45	17	0.39	0.007	0.05					0.41	
46	11	0.39	0.039	63.44	-1007.45					
47	20	0.39	0.014	7.52		-91.91			153.47	
48	12	0.39	0.029	10.36					-0.06	
49	11	0.39	0.038	0.14					1.08	
50	15	0.40	0.012	30.08	-283.59					
51	18	0.40	0.005	2.68					-0.02	
52	13	0.40	0.020	10.43					264.93	
53	13	0.40	0.020	38.47			-319.28			
54	13	0.40	0.020	59.64		-616.21				
55	14	0.40	0.015	1.35					168.50	
56	19	0.41	0.003	-0.35				257.45		
57	27	0.41	0.002	7.70	-129.10			237.60		
58	11	0.41	0.034	10.52						-101.13
59	14	0.41	0.013	1.77					0.78	

60	11	0.41	0.033	5.60				83.69			
61	18	0.42	0.004	45.41		-438.98					
62	12	0.42	0.022	34.93	-429.43						
63	17	0.43	0.004	15.57			-91.88				
64	15	0.43	0.008	5.64						1.75	
65	19	0.43	0.011	26.41	-433.58			276.36			
66	13	0.44	0.014	62.01		-475.79					
67	14	0.44	0.009	-1.18						1.03	
68	32	0.45	0.000	51.98		-645.54		373.54			
69	14	0.45	0.009	38.34		-376.47					
70	17	0.45	0.003	29.38						-0.20	
71	11	0.46	0.022	22.88		-111.25					
72	11	0.47	0.021	44.11		-492.73					
73	15	0.47	0.005	14.85						1.94	
74	16	0.48	0.015	39.81		-470.10		273.97			
75	13	0.48	0.009	7.96					460.53		
76	17	0.48	0.002	22.16			-132.73				
77	35	0.48	0.000	40.14		-477.09		275.59			
78	15	0.49	0.004	27.47						-0.21	
79	17	0.49	0.002	35.03		-295.55					
80	14	0.49	0.005	19.17						0.06	
81	18	0.49	0.006	28.96		-353.92		240.10			
82	19	0.49	0.004	9.22		-118.88					201.10
83	18	0.49	0.006	49.72			-602.81		455.25		
84	19	0.50	0.001	50.33		-451.64					
85	18	0.50	0.006	12.96	-122.83						319.30
86	12	0.50	0.010	8.67					402.16		
87	14	0.50	0.004	11.74							508.30
88	10	0.51	0.021	2.19					98.66		
89	17	0.51	0.001	1.68						0.77	
90	18	0.51	0.004	6.51			-183.11				339.58
91	13	0.52	0.006	3.74					855.63		
92	22	0.52	0.001	56.72			-2108.12	1734.11			
93	11	0.53	0.012	1.35				85.30			
94	13	0.53	0.005	23.96		-165.02					
95	25	0.53	0.001	30.31	-385.03						-1323.36
96	12	0.53	0.007	10.97						-0.08	
97	16	0.54	0.001	-0.48							1522.94
98	15	0.54	0.002	8.62						-0.05	

99	10	0.55	0.014	2.88					606.57		
100	10	0.56	0.013	2.87				-44.73			
101	14	0.56	0.002	16.62						0.05	
102	14	0.56	0.002	67.08		-652.25					
103	16	0.56	0.005	13.48			-219.40		554.99		
104	14	0.56	0.002	1.79							127.28
105	11	0.57	0.007	2.47						0.78	
106	100	0.57	0.000	30.53	989.64	-2344.13	1396.67			1.19	
107	20	0.57	0.001	103.05		-1836.71		1291.41			
108	14	0.57	0.002	14.52						0.05	
109	13	0.58	0.002	0.95						0.58	
110	16	0.59	0.003	1.21						0.58	-255.59
111	11	0.59	0.006	0.83				93.44			
112	20	0.60	0.000	49.01		-613.41			664.23		
113	15	0.60	0.004	4.40					1958.17		-1406.61
114	17	0.61	0.001	12.90			394.93	-298.79			
115	18	0.61	0.001	16.03	-160.91						232.93
116	13	0.61	0.009	30.19		-410.98				1.84	
117	11	0.62	0.004	2.56				126.47			
118	13	0.62	0.001	80.32			-649.61				
119	10	0.62	0.007	4.78				-62.02			
120	14	0.62	0.005	19.73			-274.85	425.91			
121	10	0.63	0.006	26.01			-154.14				
122	16	0.64	0.000	31.67			-178.74				
123	51	0.64	0.000	37.05		-468.49		238.52		-0.02	
124	17	0.65	0.001	62.24		-606.39					495.34
125	14	0.65	0.000	28.07				-417.35			
126	13	0.66	0.005	27.46			-416.33	381.38			
127	12	0.66	0.001	11.70						-0.10	
128	12	0.68	0.001	21.35						-0.09	
129	10	0.69	0.003	-0.90						2.32	
130	11	0.70	0.001	2.55					199.74		
131	15	0.70	0.000	0.19					517.15		
132	10	0.71	0.002	3.52						-0.02	
133	12	0.71	0.004	15.18		-218.23		283.59			
134	16	0.71	0.000	15.30		-119.76			204.61		
135	16	0.71	0.001	8.41					-1036.41	0.31	
136	17	0.72	0.001	16.33			-270.48	482.31	-311.86		
137	16	0.72	0.000	-15.51	355.94						-238.26

138	12	0.73	0.003	-0.98			64.52			-0.01	
139	11	0.74	0.001	15.06			-70.68				
140	13	0.77	0.000	30.99				-520.84			
141	12	0.77	0.001	11.99			-88.00		324.64		
142	10	0.77	0.001	4.64						-0.03	
143	12	0.77	0.001	16.84			-158.46				1000.63
144	10	0.77	0.001	-12.16				482.98			
145	13	0.78	0.000	-2.05						0.83	
146	16	0.79	0.000	61.78			-792.72	-481.28			
147	12	0.79	0.004	9.49			-85.01		659.14		-542.67
148	11	0.79	0.000	3.62					252.65		
149	11	0.80	0.002	2.45					111.11	-0.02	
150	23	0.80	0.000	17.67						-0.26	-712.63
151	17	0.81	0.000	15.16			-96.30				134.64
152	10	0.83	0.000	-11.88					973.36		
153	12	0.85	0.000	1.24					41.95	0.00	
154	17	0.85	0.000	7.07			-23.27		1432.00		-520.33
155	14	0.86	0.000	32.43		-403.33			680.73		
156	11	0.86	0.000	10.56	-230.10					0.65	
157	10	0.87	0.000	-6.72						0.96	
158	12	0.91	0.000	0.73					-236.81	0.95	
159	10	0.92	0.001	5.24	-197.75		115.81			0.00	
160	13	0.94	0.000	11.40		1136.62	-1198.21		328.67		
161	12	0.95	0.000	18.83	-636.14		416.63			0.30	
162	10	0.95	0.000	18.06	-555.81		147.34	188.00			
163	10	0.96	0.000	8.13				338.37		0.51	
164	11	0.97	0.000	-2382.2	42498.97	-3093.11	35746.46				
165	10	0.97	0.000	54.22	-1349.40			824.83	-657.72		

ID is sampling site ID. Num. is short for sample number at each site. R^2 and P are from regression model summary. Inter. is short for intercept. Blue, Green, Red, NearIn, ShortIn1, TemC, and ShortIn2 are Landsat 5/7/8 bands of Blue, Green, Red, Near infrared, Shortwave infrared 1, brightness temperature, and shortwave infrared 2.

