Surface water is the most readily accessible water resource and provides an array of ecosystem services, but its availability and access are 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 United States. 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 Southeast.

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. 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 in three different ways. First, we 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 of 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 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 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 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 (Anderson 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 (Postel et al., 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-state 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 and 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 data sets derived from moderate resolution (30 m, 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 data sets, 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 data set; 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).

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 to 2018 using three different sets of explanatory variables—climate, anthropogenic, and their combination. The specific objectives of this study are to: (a) 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 assesses model performances; and (b) determine which variables are most important in estimating percent surface water area across all models.
2. Data and Methods

2.1. Study Area

This study estimates percent surface water area at the resolution of the 310 eight-digit hydrologic unit code watersheds (hereafter, “HUCs” or “watersheds”; right map, Figure 1; Duan et al., 2017) in the southeastern U.S. The watersheds can be aggregated into four water resource regions (two-digit hydraulic unit codes; lower left map, Figure 1), which represent drainage areas of major rivers or a series of rivers (Blasch et al., 2018; Seaber et al., 1987). The water resource regions are the first of six levels of hydrologic units in the hydrologic unit system developed by the U.S. Geological Survey (USGS) and the watersheds are the fourth level in the hierarchical system (Seaber et al., 1987). This study uses the water resource regions to describe hydrologically based geographic aggregations of watershed error assessment of estimated percent surface water area (Liu et al., 2021). 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, Figure 1). The entire study area is approximately 1,087,370 km², the average watershed is approximately 3,500 km², the largest is 11,790 km², and the smallest is 180 km².

The rapidly urbanizing southeastern U.S. has experienced the most LULC change of any region in the country since the start of the 21st century (Homer et al., 2020; Sleeter et al., 2018), and its population growth rate is 4% higher than the national average (McManamay et al., 2019; University of Virginia Weldon Cooper Center, 2018). Much of the LULC change is from agriculture or forest to urban area as southern cities experience rapid low-density, sprawling growth (McManamay et al., 2019; Sanchez et al., 2020; Sleeter et al., 2018; Terando et al., 2014). Population growth, paired with increasingly severe hurricanes and dry periods (Allen et al., 2016; Emanuel, 2013; Pielke et al., 2008) due to climate change, can lead to higher water demand in urban areas (Brown et al., 2019; Duan et al., 2019), many of which get their public water supply from surface water sources (McManamay et al., 2019; Missimer et al., 2014). Urban water stress can even lead to regional and inter-state
legal conflicts, as happened between Georgia, Alabama, and Florida due to a dispute about the allocation of water from Lake Lanier to the Atlanta metropolitan area, which experiences water shortages (McManamay et al., 2019; Missimer et al., 2014).

The southeastern U.S. is currently a humid subtropical region (Alnahit et al., 2020; Hernandez-Ochoa & Asseng, 2018; Ingram et al., 2013) 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°C to 29.60°C. Over the next 30 years, summer apparent temperatures—an approximation of a person's experience of temperature (Steadman, 1984)—across the region are expected to increase by 2.4°C–4.1°C. 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 1,345.39 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), forest-dominated land cover accounted for the majority of land cover (74.52%; Table 1) across the three land cover classes (forest-dominated, agricultural, and intensive) we defined using the National Land Cover Data Set (NLCD; Homer et al., 2015; Yang et al., 2018) and the Cropland Data Layer (“CropScape - NASS CDL Program”). Over half of HUCs (60%) lost more than 1% of their forest-dominated land cover between 2000 and 2018 and 14 lost more than 5%. During this time, intensive and agricultural land cover increased by more than 1% across approximately 26% and 34% of HUCs, respectively.

### 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, 2019) 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 data set derived from the U.S. Landsat Analysis Ready Data (ARD) Surface Reflectance product in the Albers Equal Area projection (Jones, 2015, 2019; Landsat Missions, 2019). Landsat ARD is comprised of the most geometrically accurate data from Landsat 4–5 TM, Landsat 7 ETM+, and Landsat 8 OLI (Landsat Missions, 2019). DSWE has been freely accessible since 2019 on the USGS Earth Explorer data portal (https://earthexplorer.usgs.gov/; U.S. Geological Survey, 2019). It was developed using Landsat imagery from across the U.S., but the DSWE model (an inundation algorithm) inputs were designed to be applied globally (i.e., not requiring scene-based training data; Jones, 2015, 2019). 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). It was also tested for temporal trends and bias as a function of hydrologic conditions (Jones, 2015). None were detected, supporting the use of DSWE for long-term surface water inundation monitoring for trend analyses over time (Jones, 2015). For this study, we used the interpreted layer with mask applied (INWM; U.S. Geological Survey, 2019) of the DSWE product. The INWM layer classifies pixels into seven groups assigned via pixel value: 0—no water; 1—water – high confidence; 2—water – moderate confidence; 3—potential wetland; 4—

