Mollie D. Gaines[1*], Mirela G. Tulbure[1,2], Vinicius Perin[1]

[1]Center for Geospatial Analytics, North Carolina State University, 2800 Faucette Drive, Raleigh, NC 27695, USA.

[2]Department of Forestry and Environmental Resources, North Carolina State University, 2820 Faucette Drive, Raleigh, NC 27695, USA.

*Corresponding author: Mollie Gaines (mdgaines@ncsu.edu)

Key Points:

- We developed top-down data-driven models to estimate percent surface water area based on Landsat imagery and climate and human drivers.
- Compared to estimates based on climate and human drivers independently, combining them reduces percent surface water estimation error.
- Natural land cover was the most influential explanatory variable to estimate percent surface water area of our climate and human variables.

Abstract

Surface water is the most readily accessible water resource and provides an array of ecosystem services, but is stressed by changes in climate, land cover, and population size. Understanding drivers of surface water dynamics in space and time is key to better managing our water resources. However, few studies estimating changes in surface water account for climate and anthropogenic drivers both independently and together. We used 19 years (2000-2018) of the newly developed Dynamic Surface Water Extent Landsat Science Product in concert with time series of precipitation, temperature, land cover, and population size to statistically model maximum seasonal percent surface water area as a function of climate and anthropogenic drivers in the Southeastern U.S. We fitted three statistical models (linear mixed effects, random forests, and mixed effects random forests) and three groups of explanatory variables (climate, anthropogenic, and their combination) to assess the accuracy of estimating percent surface water area at the watershed scale with different drivers. We found that anthropogenic drivers accounted for approximately 37% more of the variance in the percent surface water area than the climate variables. The combination of variables in the mixed effects random forest model produced the smallest mean percent errors (mean -0.17%) and the highest explained variance (R^2 0.99). Our results indicate that anthropogenic drivers have greater influence when estimating percent surface water area than climate drivers, suggesting that water management practices and land use policies can be highly effective tools in controlling surface water variations in the Southeastern U.S.

Plain Language Summary

People and the environment rely on water to exist and thrive, especially water on the Earth's surface because that is the easiest place to get it. The amount of surface water and where it is located is changing with the climate and changes in

people's water use, and our need for it is increasing. In order to plan ahead for future water needs, we need to better understand how the climate and people are changing surface water patterns both separately and together. To help improve our understanding of these changes, we modeled the amount of surface water three different ways. First, modeled based on climate data (like temperature and precipitation); second, based on human data (like land use and population); and third, based on both climate and human data together. We found that we could best model the amount surface water if we used both climate and human data together, and that human data can explain a lot of the changes in the amount of surface water. These results mean that we can work to control changes in the amount of surface water by controlling human actions through planning and policies.

1 Introduction

Water is one of, if not the, most valuable resources in the world. The inland distribution of water is naturally dynamic over time and space with changes in climate and land use and land cover (LULC) influencing these dynamics (Palmer et al., 2008; Tulbure & Broich, 2019; Vörösmarty et al., 2010). These changes can increase water stress and heighten tensions in already strained relationships, such as those between rural, agriculture-based regions and urban centers (Flörke et al., 2018). In some areas, water scarcity and intense irrigation can and have even led to armed conflicts and civil wars, the civil war in Syria being a recent and powerful example (Iceland, 2017; Müller et al., 2016). In the Southeastern United States (U.S.), the increased water demand of the Atlanta metropolitan area has led to legal disputes between Georgia, Alabama, and Florida (Jordan, 2001).

Climate change is altering the patterns of rainfall across the globe (Dai, 2013), and warming is increasing the growing length of the agricultural growing season (Kunkel et al., 2004). Globally, the largest anthropogenic use of land is agriculture (Foley et al., 2005) and it has long been recognized as a major driver of environmental change (Lark et al., 2017; Turner et al., 2007). There is potential for longer growing seasons to increase water stress, particularly when agricultural and urban areas are competing for the same water resources (Flörke et al., 2018). For example, in 2015, the majority of total water resources (surface and groundwater) across the U.S. were used for crop irrigation (Dieter et al., 2018). While climate models project an increase in precipitation in the Southeast, they also predict a greater increase in evaporation leading to a net decrease in water resources in the region (Duan et al., 2017; Ferguson et al., 2018). Additionally, climate models project increases in extreme rainfall events (Carter et al., 2018; Keellings & Engström, 2019) and in the precipitation from and intensity of tropical cyclones (Karl et al., 2009; Kossin et al., 2017). These storms paired with increases in impervious surface area in the Southeastern U.S. can lead to major flooding events (Carter et al., 2018).

Rapid economic development, urban growth, and population growth increase water demand (Jeong et al., 2015; Piao et al., 2010; Vörösmarty et al., 2010;

Wada et al., 2017), and surface water resources are the most easily accessible water source for humans to utilize, both in urban and agricultural areas (Human Appropriation of Renewable Fresh Water, 1996). Surface water resources are essential to a region's economic and urban development (Veldkamp et al., 2017), and they are impacted by both climate and anthropogenic drivers (Vörösmarty et al., 2000; Zeng et al., 2020). Understanding the spatiotemporal patterns of surface water dynamics is the first step in addressing water scarcity issues and important for developing inter- and intra-state water management policies to provide local and regional resiliency (Engström et al., 2021).

Two common approaches for quantifying surface water dynamics have been hydrological models or data-driven models, with the former increasingly becoming the focus of regional and large-scale hydrological models (Sood & Smakhtin, 2015; Wada et al., 2017); however, incorporating human activities into such models continues to be a major challenge (Vörösmarty et al., 2010; Wada et al., 2017). Large scale process-based, hydrological models have grown in complexity and sophistication, but they still need substantial improvement in simulating anthropogenic interactions with the environment and their impacts on water systems (Clark et al., 2017; Nazemi & Wheater, 2015; Pokhrel et al., 2016; Wada et al., 2017). Many of these models do not account for human drivers such as population density and LULC change or intensity (Hostetler & Alder, 2016; Thrasher et al., 2013; Wada et al., 2017). Using data-driven approaches, we are beginning to quantify the spatiotemporal distributions of surface water dynamics with climate and anthropogenic drivers independently and synergistically (L. Li et al., 2019; Tulbure & Broich, 2019; Xu et al., 2019).

Recently, hydrological analyses of large areas have been trending toward more data-driven empirical approaches, because satellite imagery is the only way to assess water systematically over large spatial and temporal scales (Palazzoli & Ceola, 2020; Pekel et al., 2016; Perin et al., 2021; Tulbure & Broich, 2019; Wada et al., 2017; Walker et al., 2020). Surface-water-specific datasets derived from moderate resolution (30m, Landsat) satellite imagery over 30-40 years are a relatively new development at the regional (Tulbure et al., 2016; Tulbure & Broich, 2013), national (Jones, 2015, 2019), and global scales (Pekel et al., 2016; Pickens et al., 2020). The spatial and temporal scale of these surface water datasets, and the similar spatial and temporal scales of LULC (the Cropland Data Layer, CDL; "CropScape - NASS CDL Program"), population (LandScan; Rose et al., 2020), and climate data (Gridded Surface Meteorological dataset; Abatzoglou, 2013), enable us to tackle the critical task of assessing the impact of climate and anthropogenic drivers on surface water. Previous research in a dryland region indicated that the combination of human and climate drivers is more impactful than climate drivers alone on estimating surface water change (Tulbure & Broich, 2019). Another study showed that urban areas growth correlate with surface water loss across the globe (Palazzoli & Ceola, 2020).

Because of rapid LULC changes and climate variability, studying the spatial and temporal dynamics of surface water and its drivers can provide new insights into

preparation and responses to droughts and floods. To assess the combined effects of climate and anthropogenic drivers of surface water extent over decades at a regional scale, in this study, we used three statistical models—linear mixed effect models (LMMs), random forest regression (RF) models, and mixed effect random forest (MERF) models to estimate seasonal percent surface water area from 2000-2018 using three different sets of explanatory variables—climate, anthropogenic, and combination. The specific objectives of this study are to: (1) compare the influence of climate drivers, anthropogenic drivers, and their combination on estimating seasonal percent surface water area in the Southeastern U.S., a region that has experienced severe water stress and flooding, across three statistical models and assess model performances; and (2) determine which variables are most important in estimating percent surface water area across all models.

2 Data and Methods

2.1 Study Area

The Southeastern U.S. outlined by Environmental Protection Agency (EPA) Region 4 encompasses eight states: Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, and Tennessee (upper left map, Fig. 1). The region is comprised of 14 level III Ecoregions, with the Southeastern Plains accounting for the majority of the land cover (U.S. Environmental Protection Agency, 2004). It includes 310 8-digit hydrologic unit code watersheds (hereafter, "HUCs" or "watersheds"), which we used as our unit of aggregation (right map, Fig. 1), across four water resource regions. The South Atlantic Gulf and Tennessee water resource regions account for approximately 64% (199) and 10% (32) of our total 310 HUCs, respectively (lower left map, Fig. 1). The HUCs of three of the four water resource regions are not entirely contained within our study region (Tennessee: 32/32, South Atlantic Gulf: 199/204, Lower Mississippi 34/82, Ohio 45/120; lower left map, Fig. 1). The entire study area is approximately 1,087,370 km^2, the average watershed is approximately 3,500 km^2, the largest is 11,790 km^2, and the smallest is 180 km^2.

The Southeastern U.S., our study area for this work, has experienced the most LULC change of any region in the country (Homer et al., 2020; Sleeter et al., 2018), and much of this change is from agriculture or forest to urban area as southern cities experience rapid growth (Sanchez et al., 2020; Sleeter et al., 2018; Terando et al., 2014). It has also experienced the largest population growth rate of any region in the U.S. (Terando et al., 2014). This trend is expected to continue with an estimated 24% increase over the next 20 years, 4% higher than the national average (McManamay et al., 2019; University of Virginia Weldon Cooper Center, 2018). Additionally, it is heavily car-dependent, lending itself to low-density, sprawling urban growth and increasing the amount of impervious surface area (McManamay et al., 2019; Terando et al., 2014). These factors can exacerbate the effects of climate change by increasing the urban heat island effect and leading to more severe flooding (McManamay et al., 2019; Walsh et al., 2005). Many of the larger urban areas in the region,

i.e. Atlanta and Knoxville, get their public water supply mainly from surface water sources (McManamay et al., 2019). Because urban areas tend to draw on resources from their surrounding areas, urban water stress can lead to regional and even inter-state legal conflicts, as has happened when Atlanta experienced water shortages (McManamay et al., 2019; Missimer et al., 2014). The region is also expected to see an increase in the intensity of hurricanes in the future and a related increase in natural disaster associated vulnerabilities (Allen et al., 2016; Emanuel et al., 2013; Pielke et al., 2008).

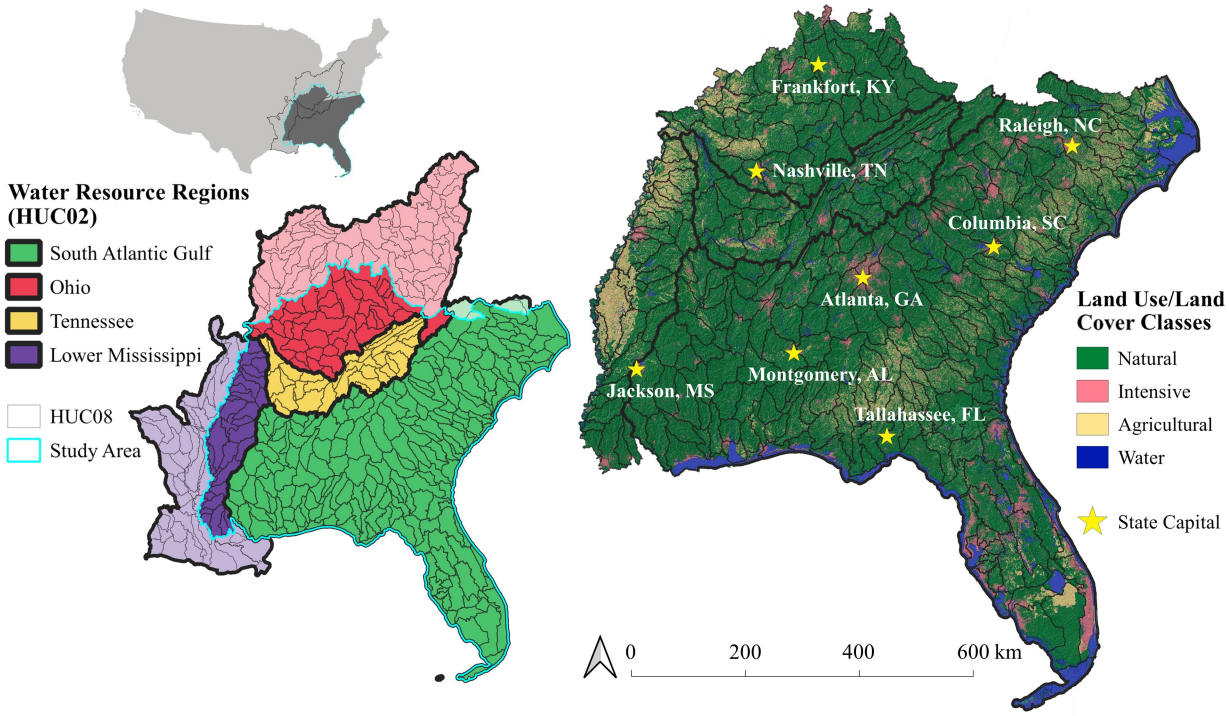

Figure 1. Our study area (outlined in cyan) is comprised of the 310 HUCs covering the Southeastern U.S. EPA Region 4, shown in dark grey in the upper left map of the U.S. The HUC 02 water resource regions are outlined in black in all three maps.