Chapter 6: Conclusions and perspectives

This dissertation used remote sensing data and techniques to map open surface water bodies at state, national, and global scales and analyzed their interannual variations and multi-decadal trends. Under the impact of climate change and anthropogenic activities, surface water area had changed significantly in the past three decades across the globe, threatening human society, natural environment, and global biodiversity. Divergent trends between surface water area and GRACE land water storage were found across the globe because of climate change, agriculture irrigation, dam construction, land reclamation, groundwater mining, etc. The Landsat 5/7/8 and Sentinel 2 satellite images were used to estimate water chlorophyll-a content across various lakes, reservoirs, and major rivers in Oklahoma. Great potential was found in remote sensing estimation of water chlorophyll-a concentration.

Chapter 2 developed robust water body mapping algorithms using the relationships between water and vegetation indices and apply them to Oklahoma to generate water body maps. Change analysis of water body area and number was carried out in four types of annual water body extents (the maximum, year-long, seasonal, and annual average extents). Statistically significant downward trends were found in both water body area (the maximum, year-long, and annual average water body area) and water body number (maximum and year-long water body number) from 1984 to 2015. The decrease in water body area was mainly attributed to the continued shrinking of large water bodies, while the decrease of water body number was caused by the vanishing of some small water bodies. Besides significant decreasing trends, remarkable inter-annual variations of water body area and number were also found. Precipitation had positive

effects on water body area and number dynamics while temperature had negative effects. Small water bodies are more vulnerable under climate-warming scenarios.

Chapter 3 modified the water body mapping algorithms and applied them to the CONUS to generate water body maps for further analysis. This chapter found the uneven water resource distribution across the CONUS with the western half US having less water body area but stronger interannual variability compared with the eastern half. Divergent trends of open surface water body area in the last three decades, mainly driven by climate, have made the water-poor regions of Southwest and Northwest US poorer, while the water-rich regions of Southeast US and far north Great Plains richer. Water resource gaps across the CONUS were enlarged by these divergent trends. Water body area change is a good indicator of land water storage dynamics in 58% of the CONUS. Surface water body shrinkage in prolonged drought years had led to massive groundwater mining and the rapid decrease of land water storage in California and the Southern Great Plains. These findings can aid decision makers and stakeholders across the CONUS, especially in the West, to develop and implement water resource planning and management in coping with the increasing water stress, unprecedented droughts, and uncertain impacts of climate change.

Chapter 4 further modified the water mapping algorithms and applied them to the entire world to generate a new multi-decadal surface water body dataset. Compared to the only existing multi-decadal water body dataset (1984–2015), released by the Joint Research Center (JRC) in 2016, our new dataset has a longer study period (1984–2017 vs 1984–2015), a larger spatial extent (60°S–85°N vs 60°S–78°N), and more Landsat data input (3.8 vs 3.1 million images). During 1984–2017, 8.5 million 0.01° grid cells had

significant increasing or decreasing trends in surface water area. These grid cells were distributed across the globe and formed interesting and surprising spatial patterns in many regions, such as northern Greenland, the Tibetan Plateau, central North America, the western US, the Great Lakes, the Gulf of Bothnia, and central South America. These interesting spatial patterns across large regions revealed unprecedented details of surface water dynamics, which were not reported in previous studies. The time-series curves of regional surface water area at 5° tiles during 1984–2017 revealed the interannual variations, the magnitude of variability, and multi-decadal trends. There were 189 five-degree tiles with significant increasing trends and 170 tiles with significant decreasing trends. Divergent trends between land water storage and regional surface water area were found in Greenland, China, the Indus Basin, and central Africa. The significant changes in surface water area and land water storage were mainly driven by climate and anthropogenic activities, which will be further aggravated by population growth and increasing food demand. The dataset and findings of this study are critical for water resource research, planning, and management in coping with water scarcity and food security associated with climate change and population growth.

Chapter 5 explored the relationship between chlorophyll-a field measurements and surface reflectance data from Landsat 5/7/8 and Sentinel 2 satellites using multiple stepwise regression analysis. Regressions models derived from Landsat data changed among water sampling sites and water bodies across Oklahoma, regulated by nutrients, temperature, water chemistry, etc. The brightness temperature band of Landsat satellites showed great potential in chlorophyll-a estimation. This indicated that temperature was among the most important factors of algal blooms in Oklahoma. The Red Edge-2 band of

Sentinel 2 Satellite showed great potential in chlorophyll-a estimation among different water sampling sites and water bodies across Oklahoma. Our results indicated that Landsat 5/7/8 and Sentinel 2 Satellite images have great potential in chlorophyll-a estimation, which can serve as an alternative and supplement to the expensive field chlorophyll-a sampling and measurement.

References

Aherne, J., T. Larssen, B.J. Cosby, and P.J. Dillon (2006), Climate variability and forecasting surface water recovery from acidification: Modelling drought-induced sulphate release from wetlands, *Science of the Total Environment*, *365*(1-3), 186-199.

Alsdorf, D.E., E. Rodriguez, and D.P. Lettenmaier (2007), Measuring surface water from space, *Reviews of Geophysics*, *45*(2), 1-24.

Argonne National Laboratory (2012), Impacts of Long-Term Drought on Power Systems in the U.S. Southwest, prepared for the U.S. Department of Energy, U.S. DOE, Washington, D.C. [Available at https://energy.gov/sites/prod/files/Impacts%20of%20Long-term%20Drought%20on%20Power%20Systems%20in%20the%20US%20Southwest%20%E2%80%93%20July%202012.pdf. Accessed December 27, 2016.

Assel, R.A., F.H. Quinn, and C.E. Sellinger (2004), Hydroclimatic factors of the recent record drop in Laurentian Great Lakes water levels, *Bulletin of the American Meteorological Society*, *85*(8), 1143-1152.

Atkinson, C.L., J.P. Julian, and C.C. Vaughn (2014), Species and function lost: Role of drought in structuring stream communities, *Biological Conservation*, *176*, 30-38.

Averyt, K., J. Meldrum, P. Caldwell, G. Sun, S. McNulty, A. Huber-Lee, and N. Madden (2013), Sectoral contributions to surface water stress in the coterminous United States, *Environmental Research Letters*, *8*(3), 035046.