### Table 1

| 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)              | 1,345.39 | 139.12             | 1,076.45| 1,230.17        | 1,342.35        | 1,434.43        | 1,731.50 |
| Forest-dominated 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/km²)       | 55.52  | 78.22              | 0.5647  | 17.15           | 29.60           | 62.55           | 677.96  |
water or wetland low confidence; 9—cloud, cloud shadow, or snow; 255—not available, fill. 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 in Python (Python Software Foundation, https://www.python.org) to calculate the surface water area for each HUC in the region for each season (each comprised of 3 months beginning with spring on March 1, summer on June 1, fall on September 1, and winter on December 1 which includes January and February in the following calendar year) from 2000 to 2018. To aggregate the DSWE INWM layer by season, we stacked the ARD raster tiles for each season within each year. There was an average of 20 raster tiles per season per year. For each season, we calculated the percent of high confidence water out of all times a value other than cloud or no data was recorded for each pixel. We recorded the percentage of high confidence water as a bin an output raster.

We then used the zonal_stats function from the rasterstats Python module (Perry, 2020) 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. We limited our analysis to using only the class 1 pixels because they were found to have the highest correlation with stream gage data of any combination of DSWE pixel classes (Walker et al., 2020). Because the number of pixels with viable pixel values (i.e., values ≤ 4) in the stacked raster varies across the raster and across tiles, we set our threshold as a percentage rather than a specific number of pixels classified as high confidence water. The 25% threshold likely omitted possible short-term flood extents, which already have a low chance of being captured by the DSWE data due to the temporal frequency of the Landsat-derived product (Heimhuber et al., 2018; Tulbure et al., 2022). However, we found from visual inspection that the threshold reduced some artifacts—likely from unmasked cloud shadows misclassified as high confidence water—that are present in the DSWE high confidence water class, while maximizing the extent of seasonal surface water detected in the DSWE data set.

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) data set from the University of Idaho (Abatzoglou, 2013), a long-term (1979–present) daily time series of moderate (4 km) spatial resolution climate data, was freely available. The data set was validated with a direct comparison to station observations from four weather station networks (Abatzoglou, 2013). When compared to five other gridded weather data sets, 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 data sets (beginning before 2015).

We used the GRIDMET data set to calculate the standardized seasonal anomaly (Equation 1) for each season (i) per year (j) on a per-pixel basis ($SSA_{ij}$) in GEE for each of our climate variables. We first subtracted the average daily climate value of the season and year ($\bar{x}_{ij}$; e.g., precipitation for the spring of 2018) from the seasonal long-term (spring 1985 through winter 2018) average ($\bar{x}_i$; e.g., average spring precipitation). We used the same 3-month seasons defined for processing the DSWE data, and then divided the difference by the seasonal long-term (1985–2018) standard deviation ($\sigma_i$; e.g., standard deviation of spring precipitations).

$$SSA_{ij} = \frac{x_{ij} - \bar{x}_i}{\sigma_i}$$  \hspace{1cm} (1)

Finally, 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 to 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 has increased over time (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). When the 16 Level II land cover classes are combined into 8 Level I classes, the overall accuracy improves by over 5% (Wickham et al., 2021), supporting Lark et al. (2017)'s recommendation of consolidating land cover classifications to improve accuracy. Class consolidation is also recommended when comparing to or using in combination with other data sets (Lark et al., 2015, 2017).

The three classes to which we consolidated the CDL and NLCD land cover classes were agricultural, forest-dominated, 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 in Supporting Information S1). For 2000 through 2007, we similarly reclassified the 2001 and 2006 NLCD data into our three categories (Table S2 in Supporting Information S1). 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 land cover proportions from the 2001 and 2006 NLCD and the 2008 CDL proportions to span 2000–2007 (Table S3 in Supporting Information S1). 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 data set, 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 1-km 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 data set as technological advances have improved 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.