The Southeastern U.S. is currently a humid subtropical region (Alnahit et al., 2020; Hernandez-Ochoa & Asseng, 2018; Ingram et al., 2012) dominated by temperate forests including the longleaf and loblolly-shortleaf pine ecosystems (Foster et al., 2019; Matusick et al., 2020). There is a latitudinal gradient in temperature (Alnahit et al., 2020), with the average maximum temperatures increasing with a decrease in latitude and ranging from 17.31 to 29.60 °C. Over the next 30 years, summer apparent temperatures—an approximation of a person's experience of temperature based on temperature, wind speed, and relative humidity (Steadman, 1984)—across the region are expected to increase by 2.4 to

4.1 °C, mirroring the current, tropical climate, summer apparent temperatures of southern Florida (Diem et al., 2017). From 1985 to 2019, the average daily maximum and minimum temperatures were 23.34 and 10.90 °C, respectively, at the HUC scale (Table 1). The average annual precipitation at the watershed scale over the 34-year time period is 1345.39 mm (Table 1) and fluctuates with the phases of the El Niño Southern Oscillation (Mourtzinis et al., 2016).

On average, at the watershed scale throughout our time period (2000-2018) natural land cover was the dominant land cover type (74.52%; Table 1) of the three land cover classes (natural, agricultural, and intensive) we defined using the National Land Cover Dataset (NLCD; Homer et al., 2015; Yang et al., 2018) and the Cropland Data Layer ("CropScape - NASS CDL Program"). Each watershed across the region and over all 19 years contained developed land, while not every instance of a watershed had agricultural land cover (Table 1). However, agricultural land cover was the second largest land cover class on average (9.78%; Table 1). Overall, the average percent of natural land cover in a watershed is decreasing, with 60% of HUCs having lost more than 1% of their natural land cover between 2000 and 2018 and 14 having lost more than 5%. The average percentages of intensive and agricultural land cover in a watershed are both increasing, with a more than 1% increase of intensive land cover across approximately 26% of HUCs and of agricultural land for approximately 34% of HUCs between 2000 and 2018.

Table 1. Distribution of Explanatory Variables across the 310 8-digit HUCs (watersheds)

@ >p(- 14) * >p(- 14) * >p(- 14) * >p(- 14) * >p(- 14) * >p(- 14) * >p(- 14) * >p(- 14) * @

Variable

& Mean & Standard deviation & Minimum & 25th Percentile & 50th Percentile & 75th Percentile & Maximum

Maximum Temp (°C)

&

23.34

&

2.73

&

17.31

&

21.25

&
23.46
&
25.32
&
29.60

Minimum Temp (°C)
&
10.90
&
3.00
&
5.02
&
8.69
&
10.76
&
12.70
&
20.88

Precipitation (mm)
& 1345.39 &
139.12
&
1076.45
&

1230.17

&

1342.35

&

1434.43

&

1731.50

Natural Land Cover (%)

&

74.52

&

18.63

&

2.71

&

67.20

&

80.23

&

88.39

&

95.48

Agriculture Land Cover (%)

&

9.78

&

13.46

&

0.00

&

0.76

&

3.96

&

13.86

&

80.17

Intensive Land Cover (%)

&

8.97

&

6.41

&

0.94

&

5.22

&

7.03

&

10.66

&

61.45

Pop. Density (People/km^2)

&

55.52

&

78.22

&

0.5647

&

17.15

&

29.60

&

62.55

&

677.96

2.2 Surface Water Data

We used the USGS Earth Resources Observation Center Landsat Level 3 Dynamic Surface Water Extent (DSWE) Science Product (U.S. Geological Survey, 2019b) to calculate seasonal surface water area for each year in each HUC across our study area and time period (2000-2018). DSWE is a relatively new, high temporal (approximately 8 days) and moderate spatial resolution (30 m), long-term (1984–present) terrestrial surface water inundation dataset derived from the U.S. Landsat Analysis Ready Data (ARD) Surface Reflectance product in the Albers Equal Area projection (Jones, 2015, 2019; U.S. Geological Survey, 2019a). Landsat ARD is comprised of the most geometrically accurate data from Landsat 4-5 TM, Landsat 7 ETM+, and Landsat 8 OLI (U.S. Geological Survey, 2019a). DSWE has been freely accessible since 2019 on the USGS Earth Explorer data portal (https://earthexplorer.usgs.gov/; U.S. Geological Survey, 2019b). The DSWE product was validated to ensure the DSWE algorithm's accuracy in detecting partial surface water and inundation in vegetated wetlands (Jones, 2015, 2019). Jones (2015) detected no temporal trend in the DSWE algorithm performance and no bias as a function of hydrologic conditions, supporting the use of DSWE for long-term surface water inundation monitoring for trend analyses over time. The overall agreement rate between DSWE classifications and gage inundation states was 77% across over 3,500 observations (Jones, 2019). While the DSWE dataset was developed using Landsat imagery from across the U.S., the DSWE model (an inundation algorithm) inputs were designed to be applied globally (i.e. not requiring scene-based training data;

Jones, 2015, 2019). For this study we used the interpreted layer with mask applied (INWM; U.S. Geological Survey, 2019b) of the DSWE product. The INWM layer classifies pixels into seven groups assigned via pixel value: 0-not water; 1-water – high confidence; 2-water – moderate confidence; 3-potential wetland; 4-water or wetland low confidence; 9-cloud, cloud shadow, or snow; 255-not available, fill. These categories are based on the results of five tests of water and vegetation indices (Jones, 2019; U.S. Geological Survey, 2019a). The high confidence water classification has been independently shown to have >80% overall accuracy assessed across a set of randomly sampled and manually interpreted pixels over multiple years (Soulard et al., 2020).

We used a pixel-based analysis to calculate the surface water area for each HUC in the region for each season (beginning March 1, June 1, September 1, and December 1 with winter including January and February from the following calendar year) from 2000-2018. To aggregate the DSWE INWM layer by season, we stacked each ARD raster tile for a season within a year, of which there was an average of 20. For each season, we calculated each pixel's frequency of "water–high confidence", any classification of water (values 1–4), "not water", "cloud, cloud shadow, and snow", and the percent of high confidence water out of all times a value other than cloud or no data was recorded. We recorded each of these seasonal frequency counts and the percentage of high confidence water as bands in an output raster.

We then used the zonal_stats function from the rasterstats module (Perry, 2020) in Python (Python Software Foundation, https://www.python.org) to calculate, within each HUC, the number of pixels from our aggregated DSWE where the percent of high confidence water values was greater than or equal to 25%. In other words, we counted the number of pixels within each HUC that had been classified as high confidence water in the DSWE algorithm (Jones, 2019) at least 25% of the time they were able to be classified (i.e., not covered by clouds) during their respective year and season. Because the number of viable pixel values at the pixel level varies even more than the number of Landsat scenes in a tile per season (mean 19.9, standard deviation 6.7), we set our threshold as a percentage rather than a specific number of pixels classified as high confidence water. We chose 25% in order to reduce some of the artifacts from cloud shadows that are present in the DSWE high confidence water class, while maximizing the extent of the surface water detected (i.e., capturing flood events) in the DSWE dataset. We completed these summations by concurrently calculating for each ARD tile the number of pixels within each HUC that met our threshold and then summing together each of these tile-specific HUC values. To calculate the percentage of each HUC that was covered in surface water, we converted the pixel count to square kilometers (multiplied by 0.0009) and divided the resulting surface water area by the total area (km^2) of the HUC.

2.3 Climate Data

To account for potential climate drivers of surface water, we calculated the standardized seasonal anomalies of three climate related variables—maximum

temperature, minimum temperature, and precipitation—for each season in each year of our timescale (2000-2018). We obtained and processed these data in Google Earth Engine (GEE) where the Gridded Surface Meteorological (GRIDMET) dataset from the University of Idaho (Abatzoglou, 2013), a long-term (1979 to present) daily time series of moderate (4 km) spatial resolution climate data, was freely available. The dataset was validated with a direct comparison to station observations from four weather station networks (Abatzoglou, 2013). For the contiguous U.S., minimum and maximum daily gridded temperature both had median correlations of 0.97 and daily gridded precipitation had a median correlation of 0.74 (Abatzoglou, 2013). When compared to five other gridded weather datasets, Blankenau et al. (2020) found that GRIDMET had the smallest median station bias (+0.54 °C), the highest median station correlation (0.87), and the smallest variability error for near-surface air temperature among the long-term datasets (beginning before 2015).

We used the GRIDMET dataset to calculate pixel-based standardized seasonal anomalies (Eq. 1) in GEE for each of our climate variables. For each season (i) per year (j), we calculated the standardized seasonal anomaly ($\mathrm{SSA_{ij}}$). We first subtracted the average daily climate value of the season and year (x_{ij}; e.g. precipitation for the spring of 2018) from the seasonal long-term (spring 1985 through winter 2018) average ($\overline{x}_i$; e.g. average spring precipitation). We used the same three-month seasons defined for processing the DSWE data, and then divided the difference by the seasonal long-term (1985-2018) standard deviation (σ_i; e.g. standard deviation of spring precipitations).

$$\mathrm{SSA_{ij}} = \frac{x_{\mathrm{ij}} - \overline{x}_i}{\sigma_i} \quad (1)$$

Lastly, we summarized these pixel-based anomalies in GEE by calculating their average across all pixels within each HUC to get the average HUC-level standardized seasonal anomaly for each season and year. The GEE function we applied (image.ReduceRegions() with ee.Reducer.mean()) includes pixels on the edge of the polygon if at least 0.5% of the pixel is within the polygon (Google Earth Engine, 2021). When calculating the mean, these edge pixel values are weighted based on the fraction of the pixel that is within the polygon (Google Earth Engine, 2021).

2.4 Anthropogenic Data

As proxies for anthropogenic drivers of surface water change from 2000-2018, we used LULC data derived from the U.S. Department of Agriculture National Agricultural Statistics Services (NASS) Cropland Data Layer ("CropScape - NASS CDL Program") and population density data from LandScan produced by Oak Ridge National Laboratory (Rose et al., 2020). Because the CDL was not available at the conterminous U.S. level until 2008 (Lark et al., 2017), we used the 2001 and 2006 components of the NLCD2016 database (Homer et al., 2020) to calculate our LULC variables from 2000 to 2007. Unlike the climate and DSWE data, the anthropogenic data have a yearly scale.

The CDL is available annually at a 30 m spatial resolution with the expressed purpose of quantifying crop types around the middle of the year (Lark et al., 2017). It uses NLCD data as part of its training data, specifically for the non-crop land cover types (Lark et al., 2017). The crop classification accuracies for major crops (i.e., corn, cotton, rice, soy beans, and wheat) are generally over 90%, and the accuracy of the layers have increased over time (Lark et al., 2017). The CDL has a slight negative bias (i.e., underestimating cropland), but this bias has been decreasing since 2008 (Lark et al., 2017). To improve the accuracies of both crop and non-crop classifications, Lark et al. (2017) recommended consolidating classifications, observing an overall accuracy of 91.8% when conducting a direct (pixel-level) change assessment on consolidated CDL classes.

The NLCD has a 30 m spatial resolution and a temporal resolution of approximately 5 years with the purpose of recording long-term land cover at the conterminous U.S. scale (Yang et al., 2018). The NLCD2016 database had a Level II (16 land cover classes) overall accuracy of 83.7% for 2001 and 83.6% for 2006 (Wickham et al., 2021). The standard error for both years was 0.5. For Level I (8 land cover classes), the NLCD2016 database had an overall accuracy of 89.2% and a standard error of 0.5 for both 2001 and 2006 (Wickham et al., 2021). The Level I land cover classes are combinations of Level II classes and the improved overall accuracy supports Lark et al. (2017)'s recommendation of consolidating land cover classifications to improve accuracy. Class consolidation is also recommended for when comparing to or using in combination with other datasets (Lark et al., 2015, 2017).