Bates, B.C., Z.W. Kundzewicz, S. Wu, and J.P. Palutikof (2008), *Climate Change and Water. Technical Paper of the Intergovernmental Panel on Climate Change*, IPCC Secretariat, Geneva.

Beard, G., E. Chandler, A. Watkins, and D. Jones (2011), How does the 2010–11 La Niña compare with past La Niña events, *Bulletin of the Australian Meteorological and Oceanographic Society*, *24*, 17-20.

Bhagat, V.S., and K.R. Sonawane (2011), Use of Landsat ETM plus data for delineation of water bodies in hilly zones, *Journal of Hydroinformatics*, *13*(4), 661-671.

Blanc, E., K. Strzepek, A. Schlosser, H. Jacoby, A. Gueneau, C. Fant, S. Rausch, and J. Reilly (2014), Modeling US water resources under climate change, *Earths Future*, *2*(4), 197-224.

Blum, M.D., and H.H. Roberts (2009), Drowning of the Mississippi Delta due to insufficient sediment supply and global sea-level rise, *Nature Geoscience*, *2*(7), 488-491.

Boland, D.H.P. (1976), Trophic classification of lakes using Landsat-1 (ERTS-1) multispectral scanner data, U.S. Environmental Protection Agency, Office of Research and Development, Corvallis Environmental Research Laboratory, Assessment and Criteria Development, Corvallis, Oregon.

Brena-Naranjo, J.A., A.D. Kendall, and D.W. Hyndman (2014), Improved methods for satellite-based groundwater storage estimates: A decade of monitoring the high plains aquifer from space and ground observations, *Geophysical Research Letters*, *41*(17), 6167-6173.

Brown, C., and U. Lall (2006), Water and economic development: The role of variability and a framework for resilience, *Natural Resources Forum*, *30*(4), 306-317.

Brunner, P., P.G. Cook, and C.T. Simmons (2009), Hydrogeologic controls on disconnection between surface water and groundwater, *Water Resources Research*, *45*, W01422.

Caldwell, P.V., G. Sun, S.G. McNulty, E.C. Cohen, and J.A.M. Myers (2012), Impacts of impervious cover, water withdrawals, and climate change on river flows in the conterminous US, *Hydrology and Earth System Sciences*, *16*(8), 2839-2857.

California Department of Water Resources (2015), California's Most Significant Droughts: Comparing Historical and Recent Conditions, pp. 1-136, California Department of Water Resources, Sacramento, CA. [Available at http://www.water.ca.gov/waterconditions/docs/California_Signficant_Droughts_2015_small.pdf. Accessed September 18, 2017.]

Campfield, Z. (2013), Drought sinks water levels at Oklahoma City's Lake Hefner to all-time low, NewsOK, Oklahoma. [Available at http://newsok.com/article/3744544. Accessed December 27, 2016.]

Carlson, R.E. (1977), A trophic state index for lakes, *Limnology and Oceanography*, *22*(2), 361-369.

Carr, J.R., C. Stokes, and A. Vieli (2014), Recent retreat of major outlet glaciers on Novaya Zemlya, Russian Arctic, influenced by fjord geometry and sea-ice conditions, *Journal of Glaciology*, *60*(219), 155-170.

Cayan, D.R., T. Das, D.W. Pierce, T.P. Barnett, M. Tyree, and A. Gershunov (2010), Future dryness in the southwest US and the hydrology of the early 21st century drought, *Proceedings of the National Academy of Sciences of the United States of America*, *107*(50), 21271-21276.

Chabwela, H., C. Chomba, and L. Thole (2017), The Habitat Structure of Lukanga Ramsar Site in Central Zambia: An Understanding of Wetland Ecological Condition, *Open Journal of Ecology*, *7*(06), 406-433.

Chander, G., B.L. Markham, and D.L. Helder (2009), Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors, *Remote Sensing of Environment*, *113*(5), 893-903.

Chen, J.L., T. Pekker, C.R. Wilson, B.D. Tapley, A.G. Kostianoy, J.F. Cretaux, and E.S. Safarov (2017), Long-term Caspian Sea level change, *Geophysical Research Letters*, *44*(13), 6993-7001.

Cheng, G.D., X. Li, W.Z. Zhao, Z.M. Xu, Q. Feng, S.C. Xiao, and H.L. Xiao (2014), Integrated study of the water-ecosystem-economy in the Heihe River Basin, *National Science Review*, *1*(3), 413-428.

Crist, E.P. (1985), A Tm Tasseled Cap Equivalent Transformation for Reflectance Factor Data, *Remote Sensing of Environment*, *17*(3), 301-306.

Dale, J., C.B. Zou, W.J. Andrews, J.M. Long, Y. Liang, and L. Qiao (2015), Climate, water use, and land surface transformation in an irrigation intensive watershed - Streamflow responses from 1950 through 2010, *Agricultural Water Management*, *160*, 144-152.

Dall'Olmo, G., A.A. Gitelson, and D.C. Rundquist (2003), Towards a unified approach for remote estimation of chlorophyll-a in both terrestrial vegetation and turbid productive waters, *Geophysical Research Letters*, *30*(18), 1-4.

Deng, H.J., and Y.N. Chen (2017), Influences of recent climate change and human activities on water storage variations in Central Asia, *Journal of Hydrology*, *544*, 46-57.

DESA (2017), *World Population Prospects: The 2017 Revision, Data Booklet. ST/ESA/SER.A/401*, United Nations Department of Economic and Social Affairs, Population Division, New York.

Dong, J.W., et al. (2015), Tracking the dynamics of paddy rice planting area in 1986-2010 through time series Landsat images and phenology-based algorithms, *Remote Sensing of Environment*, *160*, 99-113.

Dong, X., B. Xi, A. Kennedy, Z. Feng, J.K. Entin, P.R. Houser, R.A. Schiffer, T. L'Ecuyer, W.S. Olson, and K.l. Hsu (2011), Investigation of the 2006 drought and 2007 flood extremes at the Southern Great Plains through an integrative analysis of observations, *Journal of Geophysical Research: Atmospheres*, *116*(D3), D03204.

Du, Y., H.P. Xue, S.J. Wu, F. Ling, F. Xiao, and X.H. Wei (2011), Lake area changes in the middle Yangtze region of China over the 20th century, *Journal of Environmental Management*, *92*(4), 1248-1255.

Du, Z.Q., L.H. Bin, F. Ling, W.B. Li, W.D. Tian, H.L. Wang, Y.M. Gui, B.Y. Sun, and X.M. Zhang (2012), Estimating surface water area changes using time-series Landsat data in the Qingjiang River Basin, China, *Journal of Applied Remote Sensing*, *6*, 1-16.