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 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 (Figure 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, forest-dominated, 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 (Equation 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, 2013). 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. For possible fixed effects—also called 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).

We ran a backward stepwise selection process to determine which random effects we should control for, and which fixed effects contributed significantly to the models. We began by considering all fixed and random effects (Figure 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; Figure S1 in Supporting Information S2), indicating intensive land use can be used as a proxy for population density. Both nonsignificantly 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. Using these grouping factors allowed us to account for the correlation of our response variable within the groups (Harrison et al., 2018; Hox et al., 2010; Oskolkov, 2020; Schielzeth & Nakagawa, 2013). 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, 2013; 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; Figure 2). Because we centered and standardized our explanatory variables, we were able to use regression coefficients ($\beta$) to assess their importance and relationships in the LMMs.

Three models with different sets of explanatory variables were fitted and run using the lme4 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 (Equation 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), $\beta$ 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 $\beta$, and $\epsilon$ is a matrix of the residuals (Introduction to Linear Mixed Models, 2020). The models were fitted using the restricted maximum likelihood function to estimate $\beta$ and $u$ (Harrison et al., 2018; Oskolkov, 2020; Sarafian, 2020).

$$y = X\beta + Zu + \epsilon$$  \hspace{1cm} (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 data set. 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 data set, 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 1,000 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 in Supporting Information S1), 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 data set, 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 (Equation 2, Equation 3).

\[ y_i = f(X_i) + Z_i b_i + e_i \]  

(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, \( Z_u \), in Equation 2 where \( Z \) 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. It is assumed that between-cluster observations are independent and that \( b \), the unknown vector of random effects, and \( e \), 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 data sets, 80% and 20%, respectively, and set each RF \( f() \) to have 1,000 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 data sets as well as the full data set 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%, Figure 3c3). 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%; Figures 3 and 4). The climate RF model performed the worst, with the largest range of HUC MPEs (−1,268.60%, 23.50%, Figures 3b1 and 4b1). 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 (Figures 3b1 and 4b1). 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%; Figure 4). The LMM and MERF models, which account for random effects, had smaller magnitude median HUC MPE across all model types compared to RF (Figure 4). All MERF models, regardless of the set of explanatory variables, had smaller magnitude median MPEs than LMM or RF models (Figure 4).

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 ($R^2c$, Table 2).

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 (Figure 4). For the MERF models, the magnitude of the median HUC MPE was the largest for the anthropogenic model ($-0.08\%$; Figure 4c2) and smallest for the combination model ($-0.05\%$; Figure 4c3). For both the RF and MERF models, the ranges of HUC MPEs were largest for the climate models and smallest for the combination models (Figures 3 and 4). For the LMMs, the distribution of HUC MPEs stayed consistent with minimum, mean, and median.

Figure 3. Overall Mean Percent Error at the HUC level for all nine models. Light to dark pink HUCs indicate an underestimation of percent surface water area with the magnitude of the underestimation increasing with the hue. Light to dark blue HUCs indicate an overestimation of percent surface water area with the magnitude of the overestimation increasing with the hue.
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%; Figure 4a2) and the smallest for the climate model (0%; Figure 4a1).

3.2. Spatial Distribution of MPEs

The MPEs were not uniformly spatially distributed across the four water resource regions (Figures 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 (Figure 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 four water resource regions, meaning the Lower Mississippi water resource region HUCs had the largest magnitudes of error (Figure 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 South Atlantic Gulf, 100% for the Tennessee, and 95.8% in total; Figure 3c3). 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 South Atlantic Gulf, 4.4% for the Ohio, and 6.3% for the Tennessee; Figure 3b1).

| Table 2 | Marginal and Conditional $R^2$ for the LMM Models |
|---------|--------------------------------------------------|
|         | Climate | Anthropogenic | Combination |
| $R^2$m  | 0.0006  | 0.3737        | 0.3731      |
| $R^2$c  | 0.9723  | 0.9639        | 0.9641      |

Most models underestimated the HUC percentage of surface water area across all water resource regions, with very few HUCs having an MPE > 1% (Figures 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, Figure 3). All LMM and MERF models had less than 2% of HUCs with MPEs > 1% across the water resource regions.
(Figures 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, Figure 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 South Atlantic Gulf water resource regions had less than 10% of HUCs with MPE < −1% for all models, 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 forest-dominated land cover was the most influential variable across all statistical models (Figure 5). The anthropogenic variables had higher feature importance 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 percent surface water change than seasonal maximum temperature anomaly (Figure 5, left and right). For all models using climate data, seasonal maximum temperature anomaly had a negative relationship with the percent surface water area.