The three classes to which we consolidated the CDL and NLCD land cover classes were agricultural, natural, and intensive land use. To combine the land cover classes for 2008-2018, we ran a reclassification function on the CDL in GEE at the pixel-level for each year, based on the CDL crop and non-crop class memberships outlined in Lark et al. (2017; Table S1). For 2000 through 2007, we similarly reclassified the 2001 and 2006 NLCD data into our three categories (Table S2). For each year, we calculated the proportion of each land cover class within each HUC by dividing the area of the land cover classified pixels by the total area in the HUC. We then calculated a weighted average of the 2001 and 2006 NLCD and the 2008 CDL proportions to span 2000-2007 (Table S3). Because these data are at a yearly temporal scale, each season within a year was assigned the same proportions of land cover classes.

The LandScan dataset, derived from a dasymetric model based on satellite observations (Allen et al., 2016; Bhaduri et al., 2007), provides annual ambient population data from 2000 to 2019 (with the next year's data published annually) at approximately 1km resolution (Rose et al., 2020). An early version of LandScan used U.S. census counts to validate the model, finding that approximately 87% of the LandScan population corresponded with county census data in the Southwestern U.S. (Dobson et al., 2000). The spatial accuracy of LandScan has improved since the 2000 dataset as technological advances (i.e., light

detection and ranging—LiDAR—building footprints) have improved the quality, accuracy, and validity of its model inputs (McKee et al., 2015). We started our models at the first year of population data availability (2000), rather than the first year of DSWE data, because we wanted to directly compare models using different groups of explanatory variables.

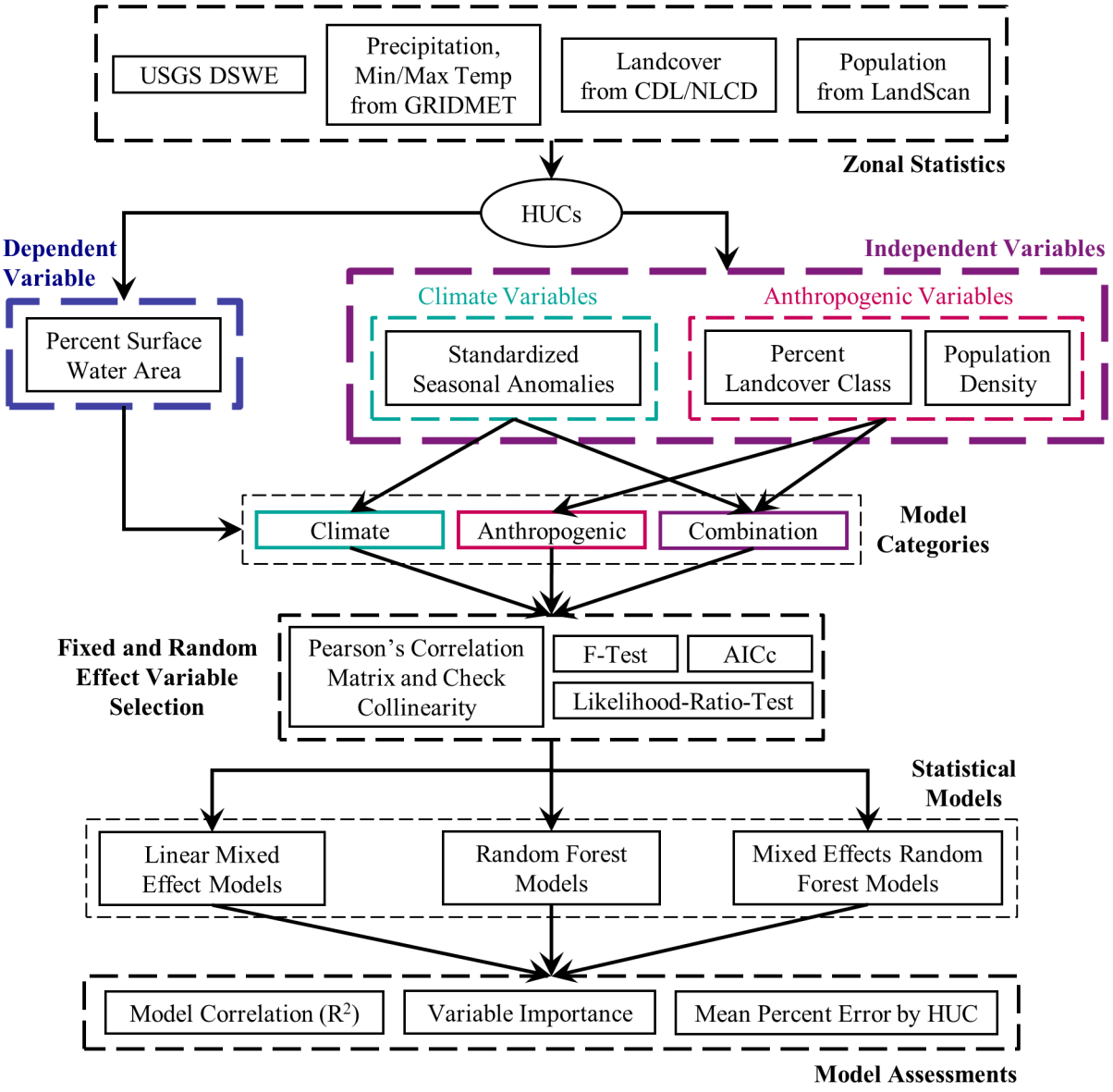

Figure 2. The overall methods workflow to estimate seasonal percent surface water area from 2000-2018 across the HUCs in the study area using a total of 9

models: three statistical models (LMM, RF, and MERF) each with three sets of explanatory variables (climate, anthropogenic, and combination).

2.5 Model Descriptions

We used three different statistical models to estimate the seasonal percent of surface water area across the 310 HUCs in the Southeast. To meet the assumptions of these models—all of which are built on linear regression—and to ensure ease of interpretation, we first centered and standardized the independent variables (Harrison et al., 2018; Hox et al., 2010). Additionally, to satisfy the assumption of the normality of the dependent variable (i.e., percent surface water) we log-transformed the percent surface water after adding a small constant (10^{-6}) to all observations in order to preserve the very rare instance ($< 0.02\%$) of a HUC surface water percentage of 0.

To compare the importance of climate and anthropogenic drivers in estimating surface water, we ran each of the three statistical models—LMM, RF, and MERF—with different sets of independent variables (Fig. 2). For the climate models, we used maximum temperature, minimum temperature, and precipitation anomaly variables. For the anthropogenic models, we used the proportions of agricultural, natural, and intensive land cover as well as population density. For the combination models, which accounted for both climate and anthropogenic drivers, we used all of the variables in the climate and anthropogenic models.

2.5.1 Linear Mixed Effect Models

Linear mixed effect models (Eq. 2) are a powerful extension of linear regression that can control for different types of clustering within the data by modeling them as random effects, also known as grouping factors (Harrison et al., 2018; Hox et al., 2010; Schielzeth & Nakagawa, 2012). These grouping factors help explain randomness in the variability of the response variable (Harrison et al., 2018). Possible grouping factors we considered were HUCs, seasons, and years, as well as years nested within seasons and seasons nested within years. Possible fixed effects, or explanatory variables, we considered the climate variables used in the climate and combination models and the anthropogenic variables used in the anthropogenic and combination models (Table 1).

To determine what fixed and random effects for which we should control, we ran a backward stepwise selection process beginning with all considered fixed and random effects (Fig. 2). We found that the HUCs were the most dominant grouping factor, and second were the nested random effects of the years within the seasons. Minimum temperature standardized seasonal anomaly did not significantly contribute to the climate or combination model ($p > 0.05$ for both models). Additionally, population density did not significantly contribute to either the anthropogenic or combination model ($p > 0.05$ for both). It was also highly correlated with intensive land use (Pearson Correlation 0.94; Fig. S1), indicating intensive land use can be used as a proxy for population density. Both non-significantly contributing variables were removed from the models. We

then ran all of our LMMs with the HUC random effects crossed with the nested year within season random effects. We were then able to separate the variance in percent surface water area explained by the fixed effects (the marginal R^2) and the variance explained by the full model (conditional R^2; Harrison et al., 2018; Schielzeth & Nakagawa, 2012; Tulbure & Broich, 2019), which aided in our comparison of the different categories of models. We also calculated the overall correlation (R^2) between our estimated and observed percent surface water areas, as well as the mean percent error (MPE) at the HUC level, for each model with a different set of explanatory variables (climate, anthropogenic, and their combination; Fig. 2). Because we centered and standardized our explanatory variables, we were able to use regression coefficients ('s) to assess their importance and relationships in the LMMs.

Three models with different sets of explanatory variables were run using the lmer4 library in R (Bates et al., 2015) and the conditional and marginal R^2 values for each model were calculated using the MuMIn library (Barton, 2009). We followed the structure of a random intercept linear mixed effects model (Eq. 2), where y is a matrix of the response variable (percent surface water area), X is a matrix of the fixed effects (the variables for each of our model categories), β is a matrix of the regression coefficients of the fixed effect variables that is calculated by the model, Z is a matrix of the grouping variables (HUCs crossed with the nested year/season), u is the complement to β, and ϵ is a matrix of the residuals (*Introduction to Linear Mixed Models*, 2020).

$$y = X\beta + Zu + \epsilon \quad (2)$$

2.5.2 Random Forest Regression

Random forests are a popular machine learning model based on an ensemble of regression trees for classification or regression and produce an assessment of variable importance (Breiman, 2001; Grömping, 2009). These regression trees are built with a random sample, with replacement, of data from the full dataset and each decision, or split, within each tree is based on a random sample of features, or dependent variables (Breiman, 2001; Grömping, 2009). Each individual regression tree produces its own prediction of the independent variable, but it is unstable and can be overfit to its subset of data. A random forest is composed of a large number of regression trees, and the final prediction values of the independent variable for the full random forest are the average of their predicted values from the individual trees (Breiman, 2001; Grömping, 2009). Random forests converge and improve in accuracy as the number of regression trees increases (Strong Law of Large Numbers) and limit overfitting (Law of Large Numbers; Breiman, 2001).

In our study, we randomly split the data into a testing and a training dataset, with 80% of the data used for training and the remaining 20% used for testing. For each category of model, we generated a random forest consisting of 1000 regression trees using the "scikit-learn" ensemble "RandomForestRegressor" module in Python (Pedregosa et al., 2011). To assess our random forest

models, we calculated the out of bag R^2 (Table S4), which measures the correlation of predicted variables to expected variables for observations not used to generate a regression tree (Grömping, 2009). We also calculated the R^2 of the testing and training data as well as of the full dataset, and we calculated the MPE at the HUC level for each category of model. Lastly, we obtained the feature importance, or Gini importance, for each of the explanatory variables used in each RF model.

2.5.3 Mixed Effects Random Forest

The mixed effects random forest (MERF) method combines the power of controlling for clustered data exemplified by the LMMs with the power of the ensemble of regression trees from RFs (Dey, 2017; Hajjem et al., 2014). The MERF method is similar to the RF, however in MERF the regression trees are replaced with mixed-effects regression trees, that account for the random effects (Hajjem et al., 2014). Hajjem et al. (2014) defined the MERF method by incorporating regression trees into the general form of LMMs (Eq. 2, Eq. 3).

$$y_i = f(X_i) + Z_i b_i + \epsilon_i \quad (3)$$

In Equation 3, the function $f()$ represents the RF applied to the fixed effects covariates, X, within each cluster, i. The second term, $Z_i b_i$, which is similar to the second term, Zu, in Equation 2 where Z_i is a matrix of the random effects covariates for cluster i, accounts for the randomness added by the grouping factors and is assumed to be linear. They assume that between-cluster observations are independent and that b_i, the unknown vector of random effects, and ϵ_i, the vector of errors, are independent and normally distributed. These last two variables are fitted iteratively using an expectation-maximization algorithm until convergence, monitored by computing a generalized log-likelihood criterion. For a more detailed explanation of the MERF method, please see Dey (2017) and Hajjem et al. (2014).

To remain consistent with our methods for the RF models, we again split the data into training and testing datasets, 80% and 20% respectively, and set each RF ($f()$) to have 1000 regression trees. We generated the MERF models for each model category in Python using the "merf" package in the "merf" module (Manifold Inc., 2020), using the "scikit-learn RandomForestRegressor" as the fixed effects estimator. We set the number of iterations for the expectation-maximization algorithm to 200 for each MERF model. In the MERF models, we designated the random effects as the HUCs crossed with the seasons, and we used the same fixed effects as described in Section 2.5.1. Similar to our previous statistical models, we calculated the R^2 for the testing and training datasets as well as the full dataset and the MPE at the HUC level for each MERF model.

For the RF and MERF models, we calculated the Shapely Additive exPlanation (SHAP) values to determine the importance and relationship of each explanatory variable in each machine learning model (Lundberg & Lee, 2017). The SHAP values quantify the contributions of the explanatory variables in the models.

The sum of the SHAP values for each variable is equal to the difference between the prediction of the model and the null model.