Elisa, M., J.I. Gara, and E. Wolanski (2010), A review of the water crisis in Tanzania's protected areas, with emphasis on the Katuma River—Lake Rukwa ecosystem, *Ecohydrology & Hydrobiology*, *10*(2-4), 153-165.

ESRI (2018), World Countries, edited by ESRI, DeLorme Publishing Company, Inc., Yarmouth, Maine, United States

Evora, N.D., D. Tapsoba, and D. De Seve (2008), Combining artificial neural network models, geostatistics, and passive microwave data for snow water equivalent retrieval and mapping, *IEEE Transactions on Geoscience and Remote Sensing*, *46*(7), 1925-1939.

Ezer, T. (2013), Sea level rise, spatially uneven and temporally unsteady: Why the US East Coast, the global tide gauge record, and the global altimeter data show different trends, *Geophysical Research Letters*, *40*(20), 5439-5444.

Fallatah, O.A., M. Ahmed, H. Save, and A.S. Akanda (2017), Quantifying temporal variations in water resources of a vulnerable middle eastern transboundary aquifer system, *Hydrological Processes*, *31*(23), 4081-4091.

Famiglietti, J.S. (2004), Remote sensing of terrestrial water storage, soil moisture and surface waters, *The state of the planet: frontiers and challenges in geophysics*, *150*, 197-207.

Famiglietti, J.S. (2014), The global groundwater crisis, *Nature Climate Change*, *4*(11), 945-948.

Farr, T.G., et al. (2007), The shuttle radar topography mission, *Reviews of Geophysics*, *45*(2), 1-33.

Feng, L., C.M. Hu, X.L. Chen, and R.F. Li (2011), Satellite observations make it possible to estimate Poyang Lake's water budget, *Environmental Research Letters*, *6*(4), 1-7.

Feng, M., J.O. Sexton, S. Channan, and J.R. Townshend (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.

Feng, Z., L.R. Leung, S. Hagos, R.A. Houze, C.D. Burleyson, and K. Balaguru (2016), More frequent intense and long-lived storms dominate the springtime trend in central US rainfall, *Nature Communications*, *7*, 1-8.

Ferguson, I.M., and R.M. Maxwell (2012), Human impacts on terrestrial hydrology: climate change versus pumping and irrigation, *Environmental Research Letters*, *7*(4).

Fettweis, X. (2007), Reconstruction of the 1979-2006 Greenland ice sheet surface mass balance using the regional climate model MAR, *Cryosphere*, *1*(1), 21-40.

Feyisa, G.L., H. Meilby, R. Fensholt, and S.R. Proud (2014), Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery, *Remote Sensing of Environment*, *140*, 23-35.

Fisher, A., N. Flood, and T. Danaher (2016), Comparing Landsat water index methods for automated water classification in eastern Australia, *Remote Sensing of Environment*, *175*, 167-182.

Fitzpatrick, A.A.W., A.L. Hubbard, J.E. Box, D.J. Quincey, D. van As, A.P.B. Mikkelsen, S.H. Doyle, C.F. Dow, B. Hasholt, and G.A. Jones (2014), A decade (2002-2012) of supraglacial lake volume estimates across Russell Glacier, West Greenland, *Cryosphere*, *8*(1), 107-121.

Frolova, N.L., P.A. Belyakova, V.Y. Grigoriev, A.A. Sazonov, L.V. Zotov, and J. Jarsjo (2017), Runoff fluctuations in the Selenga River Basin, *Regional Environmental Change*, *17*(7), 1965-1976.

Fry, J.A., G. Xian, S. Jin, J.A. Dewitz, C.G. Homer, Y. Limin, C.A. Barnes, N.D. Herold, and J.D. Wickham (2011), Completion of the 2006 national land cover database for the conterminous United States, *Photogrammetric Engineering and Remote Sensing*, *77*(9), 858-864.

Gao, B.C. (1996), NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space, *Remote Sensing of Environment*, *58*(3), 257-266.

Gaupp, F., J. Hall, and S. Dadson (2015), The role of storage capacity in coping with intra- and inter-annual water variability in large river basins, *Environmental Research Letters*, *10*(12), 125001.

Gibson, A.M. (1981), *Oklahoma, a history of five centuries*, University of Oklahoma Press, Oklahoma.

Gond, V., E. Bartholome, F. Ouattara, A. Nonguierma, and L. Bado (2004), Monitoring and mapping of waters and wetlands in arid regions using the SPOT-4 VEGETATION imaging system, *International Journal of Remote Sensing*, *25*(5), 987-1004.

Gons, H.J., M.T. Auer, and S.W. Effler (2008), MERIS satellite chlorophyll mapping of oligotrophic and eutrophic waters in the Laurentian Great Lakes, *Remote Sensing of Environment*, *112*(11), 4098-4106.

Good, S.P., C.D. Kennedy, J.C. Stalker, L.A. Chesson, L.O. Valenzuela, M.M. Beasley, J.R. Ehleringer, and G.J. Bowen (2014), Patterns of local and nonlocal water resource use across the western US determined via stable isotope intercomparisons, *Water Resources Research*, *50*(10), 8034-8049.

Google Earth Engine (2017), USGS Landsat 5/7/8 Surface Reflectance (pre-Collection), Google Earth Engine, California. [Available at https://explorer.earthengine.google.com/#detail/LANDSAT%2FLC8_SR. Accessed July 10, 2017.]

Gronewold, A.D., V. Fortin, B. Lofgren, A. Clites, C.A. Stow, and F. Quinn (2013), Coasts, water levels, and climate change: A Great Lakes perspective, *Climatic Change*, *120*(4), 697-711.

Hall, J.W., D. Grey, D. Garrick, F. Fung, C. Brown, S.J. Dadson, and C.W. Sadoff (2014), Water Security. Coping with the curse of freshwater variability, *Science*, *346*(6208), 429-430.

Hambright, K.D., R.M. Zamor, J.D. Easton, K.L. Glenn, E.J. Remmel, and A.C. Easton (2010), Temporal and spatial variability of an invasive toxigenic protist in a North American subtropical reservoir, *Harmful Algae*, *9*(6), 568-577.

Heimhuber, V., M.G. Tulbure, and M. Broich (2017), Modeling multidecadal surface water inundation dynamics and key drivers on large river basin scale using multiple time series of Earth-observation and river flow data, *Water Resources Research*, *53*(2), 1251-1269.

Henits, L., C. Jurgens, and L. Mucsi (2016), Seasonal multitemporal land-cover classification and change detection analysis of Bochum, Germany, using multitemporal Landsat TM data, *International Journal of Remote Sensing*, *37*(15), 3439-3454.

Hoerling, M., S. Schubert, K. Mo, A. Aghakouchak, H. Berbery, and J. Dong (2013), An interpretation of the origins of the 2012 central great plains drought, NOAA Drought Task Force Narrative Team, Silver Spring, MD. [Available at https://www.esrl.noaa.gov/psd/csi/factsheets/pdf/noaa-gp-drought-assessment-report.pdf.]