All of the percent land cover classes (forest-dominated, 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 (Figure 5, middle and right). For each of these models, the percent of forest-dominated land cover was consistently the most
influential, followed by the percent of agricultural land cover and the percent intensive land cover (Figure 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 (Figure 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 (Figure 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 to 2018 can be noticed in the observed percent surface water area (blue points in scatterplots, Figure 6). The 2006–2009 drought was the longest duration of continuous drought in Georgia during this time (National Integrated Drought Information Systems 2021), and we can see the steady decrease in percent surface water area (blue points in scatterplots, Figure 6). We can also observe shorter, but severe droughts such as in 2011 and 2012 (blue points in scatterplots, Figure 6).

The LMMs of each category estimate the percent of surface water almost identically, regardless of the independent variables (top row of scatterplots, Figure 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, Figure 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, Figure 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, Figure 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.

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, 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. Some process-based models were developed to assess the impacts on ecological mechanisms by climate and land use changes in water management scenarios (Freeman et al., 2013). They highlighted the need for remotely sensed data to quantify sub-yearly changes in land cover to improve the physical components of their dynamic landscape models (Freeman et al., 2013). 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-/Overestimations 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 (Figure 3). Similarly, for seven out of nine models the Ohio water resource region had around a quarter of its study-area HUCs underestimating percent surface water area with an MPE < −1%. A possible explanation for the poorer model performance in the Lower Mississippi and Ohio water resource regions is that both regions have less than 50% of their HUCs in the study area. Because the water resource regions describe drainage areas for major rivers or series of rivers, it is likely that seasonal trends in percent surface water area will be similar. Therefore, excluding over half of the watersheds in the Lower Mississippi and Ohio water resource regions reduces the number of samples of these trends and could lead to greater uncertainty. Additionally, we found that the HUCs within our study area of these water resource regions saw the greatest reductions of forest-dominated land cover (an average of 2.4% and 1.6% loss from 2001 to 2018 for the Ohio and Lower Mississippi, respectively). The Lower Mississippi water resource region HUCs within our study area also had the lowest average area of forest-dominated land cover (64.2%). Because we found forest-dominated land cover to be the most influential variable, the lower proportion of this land cover in study-area HUCs in the Lower Mississippi and the reduction of this land cover HUCs for both the Lower Mississippi and Ohio water resource regions likely influenced the higher MPEs from the models.

Additional uncertainty in models may come from an increase in irrigated areas in the study region (Yasarer et al., 2020). The regional economy in the Lower Mississippi water resource region 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 study region has increased (Yasarer et al., 2020) and may have enabled the increase in corn, which is a water-demanding crop (Smidt et al., 2016), grown in the region (there was over 12,000 ha more land cover classified as corn in 2018 than in 2008; CropScape - NASS CDL Program, 2021). 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 have 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 (U.S. Army Corps of Engineers' Inventory of Dams, 2020) there are 24,222 dams in our study region. 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 portion of the Lower Mississippi water resource region in our study area, 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. 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 and reservoir construction in our study area impacted surface water area, and will continue to do so as the number of small reservoirs has been projected to increase in agricultural areas of the U.S. (Downing et al., 2006). The increase in the number of dams and reservoirs in this area over time could help explain why our models skew toward underestimating surface water area across the HUCs (Figures 3 and 4). We do note, however, that depending on the steepness of the containment edges, we may not be able to detect fluctuations in surface water area despite changes in volume in these dams and reservoirs.

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 for 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). Because we used seasons as a grouping factor, we were able to account for seasonal variability in all the LMM and MERF models including those that did not have seasonally changing explanatory variables (i.e., the anthropogenic-only models). 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 forest-dominated 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., 2021; Wei et al., 2017; Zeng et al., 2020).