3 Results

3.1 Driver Influence and Model Assessment

Overall, we found that the MERF model using a combination of climate and anthropogenic drivers provided the best estimates of seasonal percent surface water area across our study area and per season per year time scale, from 2000 to 2018. Of all nine models, the combination MERF model had the smallest range of HUC MPE (-5.30% to 0.90%, bottom right map, Fig. 3). The combination MERF model also had the largest percent of HUCs with an MPE between -1% and 1% (95.81%), the smallest magnitude median HUC MPE (-0.06%), and the smallest magnitude mean MPE (-0.17%; Figs. 3 and 4). The climate RF model performed the worst, with the largest range of HUC MPEs (-1268.60%, 23.50%, middle left, Fig. 3 and fourth boxplot from the top, Fig. 4). The climate RF model also had the largest percent of HUCs with MPE $<$ -1% (41.94%), underestimating percent surface water area, and the largest percent of HUCs with MPE $>$ 1% (43.87%), overestimating surface water area (middle left, Fig. 3 and fourth boxplot from the top, Fig. 4). The combination models for each statistical model had the smallest magnitude median of HUC MPE between the model types (LMM: -0.11%, RF: -0.16%, MERF: -0.06%; Fig. 4). The LMM and MERF models, which account for random effects, had smaller magnitude median HUC MPE across all model types compared to RF (Fig. 4). All MERF models, regardless of set of explanatory variables, had smaller magnitude median MPEs than LMM or RF models (Fig. 4).

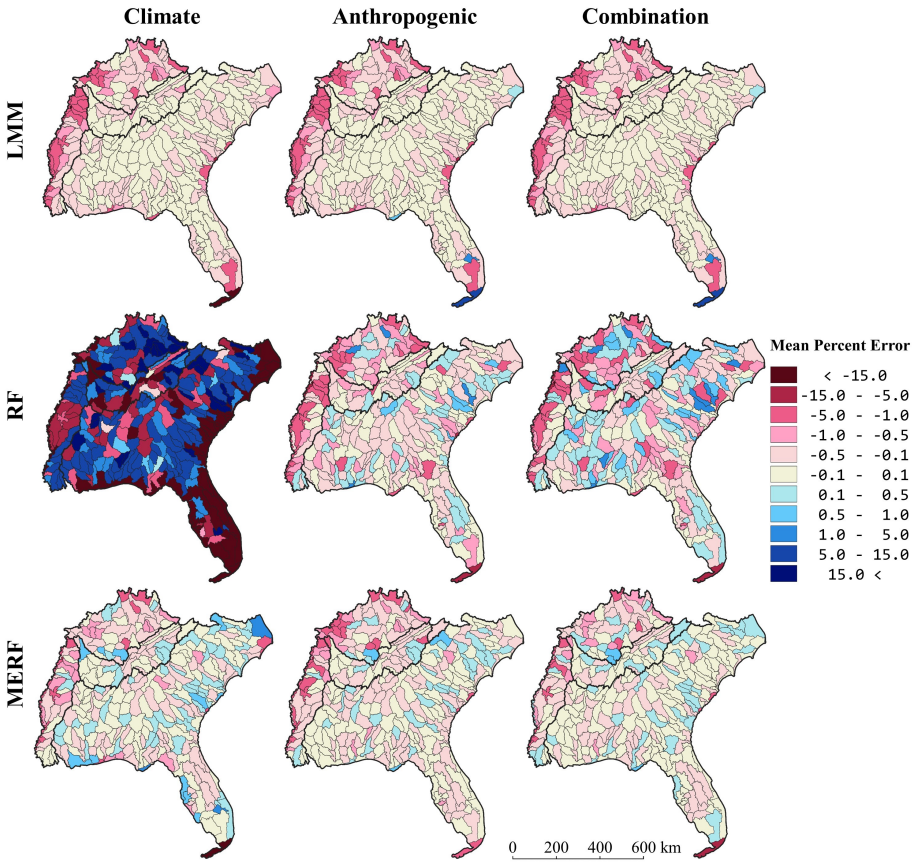

Figure 3. Overall Mean Percent Error at the HUC level for all 9 models. Orange to red HUCs indicate an underestimation of percent surface water area with the magnitude of the underestimation increasing with the hue. Light blue to dark blue HUCs indicate an overestimation of percent surface water area with the magnitude of the overestimation increasing with the hue.

The explanatory power of anthropogenic drivers was greater than that of climate drivers and approximately equal to that of the combination of climate and anthropogenic drivers, according to our LMM results. We directly compared the amount of variance explained by the fixed effects using the marginal R^2. The explanatory variables, or fixed effects, of the climate LMM accounted for <1% of the variance in the model (Climate R^2m, Table 2), indicating that most of the variance in climate LMM was found between HUCs, seasons, and years. In contrast, the anthropogenic and combination LMM explanatory variables explained approximately 37% of the model variance (Anthropogenic and Combination R^2m, Table 2). In LMMs, anthropogenic drivers account for 37.31% more of the variance in the percent surface water area than climate drivers, when the variances between HUCs, seasons, and years (random effects) are all

controlled. For each of the models, the fixed and random effects combined explained >95% of the total variance (Climate R^2c, Table 2).

Table 2. Marginal and Conditional R^2 for the LMM models

	Climate	Anthropogenic	Combination
R^2m	0.0006	0.3737	0.3731
R^2c	0.9723	0.9639	0.9641

Overall, the distribution of the HUC MPEs suggests that the most accurate model category is the combination model, which uses both climate and anthropogenic drivers, and generally the anthropogenic models are more accurate than the climate models. For both the LMMs and the RF models, the magnitude of the median HUC MPE was the smallest for the combination models and the largest for the climate models (Fig. 4). For the MERF models, the magnitude of the median HUC mean percent error was the largest for the anthropogenic model (-0.08%; Fig. 4) and smallest for the combination model (-0.05%; Fig. 4). For both the RF and MERF models, the ranges of HUC MPEs were largest for the climate models and smallest for the combination models (Figs. 3 and 4). For the LMMs, the distribution of HUC MPEs stayed consistent with minimum, mean, and median statistics all varying <1% for all three model categories; however, the maximum HUC MPE varied more with the largest from the anthropogenic model (13.70%) and the smallest for the climate model (0%; Fig. 4).

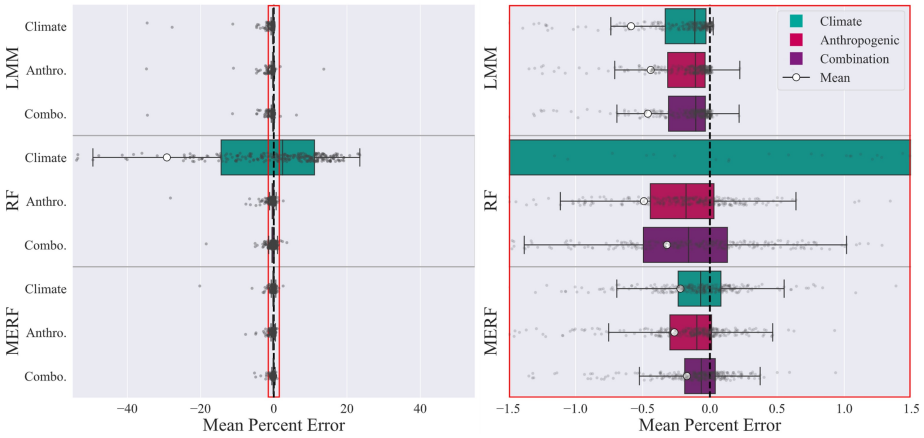

Figure 4. Boxplots of the mean percent error at the HUC level for all 9 models. The full range of MPEs, excluding those in climate RF <-55%, are shown on the left. The distribution of MPEs between -1.5 and 1.5 (red bounding box on both left and right) are shown on the right. The individual HUC MPEs (transparent grey dots) are shown on both left and right.

3.2 Spatial Distribution of Mean Percent Errors

The MPEs were not uniformly spatially distributed across the four water resource regions (Figs. 1 and 3). Excluding the climate RF model, a large majority (>80%) of HUCs for each model had MPEs between -1% and 1%; however, the MPEs varied between the water resource regions (Fig. 3). For the Lower Mississippi water resource region, only the MERF models had a majority (>50%) of HUCs with MPEs between -1% and 1% and none had more than 80% of HUCs with MPEs in that range. These are the smallest proportions of HUCs with MPE between -1% and 1% across the 4 water resource regions, meaning the Lower Mississippi water resource region had the largest magnitudes of error (Fig. 3). The Tennessee and South Atlantic Gulf water resource regions had over 90% of HUCs with MPEs between -1% and 1% across all models excluding the climate RF model. The Ohio water resource region had between 70% and 94% of HUCs with MPEs within this range for each model not including the climate RF model. Only the climate and combination MERF models had more than 90% of HUCs with MPEs between -1% and 1% for the Ohio water resource region. For each water resource region and in total, the combination MERF model had the most HUCs with MPE between -1% and 1% across all water resource regions (79.4% for the Lower Mississippi, 93.3% for the Ohio, 98.49% for the Atlantic, 100% for the Tennessee, and 95.8% in total; bottom row, right column, Fig. 3). The climate RF model had the smallest proportion of HUCs with MPE between -1% and 1% (0% for the Lower Mississippi, 3.5% for the Atlantic, 4.4% for the Ohio, and 6.3% for the Tennessee; middle row, left column, Fig. 3).

Most models underestimated the percentage of surface water area across all water resource regions, with very few HUCs having a MPE >1% (Figs. 3 and 4), overestimating the percent surface water area. The climate RF model was the exception. It had the highest proportion of MPEs >1% across all water resource regions (between 37% and 50%) and 43.9% of HUCs overall (middle row, Fig. 3). All LMM and MERF models had less than 2% of HUCs with MPEs >1% across all water resource regions (Figs. 3 and 4). The climate RF also underestimated the percent surface water area, HUCs with MPE <-1%, the most across all water resource regions (between 28% and 50%) and 41.9% of HUCs overall (middle row, Fig. 3). Excluding the climate RF model, less than 15% of HUCs for each model had MPE <-1% across the study area. However, underestimation of surface water was not evenly distributed across the water resource regions. In the Lower Mississippi water resource region, all LMM and RF models excluding the combination RF model had greater than or equal to 50% of HUCs with MPE <-1% and only the MERF models had less than one third of HUCs with MPEs <-1%. The Ohio water resource region had less than 30% of HUCs with MPE <-1% with climate and combination MERF models with less than 10% of HUCs in this range. The Tennessee and Southern Atlantic Gulf water resource regions have less than 10% of HUCs with MPE <-1%, excluding the climate RF model. The combination MERF model had the smallest percentage of HUCs with MPE <-1% for each water resource region

(between 0% and 21%) and across the whole study area (4.2%).

3.3 Variable Importance

Overall, the percent of natural land cover was the most influential variable across all statistical models (Fig. 5). The anthropogenic variables had higher feature importance's than the climate variables in the combination models. Because we used seasonal standardized anomalies, these results mean that centered and standardized anthropogenic variables had a stronger influence on estimating percent seasonal surface water area than centered and standardized climate anomalies (not the raw average seasonal precipitation or temperature data). As discussed in the previous two sections, the combination models were better (i.e., had the smallest error) than the climate and anthropogenic-only models

Despite anthropogenic variables having higher feature importance than climate variables, the latter were still found to be significant in estimating percent surface water area in both climate-only and combination models. In our LMM variable selection for the climate-only and combination models, both the precipitation and maximum temperature anomalies had a significant effect on the estimation of percent surface water area ($p < 0.01$). Seasonal precipitation anomaly had a positive impact on the estimate of percent surface water area across the LMM and MERF statistical models with both climate-only and a combination of climate and anthropogenic explanatory variables. In the climate and combination LMM and MERF models, precipitation anomaly had a larger influence on estimating precent surface water change than seasonal maximum temperature anomaly (Fig. 5 left and right). For all models using climate data, seasonal maximum temperature anomaly had a negative relationship with the percent surface water area.

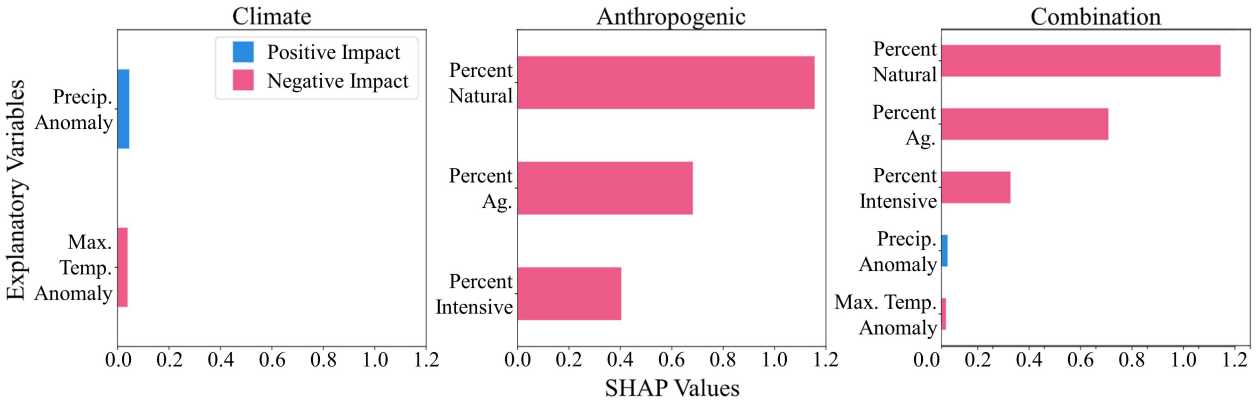

Figure 5. Absolute values of mean SHAP values for the MERF models with different sets of explanatory variables. From left to right: climate variables, anthropogenic variables, and both climate and anthropogenic variables.