Hoerling, M., J. Eischeid, A. Kumar, R. Leung, A. Mariotti, K. Mo, S. Schubert, and R. Seager (2014), Causes and Predictability of the 2012 Great Plains Drought, *Bulletin of the American Meteorological Society*, *95*(2), 269-282.

Homer, C., C.Q. Huang, L.M. Yang, B. Wylie, and M. Coan (2004), Development of a 2001 National Land-Cover Database for the United States, *Photogrammetric Engineering and Remote Sensing*, *70*(7), 829-840.

Homer, C., J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. McKerrow, J.N. VanDriel, and J. Wickham (2007), Completion of the 2001 National Land Cover Database for the conterminous United States, *Photogrammetric Engineering and Remote Sensing*, *73*(4), 337-341.

Homer, C., J. Dewitz, L.M. Yang, S. Jin, P. Danielson, G. Xian, J. Coulston, N. Herold, J. Wickham, and K. Megown (2015), Completion of the 2011 National Land Cover Database for the Conterminous United States - Representing a Decade of

Land Cover Change Information, *Photogrammetric Engineering and Remote Sensing*, *81*(5), 345-354.
Howitt, R., J. Medellín-Azuara, D. MacEwan, J.R. Lund, and D. Sumner (2014), *Economic analysis of the 2014 drought for California agriculture*, Center for Watershed Sciences University of California, Davis, CA.
Hui, F.M., B. Xu, H.B. Huang, Q. Yu, and P. Gong (2008), Modelling spatial-temporal change of Poyang Lake using multitemporal Landsat imagery, *International Journal of Remote Sensing*, *29*(20), 5767-5784.
Hulley, G.C., S.J. Hook, E. Abbott, N. Malakar, T. Islam, and M. Abrams (2015), The ASTER Global Emissivity Dataset (ASTER GED): Mapping Earth's emissivity at 100 meter spatial scale, *Geophysical Research Letters*, *42*(19), 7966-7976.
Ji, L., L. Zhang, and B. Wylie (2009), Analysis of Dynamic Thresholds for the Normalized Difference Water Index, *Photogrammetric Engineering and Remote Sensing*, *75*(11), 1307-1317.
Johansson, M.M., K.K. Kahma, H. Boman, and J. Launiainen (2004), Scenarios for sea level on the Finnish coast, *Boreal Environment Research*, *9*(2), 153-166.
Johnson, K.S., and K.V. Luza (2008), *Earth Sciences and Mineral Resources of Oklahoma. Educational Publication EP-9*, Oklahoma Geological Survey, Norman, Oklahoma.
Junk, W.J., and C.N. de Cunha (2005), Pantanal: a large South American wetland at a crossroads, *Ecological Engineering*, *24*(4), 391-401.
Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. Hnilo, M. Fiorino, and G. Potter (2002), Ncep–doe amip-ii reanalysis (r-2), *Bulletin of the American Meteorological Society*, *83*(11), 1631-1644.
Kogan, F., and W. Guo (2015), 2006-2015 mega-drought in the western USA and its monitoring from space data, *Geomatics Natural Hazards & Risk*, *6*(8), 651-668.
Kong, D.X., C.Y. Miao, A.G.L. Borthwick, Q.Y. Duan, H. Liu, Q.H. Sun, A.Z. Ye, Z.H. Di, and W. Gong (2015), Evolution of the Yellow River Delta and its relationship with runoff and sediment load from 1983 to 2011, *Journal of Hydrology*, *520*, 157-167.
Krueger, E.S., Y.T. Yimam, and T.E. Ochsner (2017), Human factors were dominant drivers of record low streamflow to a surface water irrigation district in the US southern Great Plains, *Agricultural Water Management*, *185*, 93-104.
Kuang, X.X., and J.J. Jiao (2016), Review on climate change on the Tibetan Plateau during the last half century, *Journal of Geophysical Research-Atmospheres*, *121*(8), 3979-4007.
Laghari, A.N., D. Vanham, and W. Rauch (2012), The Indus basin in the framework of current and future water resources management, *Hydrology and Earth System Sciences*, *16*(4), 1063-1083.
LaPotin, P., R. Kennedy, T. Pangburn, and R. Bolus (2001), Blended spectral classification techniques for mapping water surface transparency and chlorophyll concentration, *Photogrammetric Engineering and Remote Sensing*, *67*(9), 1059-1065.
Lehner, B., and P. Döll (2004), Development and validation of a global database of lakes, reservoirs and wetlands, *Journal of Hydrology*, *296*(1-4), 1-22.

Li, E.R., P.J. Du, A. Samat, J.S. Xia, and M.Q. Che (2015), An automatic approach for urban land-cover classification from Landsat-8 OLI data, *International Journal of Remote Sensing*, *36*(24), 5983-6007.

Li, W.B., Z.Q. Du, F. Ling, D.B. Zhou, H.L. Wang, Y.M. Gui, B.Y. Sun, and X.M. Zhang (2013), A Comparison of Land Surface Water Mapping Using the Normalized Difference Water Index from TM, ETM plus and ALI, *Remote Sensing*, *5*(11), 5530-5549.

Liao, A.P., L.J. Chen, J. Chen, C.Y. He, X. Cao, J. Chen, S. Peng, F.D. Sun, and P. Gong (2014), High-resolution remote sensing mapping of global land water, *Science China-Earth Sciences*, *57*(10), 2305-2316.

Lin, B., P. Stackhouse, W.B. Sun, Y.X. Hu, Z.Y. Liu, and T.F. Fan (2013), Is Oklahoma getting drier?, *Journal of Quantitative Spectroscopy & Radiative Transfer*, *122*, 208-213.

Liu, H.Y., Y. Yin, S.L. Piao, F.J. Zhao, M. Engels, and P. Ciais (2013), Disappearing Lakes in Semiarid Northern China: Drivers and Environmental Impact, *Environmental Science & Technology*, *47*(21), 12107-12114.

MacDonald, G.M., D.W. Stahle, J.V. Diaz, N. Beer, S.J. Busby, J. Cerano‐Paredes, J.E. Cole, E.R. Cook, G. Endfield, and G. Gutierrez‐Garcia (2008), Climate Warming and 21st‐Century Drought in Southwestern North America, *EOS, Transactions American Geophysical Union*, *89*(9), 82-82.

Matsushita, B., W. Yang, G.L. Yu, Y. Oyama, K. Yoshimura, and T. Fukushima (2015), A hybrid algorithm for estimating the chlorophyll-a concentration across different trophic states in Asian inland waters, *Isprs Journal of Photogrammetry and Remote Sensing*, *102*, 28-37.

McFeeters, S.K. (1996), The use of the normalized difference water index (NDWI) in the delineation of open water features, *International Journal of Remote Sensing*, *17*(7), 1425-1432.