The relationships of our climate variables, positive for precipitation anomalies and negative for maximum temperature anomalies (Figure 5), to percent surface water is similar to what other studies have found (Liu et al., 2021; 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., 2021; 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 & Broich, 2019; 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, forest-dominated land cover, to percent surface water supports the results of other studies (Liu et al., 2021; 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 influences 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 areas. 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 U.S., Mexico, and China found that forest-dominated land cover diminishes runoff and improves water quality (Liu et al., 2021; Orozco et al., 2020; Zeng et al., 2020). Less variability in surface water in response to forest-dominated land cover supports the negative relationship we found between percent forest-dominated 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 forest-dominated land cover is the most influential fixed effect because it is the dominant land cover type in our study area (Figure 1). However, because we centered and normalized our independent variables, this should have less influence.

LULC change plays a large role in surface water dynamics (Liu et al., 2021; 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 (Figures 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 (Figures 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. Limitations and Future Work

Uncertainty is present in all of the data sets 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. One limitation of our methods was the use of only the high confidence water classification from the DSWE product. Walker et al. (2020) found the strongest correlation between the aggregation of these class 1 pixels and stream-gage-measured discharge amounts than any other combination of water or wetland classes for stream gages in the Central Valley, California. However, just under half of these gages had stronger correlations with the combination of high confidence and moderate confidence (class 1 + class 2; Walker et al., 2020). Omitting class 2 from our method likely reduced our ability to capture shorter-term water over vegetated areas. Including class 2 and comparing results with only class 1 pixels should be considered in future work.

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. However, spatially aggregated irrigation data does exist for large areas within our study region (Torak & Painter, 2011) and efforts are being made to increase available data on irrigation water use across the U.S. (Painter et al., 2021). 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 gage 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). Additionally, error was likely introduced with the annual application of the ~5-year NLCD data before 2007. Because this accounts for many of our total years, we may have reduced the variability in some percentages of land cover classes that fluctuate more than we captured. For example, forested areas grown as a crop could be cleared during a year that was not captured in the NLCD data and could have been classified as something other than forest-dominated. By artificially removing land cover variance, we may have increased the error in the estimates. Future studies should use annual data from the USGS Land Change Monitoring, Assessment, and Projection project which is based on Landsat data and thus goes from 1985 to 2020 in the most recent collection (1.2; U.S. Geological Survey, 2021). Another limitation and avenue for future work is the distinction between natural and man-made/modified surface water bodies. The DSWE data set 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 has 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. Finally, we found that the percent of forest-dominated 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.

Conflict of Interest
The authors declare no conflicts of interest relevant to this study.

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 data set from the University of Idaho (Abatzoglou, 2013), the CDL data set 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/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://landsat.ornl.gov/landscan-datasets) after registering an account.

Acknowledgments
Funding for this work came from a North Carolina State University start-up package to Tulbure.

References
Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, 33(11), 121–131. https://doi.org/10.1002/joc.3413
Albassam, M., Lawrence, C. B., Richardson, S., & Pendill, E. J. (2019). The Mississippi Alluvial Plain aquifers—An engine for economic activity. *Fact Sheet*. https://doi.org/10.3133/fs2019003
Allen, 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/ENERGY.2016.103
Alnhami, 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.2020.139945
Allinblik, 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. https://doi.org/10.2166/wst.2002.0172
Anderson, D. M., Boesch, D. F., Burkett, V. R., Carter, L. M., Cohen, S. J., Grimm, N. B., et al. (2009). In T. R. Karl, J. M. Melillo, & T. C. Peterson (Eds.), *Global climate change impacts in the United States*. Cambridge University Press.
Barton, K. (2009). Mu-Mn: Multi-model inference. Retrieved from http://forge.t-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.i01
Bhaduri, 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*, 69(1), 103–117. https://doi.org/10.1007/S10708-007-9105-9
Biao, 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.2009.08.004
Blankenau, 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.2020.106376
Blasch, K., Hund, S., Wurster, P., Sando, R., & Berthelote, A. (2018). Streamflow contributions from tribal lands to major river basins of the United States. *PLOS One*, 13(9), e0203872. https://doi.org/10.1371/journal.pone.0203872
Breiman, L. (2001). Random forests. In *Machine learning* (Vol. 45). https://doi.org/10.1023/a:1010933404324
Brown, T. C., Mahat, V., & Ramirez, I. A. (2019). Adaptation to future water shortages in the United States caused by population growth and climate change. *Earth's Future*, 7(3), 219–234. https://doi.org/10.1029/2018EF001091
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 C. E. Stringer, K. W. Krauss, & J. S. Latimer (Eds.), *Science of the Total Environment*, 67(1), 1–48. https://doi.org/10.1016/j.scitotenv.2019.05.033
Carter, L. M., Terando, A., Dow, K., Hiers, K., Kunkel, K. E., Lascuain, A., et al. (2018). Southeast. In D. R. Reidmiller, C. W. Averv, D. R. Easterling, K. E. Kunkel, K. L. M. Lewis, T. K. Maycock, & B. C. Stewart (Eds.), *Impacts, risks, and adaptation in the United States: The fourth national climate assessment: Vol. II* (pp. 743–751). U.S. Global Change Research Program. https://doi.org/10.7930/NCA4.2018.CH19
Chaudhary, S., Felfeli, F., Shin, S., & Potthast, 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.jhydrology.2018.03.034
Ingram, K. T., Carter, L., Dow, K., Anderson, J., Asseng, S., Hopkinson, C., et al. (2013). Climate change in the southeast USA: Executive summary. In K. T. Ingram, K. Dow, L. Carter, & J. Anderson (Eds.), Climate of the southeast United States: Variability, change, impacts, and vulnerability (pp. 1–7). Island Press. https://doi.org/10.5822/978-1-61091-509-0_1
Water Resources Research