All of the percent land cover classes (natural, agricultural, and intensive) had a significant effect on the estimation of percent surface water area ($p < 0.01$)

for both the anthropogenic-only and combination LMMs. All three of these explanatory variables had a negative relationship with estimating percent surface water change across all three statistical models for both anthropogenic-only and combination explanatory variables (Fig. 5 middle and right). For each of these models, the percent of natural land cover was consistently the most influential, followed by percent of agricultural land cover, and percent intensive land cover (Fig. 5 middle and right). We found that for each of the statistical models with both climate and anthropogenic explanatory variables, all the anthropogenic variables had higher variable importance than either of the climate variables (Fig. 5 right).

3.4 HUC Example

We selected a HUC in the center of our study area to illustrate the results of each of our models (Fig. 6). This HUC is one of Georgia's main watersheds and includes part of the metropolitan Atlanta area as well as Lake Lanier, a reservoir that supplies water to the area. The impact of the major droughts from 2000-2018 can be noticed in the observed percent surface water area (blue points in scatterplots, Fig. 6). The 2006–2009 drought was the longest duration of continuous drought in Georgia during this time (National Integrated Drought Information Systems), and we can see the steady decrease in percent surface water area (blue points in scatterplots, Fig. 6). We can also observe a shorter, but severe drought in 2011 and 2012 (blue points in scatterplots, Fig. 6).

The LMMs of each category estimate the percent of surface water almost identically, regardless of the independent variables (top row of scatterplots, Fig. 6). We also see that, except for the climate RF model, the percent error range is larger for the LMMs than for any other model (top row of boxplots, Fig. 6). We see that the climate RF model is consistently underestimating surface water, which is highlighted in the distribution of the percent errors (middle left scatterplot and middle right boxplot, Fig. 6). We also see that the anthropogenic RF model estimates do not capture the seasonality of the percent surface water, because the anthropogenic variables do not have seasonal differences. The seasonality of the climate variables likely helps reduce the range of the percent errors in the combination RF model. Overall, the MERF models have the smallest range of percent errors when compared with the other statistical models, by category (top row of boxplots, Fig. 6). Despite an outlier in the summer of 2004, which appears in both the climate and combination MERF models, the combination MERF model has the most compact distribution of percent errors, making it the model with the smallest amount of error for the watershed.

These graphs highlight how we are able to accurately estimate the percent surface water area for each season in a year at the HUC level using these different categories and statistical models. They highlight some of the limitations of each model, as well as the ability of these models to capture small changes in percent surface water area between seasons.

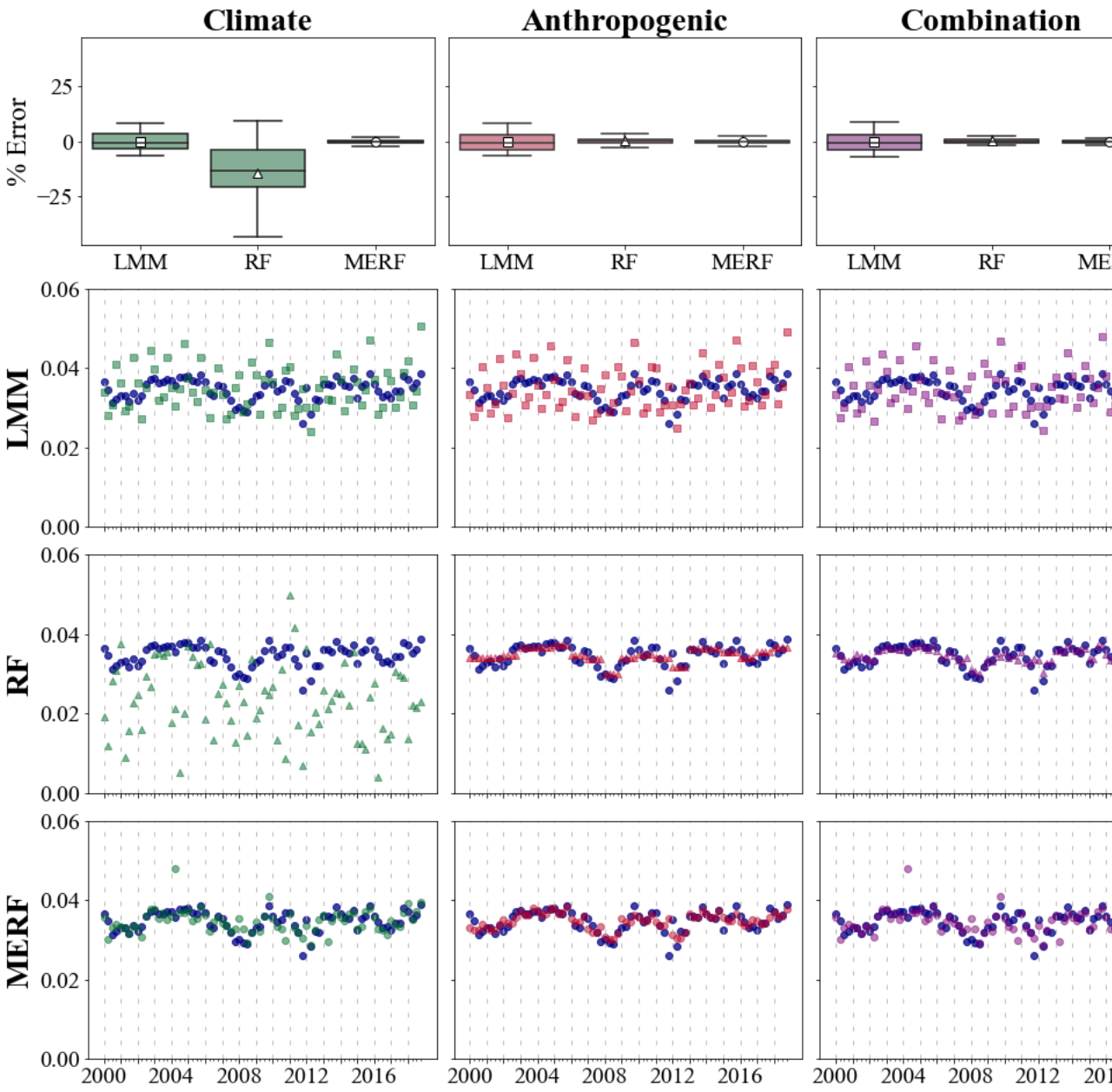

Figure 6. An example of model outputs at the HUC level. The HUC in question (03130001) is highlighted in red on the map in the upper right corner. This particular HUC contains part of Atlanta and Lake Lanier. The row of boxplots at the top show the distribution of the percent errors (y-axis) for each statistical model (x-axis) and category of model (column). The column of boxplots on the right shows the distribution of percent errors (y-axis) for each statistical model

(row) and category of model (x-axis). The 9 scatterplots show the observed percent surface water area in blue and the estimated percent surface water area in colors and shapes relative to the category and statistical model used. Climate models are all shown in green, anthropogenic models are all shown in red, and the combination models are shown in purple. Square points indicate LMMs, triangles indicate RFs, and non-blue circles indicate MERFs.

4 Discussion

The novelty of this study is to use top-down data-driven models to assess how different climate and anthropogenic drivers affect the variability of surface water in a region of the U.S. experiencing more LULC change and population growth than any other in the country (McManamay et al., 2019; Terando et al., 2014). The Southeastern U.S. is also experiencing increasingly severe hurricanes and dry periods (Allen et al., 2016; Emanuel et al., 2013; Pielke et al., 2008) leading to billions of dollars in recovery costs (Smith, 2020) and legal conflicts between states (McManamay et al., 2019; Missimer et al., 2014). These models can help identify areas where land use and water management practices, which are easier to regulate than climate change, are crucial to mitigating and/or adapting to water stress. Understanding the dynamics of surface water in this region can also help decision makers prepare for different water stresses, from either drier or wetter periods. This knowledge can be particularly effective in an area where a large portion of counties are listed under the highest level of social vulnerability in the Centers for Disease Control's 2016 Social Vulnerability Index (Flanagan et al., 2011).

4.1 Spatial Distribution of Model Surface Water Under-/Over- Estimations Across Water Resource Regions

We found that MPEs were not evenly distributed across the water resource regions. The Lower Mississippi water resource region is the region that had the largest range of MPEs across its HUCs and consistently underestimated the percent surface water area in the HUCs (Fig. 3). The regional economy in the Lower Mississippi is driven by agricultural production (Alhassan et al., 2019). To reduce crop stress and optimize crop yields using less cropland area, farmers rely on irrigation from groundwater and/or surface water (Massey et al., 2017; Yasarer et al., 2020). The amount of irrigation in the region has increased (Yasarer et al., 2020) and may have enabled an increase in planting corn, which is a water-demanding crop (Smidt et al., 2016). Increased irrigation, groundwater or surface water, can lead to streamflow depletion (Killian et al., 2019), decreased baseflow, and more frequent low flow conditions (Yasarer et al., 2020). All of these impacts of irrigation, and the increase in irrigation throughout the region, can contribute to changes in surface water area.

Building dams and reservoirs has been the main response to the growing water needs for domestic, industrial, and irrigation purposes in urban and agricultural areas (Altinbilek, 2002; Hubacek et al., 2009). These dams and reservoirs are the primary source of the increase in surface water across the globe (Pekel et al.,

2016). According to the US Army Corps of Engineers' Inventory of Dams, there are 24,222 dams in our study region (US Army Corps of Engineers). Of the dams with completion dates, 708 were built in the region after 2000, 1,759 were completed between 1985 and 2000, and 13,254 were built before 1985. In the Lower Mississippi water resource region, 163 dams have been built since 2000, the largest proportional increase (4.69%) of all the water resource regions in our study area. There are also 8,501 dams without recorded completion dates. Small water bodies, such as small reservoirs, have higher fluctuation in surface water than larger water bodies in both seasons of high rain and of increased drought risk (Zeng et al., 2020; Zou et al., 2017). Zeng et al. 2020 found that building dams and reservoirs helped explain permanent surface water increases in their study area. It is likely that dam construction in our study area impacted surface water area. Because the number of dams in this area has increased over time, it could help explain why our models skew toward underestimating surface water area across the HUCs (Figs. 3 and 4).

4.2 Drivers of Surface Water

Our work leverages three robust statistical methods, unlike other studies that have modeled surface water change with climate and human variables (Xu et al., 2019; Zeng et al., 2020). Similar studies have been able to assess the importance of explanatory variables to estimating surface water but have not captured the fluctuations in surface water at the seasonal level (Zeng et al., 2020). There can be strong variability in surface water between seasons (Pekel et al., 2016), which can adversely impact models that do not account for these variabilities (Zeng et al., 2020). We found significant differences in seasonal variability of surface water, which led us to using seasons as a grouping factor in our LMMs and MERF models. Controlling for the variance in seasonality likely improved our models.

While many studies focus on the impacts of climate change on water resources (Xu et al, 2019), we assessed the impact of climate and anthropogenic drivers independently and in combination. We found that anthropogenic drivers were more influential in estimating percent surface water area than climate drivers in our study region. Our findings are supported by similar studies identifying forested and urban land cover as the most important variables in estimating permanent surface water in the Northeast and Loess Plateau in China (Zeng et al. 2020), identifying increased water management from an expansion of agricultural area as the major source of lake volume decline in Urmia Lake in northwestern Iran (Chaudhari et al., 2018), and finding that drastic changes in vegetation cover changes the patterns of surface water in Canada (Egginton et al., 2014). Our finding that natural land cover, which is mainly composed of forested land cover, was the most important variable of all our climate and anthropogenic variables, supports other studies that found that surface water is significantly correlated with forested vegetation (Caldwell et al., 2016; Liu et al., 2020; Wei et al., 2017; Zeng et al., 2020).

The relationships of our climate variables, positive for precipitation anomalies

and negative for maximum temperature anomalies (Fig. 5), to percent surface water is similar to what other studies have found (Liu et al., 2020; Lockaby et al., 2013; Tulbure & Broich, 2019). In contrast to our findings, Xu et al, 2019 found that climate change had a greater impact on water retention than land cover change in the Upper Yangtze River Basin in western China. However, they used a process-based model and used land cover to calculate evapotranspiration as a proxy for anthropogenic drivers (Xu et al., 2019). We used data-driven models that empirically incorporated land cover. Similar to our results, Xu et al. (2019) found that both climate and land cover change were impactful on water retention, and they found that the impacts of each vary significantly across their study region. They and others have also concluded that adjusting land cover can be used as an effective and direct way to mitigate the impacts of climate change (Xu et al., 2019; Zeng et al., 2020). Water resource management can impact everything from urban water supply to irrigation to ecosystem services (Hester & Larson, 2016; Jeong et al., 2015; Liu et al., 2020; Xu et al., 2019; Yasarer et al., 2020). Additionally, other research across the globe (the Turbio river sub-basin in Mexico, the Murray Darling Basin in Australia, and the Northeast and Loess Plateau in China, respectively; Orozco et al., 2020; Tulbure and Broich, 2019; and Zeng et al., 2020) found that land cover had a larger impact on water resources than climate variables.