McGuire, V.L. (2014), Water-level changes and change in water in storage in the High Plains aquifer, predevelopment to 2013 and 2011-13: U.S. Geological Survey Scientific Investigations Report 2014-5218*Rep. 2328-0328*, Department of the Interior U.S. Geological Survey, Washington, D.C.]

McPherson, R.A., et al. (2007), Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet, *Journal of Atmospheric and Oceanic Technology*, *24*(3), 301-321.

Mei, X.F., Z.J. Dai, J.Z. Du, and J.Y. Chen (2015), Linkage between Three Gorges Dam impacts and the dramatic recessions in China's largest freshwater lake, Poyang Lake, *Scientific Reports*, *5*, 18197.

Melillo, J.M., T. Richmond, and G. Yohe (2014), *Climate change impacts in the United States: Third National Climate Assessment*, U.S. Global Change Research Program, Washington, D.C.

Mercier, F., A. Cazenave, and C. Maheu (2002), Interannual lake level fluctuations (1993-1999) in Africa from Topex/Poseidon: connections with ocean-atmosphere interactions over the Indian Ocean, *Global and Planetary Change*, *32*(2-3), 141-163.

Micklin, P. (2007), The Aral Sea disaster, *Annual Review of Earth and Planetary Sciences*, *35*, 47-72.

Miller, W.P., and T.C. Piechota (2011), Trends in Western US Snowpack and Related Upper Colorado River Basin Streamflow, *Journal of the American Water Resources Association*, *47*(6), 1197-1210.
Ministry of Water Resources (2013), *Bulletin of first national census for water*, China Water Power Press, Beijing.
Moore, P., and S.D.P. Williams (2014), Integration of altimetric lake levels and GRACE gravimetry over Africa: Inferences for terrestrial water storage change 2003-2011, *Water Resources Research*, *50*(12), 9696-9720.
Mueller, N., et al. (2016), Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia, *Remote Sensing of Environment*, *174*, 341-352.
Muller, J., M. Naeimi, O. Gitlein, L. Timmen, and H. Denker (2012), A land uplift model in Fennoscandia combining GRACE and absolute gravimetry data, *Physics and Chemistry of the Earth*, *53-54*, 54-60.
Murray, K.E. (2013), State-Scale Perspective on Water Use and Production Associated with Oil and Gas Operations, Oklahoma, U.S., *Environmental Science & Technology*, *47*(9), 4918-4925.
Murti, M., E. Yard, R. Kramer, D. Haselow, M. Mettler, R. McElvany, and C. Martin (2016), Impact of the 2012 extreme drought conditions on private well owners in the United States, a qualitative analysis, *BMC Public Health*, *16*, 1-10.
National Atlas of the United States (2006), Major Dams of the United States, National Atlas of the United States, Reston, VA.
Necsoiu, M., C.L. Dinwiddie, G.R. Walter, A. Larsen, and S.A. Stothoff (2013), Multi-temporal image analysis of historical aerial photographs and recent satellite imagery reveals evolution of water body surface area and polygonal terrain morphology in Kobuk Valley National Park, Alaska, *Environmental Research Letters*, *8*(2).
NOAA (2017), Climate at a Glance: U.S. Time Series, National Centers for Environmental information, Asheville, NC. [Available at https://www.ncdc.noaa.gov/cag/. Accessed August 3, 2017.]
O'Reilly, J.E., S. Maritorena, B.G. Mitchell, D.A. Siegel, K.L. Carder, S.A. Garver, M. Kahru, and C. McClain (1998), Ocean color chlorophyll algorithms for SeaWiFS, *Journal of Geophysical Research-Oceans*, *103*(C11), 24937-24953.
Ogilvie, A., G. Belaud, C. Delenne, J.S. Bailly, J.C. Bader, A. Oleksiak, L. Ferry, and D. Martin (2015), Decadal monitoring of the Niger Inner Delta flood dynamics using MODIS optical data, *Journal of Hydrology*, *523*, 368-383.
Oklahoma Climatological Survey (2016), Oklahoma Statewide Annual Precipitation and Temperature History with 5-year Tendencies 1895-2015, Oklahoma Climatological Survey, Norman, Oklahoma. [Available at http://climate.ok.gov/index.php/climate/climate_trends/precipitation_history_annual_statewide/CD00/prcp/Annual. Accessed May 5, 2016.]
Oklahoma Water Resources Board (2011), Oklahoma comprehensive water plan supplemental report, agricultural issues & recommendations, pp. 1-17, Oklahoma Water Resources Board, Norman, Oklahoma. [Available at http://www.owrb.ok.gov/supply/ocwp/pdf_ocwp/WaterPlanUpdate/draftreports/OCWP_AgriculturalWater_IssuesRecs.pdf. Accessed November 10, 2016.]

OWRB (2015), 2015 Oklahoma Lakes Report Beneficial Use Monitoring Program, pp. 1-216, Oklahoma Water Resources Board. [Available at https://www.owrb.ok.gov/quality/monitoring/bump/pdf_bump/Reports/BUMP%20Lakes%20Report%202015.pdf.]

OWRB and ODWC (2015), *Lakes of Oklahoma*, Oklahoma Water Resources Board and Oklahoma Department of Wildlife Conservation, Oklahoma City, US.

Pekel, J.F., A. Cottam, N. Gorelick, and A.S. Belward (2016), High-resolution mapping of global surface water and its long-term changes, *Nature*, *540*(7633), 418-422.

Proulx, R.A., M.D. Knudson, A. Kirilenko, J.A. VanLooy, and X.D. Zhang (2013), Significance of surface water in the terrestrial water budget: A case study in the Prairie Coteau using GRACE, GLDAS, Landsat, and groundwater well data, *Water Resources Research*, *49*(9), 5756-5764.

Rodell, M., J.S. Famiglietti, D.N. Wiese, J.T. Reager, H.K. Beaudoing, F.W. Landerer, and M.H. Lo (2018), Emerging trends in global freshwater availability, *Nature*, *557*(7707), 650-659.

Rouse Jr, J.W., R. Haas, J. Schell, and D. Deering (1974), Monitoring vegetation systems in the Great Plains with ERTS, *NASA Special Publication*, *351*, 309-317.

Santoro, M., U. Wegmuller, C. Lamarche, S. Bontemps, P. Defoumy, and O. Arino (2015), Strengths and weaknesses of multi-year Envisat ASAR backscatter measurements to map permanent open water bodies at global scale, *Remote Sensing of Environment*, *171*, 185-201.

Scanlon, B.R., C.C. Faunt, L. Longuevergne, R.C. Reedy, W.M. Alley, V.L. McGuire, and P.B. McMahon (2012), Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley, *Proceedings of the National Academy of Sciences of the United States of America*, *109*(24), 9320-9325.

Schaffer-Smith, D., J.J. Swenson, B. Barbaree, and M.E. Reiter (2017), Three decades of Landsat-derived spring surface water dynamics in an agricultural wetland mosaic; Implications for migratory shorebirds, *Remote Sensing of Environment*, *193*, 180-192.