Introduction to Linear Mixed Models. (2020). UCLA: Statistical Consulting Group. Retrieved from https://stats.idre.ucla.edu/other/mult-pkg/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-y

Jones, 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/rs70912503

Jones, J. W. (2019). Improved automated detection of subpixel-scale inundation-revised Dynamic Surface Water Extent (DSWE) partial surface water tests. Remote Sensing, 11(4), 374. https://doi.org/10.3390/rs11040374

Jordan, J. L. (2001). Negotiating water allocations using a comprehensive study format: The “Tri-State Water Wars. Journal of Contemporary Water Resource Education, 11(1), 38–43. https://opensuc.lib.siu.edu/jcwre/vol11/iss1/6

Keellings, D., & Engstrom, J. (2019). The future of drought in the southeastern U.S. Projections from downscaled CMIP5 models. Water, 11(259). https://doi.org/10.3390/w11020259

Killian, 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/s10040-019-01981-6

Kossin, J. P., Hall, T., Knutson, T., Kunkel, K. E., Trapp, R. J., Wilser, 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 Change Special Report: Fourth National Climate Assessment (Vol. I, pp. 257–276). U.S. Global Change Research Program. https://doi.org/10.7930/J0757KXX

Kunkel, K. E., Easterling, D. R., Hubbard, R. K., & Redmond, K. (2004). Temporal variations in frost-free season in the United States: 1895–2000. Geophysical Research Letters, 31(3). https://doi.org/10.1029/2003GL018624

Landsat Missions. (2019). Landsat dynamic surface water extent (DSWE) product guide. Retrieved from https://www.usgs.gov/media/files/landsat-dynamic-surface-water-extent-product-guide

Lark, 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), 044003. https://doi.org/10.1088/1748-9326/10/4/044003

Lark, 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.2017.06.007

Li, C., Sun, G., Caldwell, P. V., Cohen, E., Fang, Y., Zhang, Y., et al. (2020). Impacts of urbanization on watershed water balances across the conterminous United States. Water Resources Research, 56, e2019WR026574. https://doi.org/10.1029/2019WR026574

Li, L., Skidmore, A., Vrielinck, A., & Wang, T. (2019). A new dense 18-year time series of surface water fraction estimates from MODIS for the conterminous United States. Remote Sensing of Environment, 259, 3037–3056. https://doi.org/10.1016/j.rse.2019.07.034

Liang, J., Li, W., Bradford, S. A., & Šimůnek, J. (2019). Physics-informed data-driven models to predict surface runoff water quantity and quality in agricultural fields. Water, 11, 200. https://doi.org/10.3390/w11020200

Liu, N., Caldwell, P. V., Dobbs, G. R., Ford Miniat, C., Bolstad, P. V., Nelson, S. A., & Sun, G. (2021). Forested lands dominate drinking water supply in the conterminous United States. Environmental Research Letters, 16(8), 0844008. https://doi.org/10.1088/1748-9326/aae90d

Lockaby, G., Nagy, C., & Vose, J. M. (2013). Forests and water. The Forestry Chronicle, 9(6), 1050–1051.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), Advances in neural information processing systems (Vol. 30). Curran Associates, Inc. Retrieved from https://proceedings.neurips.cc/paper/2017/file/8a20a862/19786327fd6c43de62b67767-Paper.pdf