The negative relationship of our most influential variable, natural land cover, to percent surface water supports the results of other studies (Liu et al., 2020; Lockaby et al., 2013; Tulbure & Broich, 2019). There are conflicting accounts on whether forestation increases or decreases streamflow and surface water resources, indicating that there are likely other factors at play (Zeng et al., 2020; Zhang et al., 2017; Zhang & Wei, 2021). Researchers are currently working to understand how forest change effects water supply (Zhang & Wei, 2021), so studies on changes to water supply in heavily forested areas undergoing large land cover/land use change are highly relevant. Zhang et al. (2017) found that forestation reduced streamflow more often in semiarid and arid regions while these reductions were less pronounced in humid subtropical and tropical area. Other studies have indicated that increases in forest area can reduce runoff and help restore groundwater, thereby stabilizing the variability in surface water (Biao et al., 2010; Caldwell et al., 2014; C. Li et al., 2020; Zhang & Wei, 2021). Specifically, studies in the Southeastern US, Mexico, and China found that natural land cover diminishes runoff and improves water quality (Liu et al., 2020; Orozco et al., 2020; Zeng et al., 2020). Less variability in surface water in response to natural land cover supports the negative relationship we found between percent natural land cover and percent surface water area. Additionally, the negative relationship could be because both variables are a proportion in each HUC. For one to be larger, the other would have to be smaller. It is possible that natural land cover is the most influential fixed effect because it is the dominant land cover type in our study area (Fig. 1). However, because we centered and normalized our independent variables, this should have less influence.

Land use and land cover change plays a large role in surface water dynamics (Liu et al., 2020; Orozco et al., 2020; Tulbure & Broich, 2019; Xu et al., 2019). Zhou et al., 2021 also found that land use, especially agricultural activities within a watershed, impact surface water area. Orozco et al. (2020) found that urban centers and agricultural areas have the greatest water vulnerability, and that climate and land use change impacts this vulnerability. Our results showed that across statistical models, anthropogenic drivers (i.e., land use) improved model estimates of percent surface water by decreasing the magnitudes of HUC MPEs.

4.3 Climate RF Model

With the exception of the climate RF model, we were able to estimate percent surface water area accurately (Figs. 3 and 4). While we could have added complexity to the LMM and MERF models using random slopes as well as random intercepts to account for the grouping factors, they performed accurately without the added complexity and the tradeoffs with model interpretability. Across all nine models, only the climate RF model had an R^2 <0.95. The combination MERF model had the highest R^2 at 0.995. The RF models are limited in comparison to the LMM and MERF models because they do not control for variance explained by differences between HUCs, seasons, or years. Because we see the largest distribution of HUC MPEs for the climate RF model (Figs. 3 and 4) and the lowest R^2 value for the climate RF model, we can infer that for the climate models, the grouping factors are just as if not more important than the climate variables. This result is supported by our findings in the LMMs where the variance explained by the climate variables for the climate model was <1% (marginal R^2; Table 2) and the variance explained by the random effects was 97.17% (difference between conditional and marginal R^2).

4.4 Climate RF Model

Uncertainty is present in all of the datasets we used, which can lead to modeling errors. The climate data is at a coarser spatial scale than the surface water or land cover data. Because we aggregated all the data to the HUC scale, we assume that the precipitation and temperature anomalies will be reasonably captured; however, a finer scale of climate data could yield more accurate results. To improve our capture of anthropogenic drivers, future work could add irrigation data, which would change according to the seasons. At this time however, and to the best of our knowledge, a spatially explicit seasonal record of irrigation does not exist for our study area. In addition to irrigation data, future work consisting of adding explanatory variables (Liang et al., 2019; Tulbure & Broich, 2019) to improve our understanding of the driver categories should consider including soil type, soil moisture, slope, elevation, stream gauge data, and time-lag factors such as precipitation or surface water area from the proceeding season (Heimhuber et al., 2017; Liang et al., 2019; Tulbure & Broich, 2019; Walker et al., 2020). Another limitation and avenue for future work is the distinction between natural and man-made/modified surface water bodies. The DSWE dataset does not distinguish between the two, but there is some evidence that these water features exhibit different patterns of variability and

are impacted differently by climate and anthropogenic forces (Poff et al., 2007), and therefore disentangling them should be a focus of future work.

Moreover, increasing the types of machine learning techniques to compare data-driven models (Liang et al., 2019), and expanding the combination MERF model into projection models based on different future climate scenarios (Orozco et al., 2020) represent promising future directions of this work. Such machine learning models include K-Nearest Neighbor, Support Vector Machine-Non Linear, and Deep Feed Neural Networks–each of which have different strengths and weaknesses (Liang et al., 2019). Expanding the combination MERF model into projection models using data for different Representative Concentration Pathway and Shared Socioeconomic Pathways could increase the potential impact of these top-down data-driven models of surface water (Duan et al., 2019; Orozco et al., 2020; Vepraskas et al., 2020) when used for predicting the impact of future climate and land use on surface water resources.

5 Conclusions

Incorporating anthropogenic drivers with climate drivers to model surface water using top-down data-driven models expands our ability to assess surface water dynamics. In this study, we used climate and anthropogenic variables to develop nine different models to estimate percent surface water at the HUC level. Our climate-only models used precipitation and temperature variables, our anthropogenic-only models used land cover variables, and our combination models used all the climate and anthropogenic variables. We found that between climate and anthropogenic drivers, the latter is more influential and that when used together, they yield more accurate results. We also found that the MERF model, a combination of the LMM and RF models, produced better percent surface water estimates than either the LMMs or RF models. Lastly, we found that the percent of natural land cover was the overall most important explanatory variable in estimating the percent surface water.

Our time period covered both large droughts and hurricane seasons, highlighting that our data-driven models can be reasonably applied to a wide range of environmental variability for this region. The results of this study emphasize the role of land cover in estimating surface water, indicating that LULC management practices can be used to limit or adapt to future drought or flooding events. Additional research can be done to develop these models into projections to help local decision makers adapt to future environmental variability.

Acknowledgments

Funding for this work came from a North Carolina State University start-up package to Tulbure. To our knowledge, no conflict of interest are present.

Data Availability Statement

DSWE data (Jones, 2015, 2019) provided in this manuscript can be accessed from the USGS Earth Explorer data portal (https://earthexplorer.usgs.gov/), and can be downloaded after creating an account. The GRIDMET dataset

from the University of Idaho (Abatzoglou, 2013), the CDL dataset from the U.S. Department of Agriculture NASS ("CropScape - NASS CDL Program"), and the NLCD data from the USGS (Homer et al., 2020) can be accessed from Google Earth Engine (https://developers.google.com/earth-engine/datasets/catalog/IDAHO_EPSCOR_GRIDMET; https://developers.google.com/earth-engine/datasets/catalog/USDA_NASS_CDL; https://developers.google.com/earth-engine/datasets/catalog/USGS_NLCD_RELEASES_2016_REL), and can be downloaded after making an account. The LandScan data produced by Oak Ridge National Laboratory (Rose et al., 2020) can be accessed through the LandScan data portal (https://landscan.ornl.gov/landscan-datasets) after registering an account.

References

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, *33*(1), 121–131. https://doi.org/10.1002/joc.3413Alhassan, M., Lawrence, C. B., Richardson, S., & Pindilli, E. J. (2019). The Mississippi Alluvial Plain aquifers—An engine for economic activity. In *Fact Sheet*. https://doi.org/10.3133/fs20193003Allen, M. R., Fernandez, S. J., Fu, J. S., & Olama, M. M. (2016). Impacts of climate change on sub-regional electricity demand and distribution in the southern United States. *Nature Energy*, *1*. https://doi.org/10.1038/NENERGY.2016.103Alnahit, A. O., Mishra, A. K., & Khan, A. A. (2020). Quantifying climate, streamflow, and watershed control on water quality across Southeastern US watersheds. *Science of the Total Environment*, *739*, 139945. https://doi.org/10.1016/j.scitotenv.2020.139945Altinbilek, D. (2002). The role of dams in development. *Water Science and Technology : A Journal of the International Association on Water Pollution Research*, *45*(8), 169–180.Barton, K. (2009). *Mu-MIn: Multi-model inference* (R Package Version 0.12.2/r18). http://r-forge.r-project.org/projects/mumin/Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using {lme4}. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01Bhaduri, B., Bright, E., Coleman, P., & Urban, M. L. (2007). LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal 2007 69:1*, *69*(1), 103–117. https://doi.org/10.1007/S10708-007-9105-9Biao, Z., Wenhua, L., Gaodi, X., & Yu, X. (2010). Water conservation of forest ecosystem in Beijing and its value. *Ecological Economics*, *69*(7), 1416–1426. https://doi.org/10.1016/J.ECOLECON.2008.09.004Blankenau, P. A., Kilic, A., & Allen, R. (2020). An evaluation of gridded weather data sets for the purpose of estimating reference evapotranspiration in the United States. *Agricultural Water Management*, *242*, 106376. https://doi.org/10.1016/j.agwat.2020.106376Breiman, L. (2001). Random Forests. In *Machine Learning* (Vol. 45).Caldwell, Muldoon, C., Miniat, C.

F., Cohen, E., Krieger, S., Sun, G., McNulty, S., & Bolstad, P. V. (2014). Quantifying the role of National Forest System lands in providing surface drinking water supply for the Southern United States. *General Technical Report - Southern Research Station, USDA Forest Service*, *SRS-197*, 1–135.Caldwell, P. V., Miniat, C. F., Brantley, S., Elliott, K., Laseter, S., & Swank, W. (2016). Long term records provide insights on the relative influence of climate and forest community structure on water yield in the southern Appalachians. *In: Stringer, Christina E.; Krauss, Ken W.; Latimer, James S., Eds. 2016. Headwaters to Estuaries: Advances in Watershed Science and Management -Proceedings of the Fifth Interagency Conference on Research in the Watersheds*, 238.Carter, L. M., Terando, A., Dow, K., Hiers, K., Kunkel, K. E., Lascurain, A., Marcy, D. C., Osland, M. J., & Schramm, P. J. (2018). *Chapter 19 : Southeast. Impacts, Risks, and Adaptation in the United States: The Fourth National Climate Assessment, Volume II*. https://doi.org/10.7930/NCA4.2018.CH19Chaudhari, S., Felfelani, F., Shin, S., & Pokhrel, Y. (2018). Climate and anthropogenic contributions to the desiccation of the second largest saline lake in the twentieth century. *Journal of Hydrology*, *560*(342–353). https://doi.org/10.1016/j.jhydrol.2018.03.034Clark, M. P., Bierkens, M. F. P., Samaniego, L., Woods, R. A., Uijlenhoet, R., Bennett, K. E., Pauwels, V. R. N., Cai, X., Wood, A. W., & Peters-Lidard, C. D. (2017). The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism. *Hydrol. Earth Syst. Sci*, *21*, 3427–3440. https://doi.org/10.5194/hess-21-3427-2017*CropScape - NASS CDL Program*. (n.d.). https://nassgeodata.gmu.edu/CropScape/Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature Climate Change*. https://doi.org/10.1038/NCLIMATE1633Dey, S. (2017). *Mixed Effects Random Forests in Python*. Towards Data Science. https://towardsdatascience.com/mixed-effects-random-forests-6ecbb85cb177Diem, J. E., Stauber, C. E., & Rothenberg, R. (2017). Heat in the southeastern United States: Characteristics, trends, and potential health impact. *PLoS ONE*, *12*(5). https://doi.org/10.1371/journal.pone.0177937Dieter, C. A., Maupin, M. A., Caldwell, R. R., Harris, M. A., Ivahnenko, T. I., Lovelace, J. K., Barber, N. L., & Linsey, K. S. (2018). Estimated use of water in the United States in 2015. *Circular*. https://doi.org/10.3133/CIR1441Dobson, J. E., Bright, E. A., Coleman, P., Durfee, R. C., & Worley, B. A. (2000). LandScan: A Global Population Database for Estimating Populations at Risk. *PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING*, *66*(7), 849–857. https://www.researchgate.net/publication/267450852Duan, K., Caldwell, P. V., Sun, G., McNulty, S. G., Zhang, Y., Shuster, E., Liu, B., & Bolstad, P. V. (2019). Understanding the role of regional water connectivity in mitigating climate change impacts on surface water supply stress in the United States. *Journal of Hydrology*, *570*(December 2018), 80–95. https://doi.org/10.1016/j.jhydrol.2019.01.011Duan, K., Sun, G., Mcnulty, S. G., Caldwell, P. V., Cohen, E., Sun, S., Aldridge, H. D., Zhou, D., Zhang, L., & Zhang, Y. (2017). Future shift of the relative roles of precipitation and temperature in controlling annual runoff in the conterminous