Scheff, J., and D.M.W. Frierson (2012), Robust future precipitation declines in CMIP5 largely reflect the poleward expansion of model subtropical dry zones, *Geophysical Research Letters*, *39*, 18704.

Sellinger, C.E., C.A. Stow, E.C. Lamon, and S.S. Qian (2008), Recent water level declines in the Lake Michigan-Huron system, *Environmental Science & Technology*, *42*(2), 367-373.

Shafer, M., D. Ojima, J.M. Antle, D. Kluck, R.A. McPherson, S. Petersen, B. Scanlon, and K. Sherman (2014), Great Plains, in *Climate Change Impacts in the United States: The Third National Climate Assessment*, edited by J.M. Melillo, T.C. Richmond and G.W. Yohe, pp. 441-461, U.S. Global Change Research Program, United States of America.

Song, C.Q., B. Huang, and L.H. Ke (2013), Modeling and analysis of lake water storage changes on the Tibetan Plateau using multi-mission satellite data, *Remote Sensing of Environment*, *135*, 25-35.

Song, H., R. Xu, Y. Ma, and G. Li (2013), Classification of ETM+ Remote Sensing Image Based on Hybrid Algorithm of Genetic Algorithm and Back Propagation Neural Network, *Mathematical Problems in Engineering*, *2013*, 1-8.

Steele, M.K., et al. (2014), Convergent Surface Water Distributions in US Cities, *Ecosystems*, *17*(4), 685-697.

Strzepek, K., and B. Boehlert (2010), Competition for water for the food system, *Philosophical Transactions of the Royal Society B-Biological Sciences*, *365*(1554), 2927-2940.

Strzepek, K., et al. (2015), Benefits of greenhouse gas mitigation on the supply, management, and use of water resources in the United States, *Climatic Change*, *131*(1), 127-141.

Tao, S.L., J.Y. Fang, X. Zhao, S.Q. Zhao, H.H. Shen, H.F. Hu, Z.Y. Tang, Z.H. Wang, and Q.H. Guo (2015), Rapid loss of lakes on the Mongolian Plateau, *Proceedings of the National Academy of Sciences of the United States of America*, *112*(7), 2281-2286.

Thomas, B.F., J.S. Famiglietti, F.W. Landerer, D.N. Wiese, N.P. Molotch, and D.F. Argus (2017), GRACE Groundwater Drought Index: Evaluation of California Central Valley groundwater drought, *Remote Sensing of Environment*, *198*, 384-392.

Tulbure, M.G., and M. Broich (2013), Spatiotemporal dynamic of surface water bodies using Landsat time-series data from 1999 to 2011, *Isprs Journal of Photogrammetry and Remote Sensing*, *79*, 44-52.

Tulbure, M.G., S. Kininmonth, and M. Broich (2014), Spatiotemporal dynamics of surface water networks across a global biodiversity hotspot-implications for conservation, *Environmental Research Letters*, *9*(11), 114012.

Tulbure, M.G., M. Broich, S.V. Stehman, and A. Kommareddy (2016), Surface water extent dynamics from three decades of seasonally continuous Landsat time series at subcontinental scale in a semi-arid region, *Remote Sensing of Environment*, *178*, 142-157.

U.S. Census Bureau (2010), State Area Measurements and Internal Point Coordinates, U.S. Census Bureau, Suitland, Maryland. [Available at http://www.census.gov/geo/reference/state-area.html. Accessed May 24, 2016.]

US Geological Survey (2018), *Landsat Collection 1 Level-1 Quality Assessment Band*, Department of the Interior U.S. Geological Survey, Washington, D.C.

USACE (2017), Tulsa District, US Army Corps of Engineers Hydropower, US Army Corps of Engineers, Washington, D.C. [Available at http://www.swt.usace.army.mil/Missions/Hydropower/. Accessed March 20, 2017.]

USDA-NASS (2014), 2012 Census of Agriculture, pp. 340-358, United States Department of Agriculture and National Agricultural Statistics Service, Washington, DC. [Available at https://www.agcensus.usda.gov/Publications/2012/. Accessed December 20, 2016.]

USGS (2010), Estimated Use of Water in the United States in 1985, 1990, 1995, 2000, 2005 and 2010, U.S. Geological Survey, Reston, Virginia. [Available at http://water.usgs.gov/watuse/data/. Accessed September 10, 2016.]

USGS (2012), Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) Algorithm Description Document, Department of the Interior U.S. Geological Survey, Washington, D.C. [Available at https://landsat.usgs.gov/sites/default/files/documents/ledaps_add.pdf. Accessed December 10, 2015.]

USGS (2017a), Landsat 8 Surface Reflectance Code (LaSRC) Product Guide-Version 3.4, Department of the Interior U.S. Geological Survey, Washington, D.C. [Available at https://landsat.usgs.gov/sites/default/files/documents/lasrc_product_guide.pdf. Accessed July 15, 2017.]

USGS (2017b), Landsat 4-7 Climate Data Record (CDR) Surface Reflectance Product Guide-Version 7.2, Department of the Interior U.S. Geological Survey, Washington, D.C. [Available at https://landsat.usgs.gov/sites/default/files/documents/ledaps_product_guide.pdf. Accessed July 10, 2017.]

van As, D., et al. (2018), Reconstructing Greenland Ice Sheet meltwater discharge through the Watson River (1949-2017), *Arctic, Antarctic, and Alpine Research*, *50*(1), S100010.

van Dijk, A.I.J.M., H.E. Beck, R.S. Crosbie, R.A.M. de Jeu, Y.Y. Liu, G.M. Podger, B. Timbal, and N.R. Viney (2013), The Millennium Drought in southeast Australia (2001-2009): Natural and human causes and implications for water resources, ecosystems, economy, and society, *Water Resources Research*, *49*(2), 1040-1057.

van Vliet, M.T.H., J.R. Yearsley, F. Ludwig, S. Vogele, D.P. Lettenmaier, and P. Kabat (2012), Vulnerability of US and European electricity supply to climate change, *Nature Climate Change*, *2*(9), 676-681.

Vaughn, C.C., C.L. Atkinson, and J.P. Julian (2015), Drought-induced changes in flow regimes lead to long-term losses in mussel-provided ecosystem services, *Ecology and Evolution*, *5*(6), 1291-1305.

Verpoorter, C., T. Kutser, and L. Tranvik (2012), Automated mapping of water bodies using Landsat multispectral data, *Limnology and Oceanography-Methods*, *10*, 1037-1050.

Verpoorter, C., T. Kutser, D.A. Seekell, and L.J. Tranvik (2014), A global inventory of lakes based on high-resolution satellite imagery, *Geophysical Research Letters*, *41*(18), 6396-6402.