Manifold Inc. (2020). merf - PyPi. PyPi. Retrieved from https://pypi.org/project/merf/

Massey, J. H., Stiles, C. M., Eipting, J. W., Powers, S. R., Kelly, D. B., Bowling, T. H., et al. (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/s00227-017-0543-y

Matsuick, 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/ecs2.3055

McKay, 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.1407131112

McManamy, R. A., DeRolph, C. R., Surenondra-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, 880–900. https://doi.org/10.1016/j.rser.2019.06.011

Missaime, T. M., Danser, P. A., Amy, G., & Pankratz, T. (2014). Water crisis: The metropolitan Atlanta, Georgia, regional water supply conflict. Water Policy, 16, 669–689. https://doi.org/10.2166/wp.2014.131

Mourtonis, S., Ortiz, B. V., & Damianidis, D. (2016). Climate change and ENSO effects on southeastern US climate patterns and maize yield. Scientific Reports, 6, 29777. https://doi.org/10.1038/srep29777

Müller, 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), 14932–14937. https://doi.org/10.1073/pnas.1614342113

National Integrated Drought Information Systems. (2021). Drought in Georgia from 2000—Present. U.S. Drought Monitor. Retrieved from https://www.drought.gov/states/georgia

Ninemt, 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, 63–90. https://doi.org/10.5194/hess-19-63-2015

Ortiz, I., Martinez, 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). https://doi.org/10.3390/w12061682

Oskolok, N. (2020). Linear mixed model from scratch. Today Data Science. Retrieved from https://towardsdatascience.com/linear-mixed-model-from-scratch-2982624580a4

Painter, J. A., Brandt, J. T., Caldwell, R. R., Haynes, J. V., & Read, A. L. (2021). Documentation of methods and inventory of irrigation information collected for the 2013 U.S. Geological Survey estimated use of water in the United States. https://doi.org/10.3334/ia20205139

Palazzoli, I., & Ceola, S. (2020). Anthropogenic and climatic controls on surface water loss across USA. (Vol. 2020). EGU General Assembly. https://doi.org/10.5194/egusphere-ega2020-254
Palmer, M. A., Reidy Liermann, C. A., Nilsson, C., Florke, 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/10.1890/060148

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(85), 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.html

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

Perin, 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, 106694. https://doi.org/10.1016/j.agwat.2020.106694

Perry, M. T. (2020). *rasterstats - PyPy*. PyPy. Retrieved from https://ppypi.org/project/rasterstats/0.7.0/

Piao, S., Ciais, P., Huang, Y., Shen, Z., Peng, S., Li, J., et al. (2010). The impacts of climate change on water resources and agriculture in China. *Nature*, 467(7311), 43–51. https://doi.org/10.1038/nature09364

Picketn, A. H., Hansen, M. C., Hancher, M., Stehman, S. V., Tyskavina, A., Potapov, P., et al. (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.2020.111792

Pielke, R. A., Gratz, J., Landsda, 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/(asce)1527-6989(2008)9:1(29)

Poff, 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 of the United States of America*, 104(14), 5732–5737. https://doi.org/10.1073/pnas.0609812104

Pokhrel, 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/wat2.1150

Postel, S. L., Daily, G. C., & Ehrlich, P. R. (1996). Human appropriation of renewable fresh water. *Science*, 271(5250), 785–788. https://doi.org/10.1126/science.271.5250.785

Rose, A. N., McKee, J. J., Sims, K. M., Bright, E. A., Reith, A. E., & Urban, M. L. (2020). *LandScan 2019*. Oak Ridge National Laboratory. Retrieved from https://landscan.ornl.gov/

Sanchez, G. M., Terando, A., Smith, J. W., Garcia, A. M., Wagner, C. R., & Meentemeyer, R. K. (2020). Forecasting water demand across a rapidly urbanizing region. *The Science of the Total Environment*, 730, 139050. https://doi.org/10.1016/j.scitotenv.2020.139050

Sarafian, R. (2020). Linear mixed models. In J. D. Rosenblatt (Ed.), *Introduction to data science*. Ben-Gurion University. Retrieved from https://bookdown.org/ronsarafian/IntrotoDS/book.html#problem-setup-2