United States. *Hydrology and Earth System Sciences*, *21*(11), 5517–5529. https://doi.org/10.5194/hess-21-5517-2017Egginton, P., Beall, F., & Buttle, J. (2014). Reforestation-climate change and water resource implications. *The Forestry Chronicle*, *90*(4), 516–524.Emanuel, K. A., Santer, B. D., & Lawrence, E. O. (2013). Downscaling CMIP5 climate models shows increased tropical cyclone activity over the 21st century. *ATMOSPHERIC, AND PLANETARY SCIENCES*, *110*. https://doi.org/10.1073/pnas.1301293110Engström, J., Praskievicz, S., Bearden, B., & Moradkhani, H. (2021). Decreasing water resources in Southeastern U.S. as observed by the GRACE satellites. *Water Policy*, *23*(4), 1017–1029. https://doi.org/10.2166/WP.2021.039Ferguson, C. R., Pan, M., & Oki, T. (2018). The Effect of Global Warming on Future Water Availability: CMIP5 Synthesis. *Water Resources Research*, *54*(10), 7791–7819. https://doi.org/10.1029/2018WR022792Flanagan, B. E., Gregory, E. W., Hallisey, E. J., Heitgerd, J. L., & Lewis, B. (2011). A Social Vulnerability Index for Disaster Management. *Journal of Homeland Security and Emergency Management*, *8*(1). https://doi.org/10.2202/1547-7355.1792Flörke, M., Schneider, C., & McDonald, R. I. (2018). Water competition between cities and agriculture driven by climate change and urban growth. *Nature Sustainability*, *1*(1), 51–58. https://doi.org/10.1038/s41893-017-0006-8Foley, J. A., Defries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N., & Snyder, P. K. (2005). Global Consequences of Land Use. *Science*, *309*(5734), 570–574. https://www.jstor.org/stable/3842335Foster, M., Peterson, M. N., Cubbage, F., & McMahon, G. (2019). Evaluating natural resource planning for longleaf pine ecosystems in the Southeast United States. *Forest Policy and Economics*. https://doi.org/10.1016/j.forpol.2018.11.008Google Earth Engine. (2021). *Statistics of an Image Region Google Earth Engine*. https://developers.google.com/earth-engine/guides/reducers_reduce_regionGrömping, U. (2009). Variable importance assessment in regression: Linear regression versus random forest. *American Statistician*, *63*(4), 308–319. https://doi.org/10.1198/tast.2009.08199Hajjem, A., Bellavance, F., & Larocque, D. (2014). Mixed-effects random forest for clustered data. *Journal of Statistical Computation and Simulation*, *84*(6), 1313–1328. https://doi.org/10.1080/00949655.2012.741599Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *2018*(5), e4794. https://doi.org/10.7717/peerj.4794Heimhuber, V., Tulbure, M. G., & Broich, M. (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). https://doi.org/10.1002/2016WR019858Hernandez-Ochoa, I. M., & Asseng, S. (2018). Cropping Systems and Climate Change in Humid Subtropical Environments. *Agronomy*, *8*(2), 19. https://doi.org/10.3390/agronomy8020019Hester, C. M., & Larson, K. L.

(2016). Time-Series Analysis of Water Demands in Three North Carolina Cities. *Journal of Water Resources Planning and Management*, *142*(8), 880–891. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000659Homer, C., Dewitz, J., Jin, S., Xian, G., Costello, C., Danielson, P., Gass, L., Funk, M., Wickham, J., Stehman, S., Auch, R., & Riitters, K. (2020). Conterminous United States land cover change patterns 2001–2016 from the 2016 National Land Cover Database. *ISPRS Journal of Photogrammetry and Remote Sensing*, *162*, 184–199. https://doi.org/10.1016/J.ISPRSJPRS.2020.02.019Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N., Wickham, J., & Megown, K. (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. https://doi.org/10.14358/PERS.81.5.345Hostetler, S. W., & Alder, J. R. (2016). Implementation and evaluation of a monthly water balance model over the US on an 800 m grid. *Water Resources Research*, *52*(12), 9600–9620. https://doi.org/10.1002/2016WR018665Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel Analysis : Techniques and Applications, Second Edition*. Taylor & Francis Group. http://ebookcentral.proquest.com/lib/ncsu/detail.action?docID=574571Hubacek, K., Guan, D., Barrett, J., & Wiedmann, T. (2009). Environmental implications of urbanization and lifestyle change in China: Ecological and Water Footprints. *Journal of Cleaner Production*, *17*, 1241–1248. https://doi.org/10.1016/j.jclepro.2009.03.011Iceland, C. (2017). *Water Stress is Helping Drive Conflict and Migration*. World Resources Institute. https://www.wri.org/insights/water-stress-helping-drive-conflict-and-migrationIngram, K. T., Carter, L., Dow, K., Anderson, J., Asseng, S., Hopkinson, C., Konrad, C., McNulty, S., Mitchell, K., Moody, K., Quattrochi Marshall, D., Schramm, P., Sun, G., & Swann, L. (2012). *Climate Change in the Southeast USA: Executive Summary.Introduction to Linear Mixed Models*. (2020). UCLA: Statistical Consulting Group. https://stats.idre.ucla.edu/other/multpkg/introduction-to-linear-mixed-models/Jeong, H., Minne, E., & Crittenden, J. C. (2015). Life cycle assessment of the City of Atlanta, Georgia's centralized water system. *International Journal of Life Cycle Assessment*, *20*(6), 880–891. https://doi.org/10.1007/S11367-015-0874-YJones, J. W. (2015). Efficient wetland surface water detection and monitoring via landsat: Comparison with in situ data from the everglades depth estimation network. *Remote Sensing*, *7*(9), 12503–12538. https://doi.org/10.3390/rs70912503Jones, J. W. (2019). Improved automated detection of subpixel-scale inundation-revised Dynamic Surface Water Extent (DSWE) partial surface water tests. *Remote Sensing*, *11*(4). https://doi.org/10.3390/rs11040374Jordan, J. L. (2001). Negotiating Water Allocations Using a Com Prehensive Study Form At : the " Tri-State W Ater Wars ." *Journal of Contemporary Water Research and Education*, *118*(1), 38–43.Karl, T. R., Melillo, J. M., & Peterson, T. C. (2009). *Global Climate Change Impacts in the United States*.Keellings, D., & Engström, J. (2019). The Future of Drought in the Southeastern U.S.: Projections from Downscaled

CMIP5 Models. *Water*, *11*(2), 259. https://doi.org/10.3390/w11020259Killian, C. D., Asquith, W. H., Barlow, J. R. B., Bent, G. C., Kress, W. H., Barlow, P. M., & Schmitz, D. W. (2019). Characterizing groundwater and surface-water interaction using hydrograph-separation techniques and groundwater-level data throughout the Mississippi Delta, USA. *Hydrogeology Journal*, *27*(6), 2167–2179. https://doi.org/10.1007/S10040-019-01981-6Kossin, J. ., Hall, T., Knutson, T., Kunkel, K. E., Trapp, R. J., Waliser, D. E., & Wehner, M. F. (2017). Extreme Storms. In D. J. Wuebbles, D. W. Fahey, K. A. Hibbard, D. J. Dokken, B. C. Stewart, & T. K. Maycock (Eds.), *Climate Science Special Report: Fourth National Climate Assessment* (Volume I, pp. 257–276). U.S. Global Change Research Program. https://doi.org/10.7930/J07S7KXXKunkel, K. E., Easterling, D. R., Hubbard, K., & Redmond, K. (2004). Temporal variations in frost-free season in the United States: 1895–2000. *Geophysical Research Letters*, *31*(3). https://doi.org/10.1029/2003GL018624Lark, T. J., Meghan Salmon, J., & Gibbs, H. K. (2015). Cropland expansion outpaces agricultural and biofuel policies in the United States. *Environmental Research Letters*, *10*(4). https://doi.org/10.1088/1748-9326/10/4/044003Lark, T. J., Mueller, R. M., Johnson, D. M., & Gibbs, H. K. (2017). Measuring land-use and land-cover change using the U.S. department of agriculture's cropland data layer: Cautions and recommendations. *International Journal of Applied Earth Observation and Geoinformation*, *62*(March), 224–235. https://doi.org/10.1016/j.jag.2017.06.007Li, C., Sun, G., Caldwell, P. V., Cohen, E., Fang, Y., Zhang, Y., Oudin, L., Sanchez, G. M., & Meentemeyer, R. K. (2020). Impacts of Urbanization on Watershed Water Balances Across the Conterminous United States. *Water Resources Management*, *56*(e2019WR926574), e2019WR026574. https://doi.org/10.1029/2019WR026574Li, L., Skidmore, A., Vrieling, A., & Wang, T. (2019). A new dense 18-year time series of surface water fraction estimates from MODIS for the Mediterranean region. *Hydrology and Earth System Sciences*, *23*(7), 3037–3056. https://doi.org/10.5194/HESS-23-3037-2019Liang, J., Li, W., Bradford, S. A., & Šimůnek, J. (2019). Physics-informed data-driven models to predict surface runoffwater quantity and quality in agricultural fields. *Water*, *11*(200). https://doi.org/10.3390/w11020200Liu, N., Dobbs, G. R., Caldwell, P. V., Miniat, C. F., Bolstad, P. V, Nelson, S., & Sun, G. (2020). Quantifying the role of State and private forest lands in providing surface drinking water supply for the Southern United States. *General Technical Report GTR-SRS-248*. https://doi.org/10.2737/SRS-GTR-248Lockaby, G., Nagy, C., & Vose, J. M. (2013). Forests and water. *Forestry Chronicle*, *79*(6), 1050–1051.Lundberg, S., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, *2017-December*, 4766–4775. https://arxiv.org/abs/1705.07874v2Manifold Inc. (2020). *merf · PyPI*. PyPI. https://pypi.org/project/merf/Massey, J. H., Stiles, C. M., Epting, J. W., Powers, S. R., Kelly, D. B., Bowling, T. H., Janes, C. L., & Pennington, D. A. (2017). Long-term measurements of agronomic crop irrigation made in the Mississippi delta portion of the lower Mississippi River Valley. *Irrigation Science*, *35*, 297–313. https://doi.org/10.1007/s00271-017-0543-yMatusick, G.,

Hudson, S. J., Garrett, C. Z., Samuelson, L. J., Kent, J. D., Addington, R. N., & Parker, J. M. (2020). Frequently burned loblolly–shortleaf pine forest in the southeastern United States lacks the stability of longleaf pine forest. *Ecosphere*, *11*(2), e03055. https://doi.org/10.1002/ecs2.3055McKee, J. J., Rose, A. N., Bright, E. A., Huynh, T., & Bhaduri, B. L. (2015). Locally adaptive, spatially explicit projection of US population for 2030 and 2050. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(5), 1344–1349. https://doi.org/10.1073/pnas.1405713112McManamay, R. A., DeRolph, C. R., Surendran-Nair, S., & Allen-Dumas, M. (2019). Spatially explicit land-energy-water future scenarios for cities: Guiding infrastructure transitions for urban sustainability. *Renewable and Sustainable Energy Reviews*, *112*(March), 880–900. https://doi.org/10.1016/j.rser.2019.06.011Missimer, T. M., Danser, P. A., Amy, G., & Pankratz, T. (2014). Water crisis: the metropolitan Atlanta, Georgia, regional water supply conflict. *Water Policy*, *16*(4), 669–689.Mourtzinis, S., Ortiz, B. V., & Damianidis, D. (2016). Climate Change and ENSO Effects on Southeastern US Climate Patterns and Maize Yield. *Scientific Reports*, *6*(1), 1–7. https://doi.org/10.1038/srep29777Müller, M. F., Yoon, J., Gorelick, S. M., Avisse, N., & Tilmant, A. (2016). Impact of the Syrian refugee crisis on land use and transboundary freshwater resources. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(52). https://doi.org/10.1073/pnas.1614342113National Integrated Drought Information Systems. (n.d.). *Drought in Georgia from 2000-Present*. U.S. Drought Monitor. Retrieved July 30, 2021, from https://www.drought.gov/states/georgiaNazemi, A., & Wheater, H. S. (2015). On inclusion of water resource management in Earth system models – Part 2: Representation of water supply and allocation and opportunities for improved modeling. *Hydrology and Earth System Sciences*, *19*(1), 63–90. https://doi.org/10.5194/hess-19-63-2015Orozco, I., Martínez, A., Ortega, V., Martinez, A., & Ortega, V. (2020). Assessment of the Water, Environmental, Economic and Social Vulnerability of a Watershed to the Potential Effects of Climate Change and Land Use Change. *Water*, *12*(1682), 1682. https://doi.org/10.3390/w12061682Palazzoli, I., & Ceola, S. (2020). *Anthropogenic and climatic controls on surface water loss across USA*. EGU General Assembly 2020. https://doi.org/https://doi.org/10.5194/egusphere-egu2020-254Palmer, M. A., Reidy Liermann, C. A., Nilsson, C., Flörke, M., Alcamo, J., Lake, P. S., & Bond, N. (2008). Climate change and the world's river basins: Anticipating management options. *Frontiers in Ecology and the Environment*, *6*(2), 81–89. https://doi.org/10.1890/060148Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.htmlPekel, J.-F. F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422. https://doi.org/10.1038/nature20584Perin, V., Tulbure, M. G., Gaines, M.