Vogelmann, J.E., S.M. Howard, L. Yang, C.R. Larson, B.K. Wylie, and N. Van Driel (2001), Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources, *Photogrammetric Engineering and Remote Sensing*, *67*(6), 650-662.

Vorosmarty, C.J., et al. (2010), Global threats to human water security and river biodiversity, *Nature*, *467*(7315), 555-561.

Walls, S.C., W.J. Barichivich, M.E. Brown, D.E. Scott, and B.R. Hossack (2013), Influence of Drought on Salamander Occupancy of Isolated Wetlands on the Southeastern Coastal Plain of the United States, *Wetlands*, *33*(2), 345-354.

Wan, Z., S. Hook, and G. Hulley (2015), MOD11C3 MODIS/Terra Land Surface Temperature/Emissivity Monthly L3 Global 0.05 Deg CMG V006, edited, NASA EOSDIS LP DAAC, Sioux Falls, SD.

Wang, H.S., L.L. Jia, H. Steffen, P. Wu, L.M. Jiang, H.T. Hsu, L.W. Xiang, Z.Y. Wang, and B. Hu (2013), Increased water storage in North America and Scandinavia from GRACE gravity data, *Nature Geoscience*, *6*(1), 38-42.

Wang, S.D., M.H.A. Baig, L.F. Zhang, H.L. Jiang, Y.H. Ji, H.Q. Zhao, and J.G. Tian (2015), A Simple Enhanced Water Index (EWI) for Percent Surface Water Estimation Using Landsat Data, *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(1), 90-97.

Wang, S.Y.S., W.R. Huang, H.H. Hsu, and R.R. Gillies (2015), Role of the strengthened El Nino teleconnection in the May 2015 floods over the southern Great Plains, *Geophysical Research Letters*, *42*(19), 8140-8146.

Wang, S.Y.S., Y.H. Lin, R.R. Gillies, and K. Hakala (2016), Indications for Protracted Groundwater Depletion after Drought over the Central Valley of California, *Journal of Hydrometeorology*, *17*(3), 947-955.

Wang, W., H.I. Liu, Y.Q. Li, and J.L. Su (2014), Development and management of land reclamation in China, *Ocean & Coastal Management*, *102*, 415-425.

Watkins, M.M., D.N. Wiese, D.N. Yuan, C. Boening, and F.W. Landerer (2015), Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons, *Journal of Geophysical Research: Solid Earth*, *120*(4), 2648-2671.

Wiese, D.N., D.-N. Yuan, C. Boening, F.W. Landerer, and M.M. Watkins (2018), JPL GRACE Mascon Ocean, Ice, and Hydrology Equivalent Water Height Release 06 Coastal Resolution Improvement (CRI) Filtered Version 1.0. Ver. 1.0., edited by P. DAAC, PO. DAAC, CA, USA.

Wolf, S., et al. (2016), Warm spring reduced carbon cycle impact of the 2012 US summer drought, *Proceedings of the National Academy of Sciences of the United States of America*, *113*(21), 5880-5885.

Wood, E.F., et al. (2011), Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, *Water Resources Research*, *47*, 1-10.

Woods, A.J., J.M. Omernik, D.R. Butler, J.G. Ford, J.E. Henley, B.W. Hoagland, D.S. Arndt, and B.C. Moran (2005), Ecoregions of Oklahoma, *map scale 1:1,250,000*, U.S. Geological Survey, Reston, Virginia.

Wulder, M.A., J.C. White, T.R. Loveland, C.E. Woodcock, A.S. Belward, W.B. Cohen, E.A. Fosnight, J. Shaw, J.G. Masek, and D.P. Roy (2016), The global Landsat archive: Status, consolidation, and direction, *Remote Sensing of Environment*, *185*, 271-283.

Xiao, M., A. Koppa, Z. Mekonnen, B.R. Pagan, S.A. Zhan, Q.A. Cao, A. Aierken, H. Lee, and D.P. Lettenmaier (2017), How much groundwater did California's Central Valley lose during the 2012-2016 drought?, *Geophysical Research Letters*, *44*(10), 4872-4879.

Xiao, X., S. Boles, S. Frolking, W. Salas, B. Moore, C. Li, L. He, and R. Zhao (2002), Observation of flooding and rice transplanting of paddy rice fields at the site to

landscape scales in China using VEGETATION sensor data, *International Journal of Remote Sensing*, *23*(15), 3009-3022.
Xiao, X.M., S. Boles, S. Frolking, C.S. Li, J.Y. Babu, W. Salas, and B. Moore (2006), Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images, *Remote Sensing of Environment*, *100*(1), 95-113.
Xiao, X.M., S. Boles, J.Y. Liu, D.F. Zhuang, S. Frolking, C.S. Li, W. Salas, and B. Moore (2005), Mapping paddy rice agriculture in southern China using multi-temporal MODIS images, *Remote Sensing of Environment*, *95*(4), 480-492.
Xu, H.Q. (2006), Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery, *International Journal of Remote Sensing*, *27*(14), 3025-3033.
Yamazaki, D., M.A. Trigg, and D. Ikeshima (2015), Development of a global similar to 90 m water body map using multi-temporal Landsat images, *Remote Sensing of Environment*, *171*, 337-351.
Yang, X.K., and X.X. Lu (2014), Drastic change in China's lakes and reservoirs over the past decades, *Scientific Reports*, *4*, 6041.
Yoshimura, C., T. Omura, H. Furumai, and K. Tockner (2005), Present state of rivers and streams in Japan, *River Research and Applications*, *21*(2-3), 93-112.
Zhang, X.H., N. Wang, Z.Y. Xie, X.L. Ma, and A. Huete (2018), Water Loss Due to Increasing Planted Vegetation over the Badain Jaran Desert, China, *Remote Sensing*, *10*(1), 134.
Zhang, Y.W., J.T. Liu, L.H. Wan, and S.Q. Qi (2015), Land Cover/Use Classification Based on Feature Selection, *Journal of Coastal Research*, *73*, 380-385.
Zhao, C., Z. Zhu, and D. Zhou (2000), *Worldwide rivers and dams*, China Water Conservancy and Hydroelectric Press, Beijing.
Zhao, G.J., X.M. Mu, Z.M. Wen, F. Wang, and P. Gao (2013), Soil Erosion, Conservation, and Eco-Environment Changes in the Loess Plateau of China, *Land Degradation & Development*, *24*(5), 499-510.
Zou, Z., J. Dong, M.A. Menarguez, X. Xiao, Y. Qin, R.B. Doughty, K.V. Hooker, and K. David Hambright (2017), Continued decrease of open surface water body area in Oklahoma during 1984-2015, *Science of the Total Environment*, *595*, 451-460.
Zou, Z., X. Xiao, J. Dong, Y. Qin, R.B. Doughty, M.A. Menarguez, G. Zhang, and J. Wang (2018), Divergent trends of open surface water body area in the contiguous US from 1984-2016, *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 3810-3815.