Schielzeth, H., & Nakagawa, S. (2013). Nested by design: Model fitting and interpretation in a mixed model era. *Methods in Ecology and Evolution*, 4, 14–24. https://doi.org/10.1111/j.2041-210X.2012.00251.x

Seaber, P. R., Kapinos, F. P., & Knapp, G. L. (1987). *Hydrologic unit maps*. In U.S. Geological Survey. Retrieved from https://landscan.ornl.gov/

Sood, A., & Smakhtin, V. (2015). Global hydrological models: A review. *Hydrological Sciences Journal*, 60(4), 549–565. https://doi.org/10.1080/02626667.2014.950580

Soulard, C. L., Walker, J. J., & Petratski, 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/rs12090984

Steadman, R. G. (1984). A universal scale of apparent temperature. *Science*, 224(4649), 785–788. https://doi.org/10.1126/science.224.4649.785

Tulbure, 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.2013.01.010

Tulbure, 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. *The Science of the Total Environment*, 658, 1574–1585. https://doi.org/10.1016/j.scitotenv.2018.11.390

Turner, 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 of the United States of America*, 104(52), 20666–20671. https://doi.org/10.1073/pnas.0704119104

University of Virginia Weldon Cooper Center. (2018). National population projections. Demographics Research Group. Retrieved from https://demographics.coopercenter.org/national-population-projections
U.S. Army Corps of Engineers. (2020). National inventory of dams. Federal Emergency Management Agency. Retrieved from https://search.library.wisc.edu/catalog/999834648302121

U.S. Geological Survey. (2019). USGS EROS archive – Landsat – Landsat level-3 dynamic surface water extent (DSWE) science product. https://doi.org/10.5066/F7445SKQ

U.S. Geological Survey. (2021). Land change monitoring, assessment, and projection (LCMAP) collection 1.2 science products for the conterminous United States. USGS data release. https://doi.org/10.5066/P9SWW9Z2

Veldkamp, T. I. E., Wada, Y., Aerts, J. C. J. H., Döll, P., Gosling, S. N., Liu, J., et al. (2017). Water scarcity hotspots travel downstream due to human interventions in the 20th and 21st century. Nature Communications, 8, 15697. https://doi.org/10.1038/ncomms15697

Vepaska, 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/s13157-019-01183-6

Vörösmarty, C. I., 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.289.5477.284

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., et al. (2010). Global threats to human water security and river biodiversity. Nature, 467(7315), 555–561. https://doi.org/10.1038/nature09440

Wada, Y., Bierkens, M. F. P., de Roo, A., Dirmeyer, P. A., Famiglietti, J. S., Hanasaki, N., et al. (2017). Human-water interface in hydrological modelling: Current status and future directions. Hydrology and Earth System Sciences, 21(8), 4169–4193. https://doi.org/10.5194/hess-21-4169-2017

Walker, 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, 101973. https://doi.org/10.1016/j.jag.2019.101973

Wei, X., Li, Q., Zhang, M., Giles-Hansen, K., Liu, W., Fan, H., et al. (2017). Vegetation cover-another dominant factor in determining global water resources in forested regions. Global Change Biology, 24, 1–795. https://doi.org/10.1111/gcb.13983

Wickham, 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.2021.112357

Xu, P., Guo, Y., & Fu, B. (2019). Regional impacts of climate and land cover on ecosystem water retention services in the Upper Yangtze River Basin. Sustainability, 11(19), 5300. https://doi.org/10.3390/su11195300

Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S. M., et al. (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.2018.09.006

Yasarer, L. M. W., Taylor, J. M., Rigby, J. R., Locke, M. A., Antczak, E., Anna, J., & Koch, M. (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.2020.00066

Zeng, 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.2020.108001

Zhang, M., Liu, N., Harper, R., Li, Q., Liu, K., Wei, X., et al. (2017). A global review on hydrological responses to forest change across multiple spatial scales: Importance of scale, climate, forest type and hydrological regime. Journal of Hydrology, 546, 44–59. https://doi.org/10.1016/j.jhydrol.2016.12.040

Zhang, M., & Wei, X. (2021). Deforestation, forestation, and water supply. Science, 371(6533), 990–991. https://doi.org/10.1126/science.abe7821

Zhou, X., Polcher, J., & Dumas, P. (2021). Representing human water management in a land surface model using a supply/demand approach. Water Resources Research, 57, e2020WR028133. https://doi.org/10.1029/2020WR028133