D., Reba, M. L., & Yaeger, M. A. (2021). On-farm reservoir monitoring using Landsat inundation datasets. *Agricultural Water Management*, *246*. https://doi.org/10.1016/j.agwat.2020.106694Perry, M. T. (2020). *rasterstats · PyPI*. PyPI. https://pypi.org/project/rasterstats/0.7.0/Piao, S., Ciais, P., Huang, Y., Shen, Z., Peng, S., Li, J., Zhou, L., Liu, H., Ma, Y., Ding, Y., Friedlingstein, P., Liu, C., Tan, K., Yu, Y., Zhang, T., & Fang, J. (2010). The impacts of climate change on water resources and agriculture in China. *Nature*, *467*(7311), 43–51. https://doi.org/10.1038/nature09364Pickens, A. H., Hansen, M. C., Hancher, M., Stehman, S. V., Tyukavina, A., Potapov, P., Marroquin, B., & Sherani, Z. (2020). Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series. *Remote Sensing of Environment*, *243*, 111792. https://doi.org/10.1016/J.RSE.2020.111792Pielke, R. A., Gratz, J., Landsea, C. W., Collins, D., Saunders, M. A., & Musulin, R. (2008). Normalized Hurricane Damage in the United States: 1900–2005. *Natural Hazards Review*, *9*(1), 29–42. https://doi.org/10.1061/_ASCE1527-6988_20089:1_29Poff, N. L., Olden, J. D., Merritt, D. M., & Pepin, D. M. (2007). Homogenization of regional river dynamics by dams and global biodiversity implications. *Proceedings of the National Academy of Sciences*, *104*(14), 5732–5737. https://doi.org/10.1073/PNAS.0609812104Pokhrel, Y. N., Hanasaki, N., Wada, Y., & Kim, H. (2016). Recent progresses in incorporating human land–water management into global land surface models toward their integration into Earth system models. *Wiley Interdisciplinary Reviews: Water*, *3*(4), 548–574. https://doi.org/10.1002/wat2.1150Human Appropriation of Renewable Fresh Water, 271 Science 785 (1996). https://doi.org/10.1126/SCIENCE.271.5250.785Rose, A. N., McKee, J. J., Sims, K. M., Bright, E. A., Reith, A. E., & Urban, M. L. (2020). *LandScan 2019* (2019 RI-). Oak Ridge National Laboratory SE - July 1, 2020. https://landscan.ornl.gov/Sanchez, G. M., Terando, A., Smith, J. W., García, A. M., Wagner, C. R., & Meentemeyer, R. K. (2020). Forecasting water demand across a rapidly urbanizing region. *Science of the Total Environment*, *730*, 139050. https://doi.org/10.1016/j.scitotenv.2020.139050Schielzeth, H., & Nakagawa, S. (2012). Nested by design: model fitting and interpretation in a mixed model era. *Methods in Ecology AndEvolution*, *4*, 14–24. https://doi.org/10.1111/j.2041-210x.2012.00251.xSleeter, B., Loveland, T. R., Domke, G. M., Herold, N., Wickham, J., & Wood, N. J. (2018). *Chapter 5: Land Cover and Land Use Change. Impacts, Risks, and Adaptation in the United States: The Fourth National Climate Assessment, Volume II*. https://doi.org/10.7930/NCA4.2018.CH5Smidt, S. J., Haacker, E. M. K., Kendall, A. D., Deines, J. M., Pei, L., Cotterman, K. A., Li, H., Liu, X., Basso, B., & Hyndman, D. W. (2016). Complex water management in modern agriculture: Trends in the water-energy-food nexus over the High Plains Aquifer. *Science of the Total Environment*, *566–567*, 988–1001. https://doi.org/10.1016/J.SCITOTENV.2016.05.127Smith, A. B. (2020). U.S. Billion-dollar Weather and Climate Disasters, 1980 - present. *NOAA National Centers for Environmental Information*. https://doi.org/10.25921/STKW-7W73Sood, A., & Smakhtin, V. (2015).

Global hydrological models: a review. *Hydrological Sciences Journal*, *60*(4), 549–565. https://doi.org/10.1080/02626667.2014.950580Soulard, C. E., Walker, J. J., & Petrakis, R. E. (2020). Implementation of a Surface Water Extent Model in Cambodia using Cloud-Based Remote Sensing. *Remote Sensing*, *12*(984). https://doi.org/10.3390/rs12060984Steadman, R. G. (1984). A universal scale of apparent temperature. *Journal of Climate & Applied Meteorology*, *23*(12), 1674–1687. https://doi.org/10.1175/1520-0450(1984)023<1674:AUSOAT>2.0.CO;2Terando, A. J., Costanza, J. K., Belyea, C., Dunn, R. R., McKerrow, A., & Collazo, J. A. (2014). The southern megalopolis: Using the past to predict the future of urban sprawl in the Southeast U.S. *PLoS ONE*, *9*(7). https://doi.org/10.1371/journal.pone.0102261Thrasher, B., Xiong, J., Wang, W., Melton, F., Michaelis, A., & Nemani, R. (2013). Downscaled Climate Projections Suitable for Resource Management. *Eos, Transactions American Geophysical Union*, *94*(37), 321–323. https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2013EO370002Tulbure, M. G., & Broich, M. (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. https://doi.org/10.1016/j.isprsjprs.2013.01.010Tulbure, M. G., & Broich, M. (2019). Spatiotemporal patterns and effects of climate and land use on surface water extent dynamics in a dryland region with three decades of Landsat satellite data. *Science of the Total Environment*, *658*, 1574–1585. https://doi.org/10.1016/j.scitotenv.2018.11.390Tulbure, M. G., Broich, M., Stehman, S. V., & Kommareddy, A. (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. https://doi.org/10.1016/j.rse.2016.02.034Turner, B. L., Lambin, E. F., & Reenberg, A. (2007). The emergence of land change science for global environmental change and sustainability. *Proceedings of the National Academy of Sciences*, *104*(52), 20666–20671. https://doi.org/10.1073/PNAS.0704119104U.S. Environmental Protection Agency. (2004). *Level III and IV Ecoregions of EPA Region 4*. 2004.U.S. Geological Survey. (2019a). *LANDSAT DYNAMIC SURFACE WATER EXTENT (DSWE) PRODUCT GUIDE*.U.S. Geological Survey. (2019b). *USGS EROS Archive - Landsat - Landsat Level-3 Dynamic Surface Water Extent (DSWE) Science Product*. https://doi.org/10.5066/F7445KQKUniversity of Virginia Weldon Cooper Center. (2018). National Population Projections. In *Demographics Research Group*. https://demographics.coopercenter.org/national-population-projectionsUS Army Corps of Engineers. (n.d.). *National inventory of dams*. Federal Emergency Management Agency. https://search.library.wisc.edu/catalog/999834648302121Veldkamp, T. I. E. I. E., Wada, Y., Aerts, J. C. J. H. C. J. H., Döll, P., Gosling, S. N., Liu, J., Masaki, Y., Oki, T., Ostberg, S., Pokhrel, Y., Satoh, Y., Kim, H., & Ward, P. J. (2017). Water scarcity hotspots travel downstream due to human interventions in the 20th and 21st century. *Nature Communications 2017 8:1*, *8*(1), 1–12. https://doi.org/10.1038/ncomms15697Vepraskas, M. J., Skaggs, & R. W., & Caldwell, P. V. (2020). Method to Assess Climate Change Impacts on Hydrologic Boundaries of Individual Wetlands. *Wetlands*, *40*, 365–376.

https://doi.org/10.1007/s13157-019-01183-6Vörösmarty, C. J., Green, P., Salisbury, J., & Lammers, R. B. (2000). Global Water Resources: Vulnerability from Climate Change and Population Growth. *Science*, *289*(5477), 284–288. https://doi.org/10.1126/SCIENCE.289.5477.284Vörösmarty, C. J., Mcintyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R., & Davies, P. M. (2010). Global threats to human water security and river biodiversity. *Nature*, *467*(7315), 555–561. https://doi.org/10.1038/nature09440Wada, Y., Bierkens, M. F. P., De Roo, A., Dirmeyer, P. A., Famiglietti, J. S., Hanasaki, N., Konar, M., Liu, J., Müller Schmied, H., Oki, T., Pokhrel, Y., Sivapalan, M., Troy, T. J., van Dijk, A. I. J. M., Van Emmerik, T., Van Huijgevoort, M. H. J., Van Lanen, H. A. J., Vörösmarty, C. J., Wanders, N., … Wheater, H. (2017). Human-water interface in hydrological modelling: current status and future directions Human-water interface in hydrological modelling. *Hydrology and Earth System Sciences*, *21*(8), 4169–4193. https://doi.org/10.5194/hess-21-4169-2017Walker, J. J., Soulard, C. E., & Petrakis, R. E. (2020). Integrating stream gage data and Landsat imagery to complete time-series of surface water extents in Central Valley, California. *International Journal of Applied Earth Observation and Geoinformation*, *84*(July 2019), 101973. https://doi.org/10.1016/j.jag.2019.101973Walsh, C. J., Roy, A. H., Feminella, J. W., Cottingham, P. D., Groffman, P. M., & Ii, R. P. M. (2005). The urban stream syndrome: current knowledge and the search for a cure. *Am. Benthol. Soc*, *24*(3), 706–723.Wei, X., Li, Q., Zhang, M., Giles-Hansen, K., Liu, W., Fan, H., Wang, Y., Zhou, G., Piao, S., & Liu, S. (2017). Vegetation cover-another dominant factor in determining global water resources in forested regions. *Global Change Biology*, 1–10. https://doi.org/10.1111/gcb.13983Wickham, J., Stehman, S. V, Sorenson, D. G., Gass, L., & Dewitz, J. A. (2021). Thematic accuracy assessment of the NLCD 2016 land cover for the conterminous United States. *Remote Sensing of Environment*, *257*, 112357. https://doi.org/10.1016/j.rse.2021.112357Xu, P., Guo, Y., & Fu, B. (2019). Regional impacts of climate and land cover on ecosystemwater retention services in the Upper Yangtze River Basin. *Sustainability (Switzerland)*, *11*(19), 5300. https://doi.org/10.3390/su11195300Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S. M., Case, A., Costello, C., Dewitz, J., Fry, J., Funk, M., Granneman, B., Liknes, G. C., Rigge, M., & Xian, G. (2018). A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. *ISPRS Journal of Photogrammetry and Remote Sensing*, *146*, 108–123. https://doi.org/10.1016/J.ISPRSJPRS.2018.09.006Yasarer, L. M. W., Taylor, J. M., Rigby, J. R., & Locke, M. A. (2020). Trends in Land Use, Irrigation, and Streamflow Alteration in the Mississippi River Alluvial Plain. *Frontiers in Environmental Science*, *8*, 66. https://doi.org/10.3389/fenvs.2020.00066Zeng, Y., Yang, X., Fang, N., & Shi, Z. (2020). Large-scale afforestation significantly increases permanent surface water in China's vegetation restoration regions. *Agricultural and Forest Meteorology*, *290*(April), 108001. https://doi.org/10.1016/j.agrformet.2020.108001Zhang, M., Liu, N., Harper, R., Li, Q., Liu, K., Wei, X., Ning, D., Hou, Y., & Liu, S. (2017). *A global*

review on hydrological responses to forest change across multiple spatial scales: Importance of scale, climate, forest type and hydrological regime. https://doi.org/10.1016/j.jhydrol.2016.12.040Zhang, M., & Wei, X. (2021). Deforestation, forestation, and water supply. *Science*, *371*(6533), 990–991. https://doi.org/10.1126/SCIENCE.ABE7821Zhou, X., Polcher, J., & Dumas, P. (2021). Representing Human Water Management in a Land Surface Model Using a Supply/Demand Approach. *Water Resources Research*, *57*(4), e2020WR028133. https://doi.org/10.1029/2020WR028133Zou, Z., Dong, J., Menarguez, M. A., Xiao, X., Qin, Y., Doughty, R. B., Hooker, K. V., Hambright, K. D., Z, Z., J, D., MA, M., X, X., Y, Q., RB, D., KV, H., & K, D. H. (2017). Continued decrease of open surface water body area in Oklahoma during 1984–2015. *Science of the Total Environment*, *595*, 451–460. https://doi.org/10.1016/j.scitotenv.2017.03.259