

Article

Regional Multivariate Indices of Water Use Potential for the Continental United States

Jonah D. White ¹ , Elizabeth A. Mack ^{1,*}, Sharon L. Harlan ², E. Scott Krayenhoff ³, Matei Georgescu ⁴ and Kyle Redican ¹

¹ Department of Geography, Environment, and Spatial Sciences, Michigan State University, East Lansing, MI 48824, USA; white115@msu.edu (J.D.W.); redicank@msu.edu (K.R.)

² Department of Health Sciences and Department of Sociology and Anthropology, Northeastern University, Boston, MA 02115, USA; s.harlan@northeastern.edu

³ School of Environmental Sciences, University of Guelph, Guelph, Ontario N1G 2W1, Canada; skrayenh@uoguelph.ca

⁴ School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ 85281, USA; matei.georgescu@asu.edu

* Correspondence: emack@msu.edu; Tel.: +1-517-432-7058

Received: 3 February 2019; Accepted: 8 April 2019; Published: 16 April 2019



Abstract: The necessity of freshwater for sustaining human life has prompted the development of numerous estimation techniques and metrics for understanding where, when, and why water is used. While estimates are valuable, techniques for estimating water use vary, and may be difficult to replicate and/or unavailable on an annual basis or at the regional scale. To address these drawbacks, this paper proposes a series of regional indices for the continental United States that could serve as proxies for water use that are based on key variables associated with water use. Regional indices at the county level are computed, compared against each other, and compared to water withdrawal estimates from the United States Geological Survey (USGS). These comparisons highlight differences amongst the derived indices and the water withdrawal estimates. They also demonstrate promise for future development and implementation of related indices, given their similarities with water withdrawal estimates. Using only a small set of variables, these indices achieve some degree of similarity (~20%) to estimates of water withdrawals. The comparative data availability and ease of estimating these indices, as well as the ability to decompose the additive indices into their constituent use categories and constituent variables, renders them practically useful to water managers and other decision makers for identification of locally specific drivers of water use and implementation of more geographically-appropriate policies to manage scarce water resources.

Keywords: water use; water demand; multivariate analysis; water withdrawals; regional

1. Introduction

Water covers 70% of the earth, but less than 1% of this water is suitable for human consumption and food production [1]. Recent studies estimate that 4 billion people suffer from severe water scarcity at least one month per year and that half these people reside in China and India [2]. This scarcity is exacerbated by the uneven distribution of water globally [3] and widespread pollution that negatively affects human health [4]. Climate change, which has increased air temperature and modified regional precipitation patterns [5], has changed the distribution of renewable water resources [6]. Population growth has also altered the demand for renewable water resources [7]. Compounding this scarcity is the urbanization and development of countries around the globe [8]. Water consumption is linked with affluence [9] and studies highlight that per capita uses of water in developed countries are orders of magnitude greater than water use in developing nations [10].

The global scarcity of freshwater, combined with the aforementioned pressures, has attracted the attention of water scientists and engineers around the world. Their research efforts are dedicated to understanding and reducing demands on water supply/demand through technological innovation while improving management of available resources [11]. A critical first step in reducing pressures on water systems is identifying locations of higher water use and the drivers of this use. To this end, researchers have created a variety of water scarcity and water stress metrics [12–15]. Several of these indices involve global assessments using country-level data, which limits our ability to understand important variations in use and water stress at a regional scale.

From a management perspective, more localized regional metrics are critical to achieving the first step in sustainable water management, which is understanding “where, when, and how water is used to satisfy our needs” ([11], p. 20). More recently, studies have begun to address the need for assessment of regional water scarcity and development of stress indicators by analyzing water demand at finer spatial scales [16,17]. This paper builds on these efforts by deriving four multivariate indices of water use based on county-level data for the United States, which is one of the nations with the highest level of water use globally [10]. Variables used to construct our multivariate indices are grouped into three categories of use: agriculture, industry, and municipal. Once derived, the multivariate indices are compared to one another and to county estimates of water withdrawals from the United States Geological Survey (USGS). Here, it is important to note that these variables indicate *the potential* for water use since the composite indices and their components are not estimates of *actual* use but are instead based on variables from the literature that are acknowledged to be drivers of use.

While metrics for understanding water use are an ongoing area of research, the composite indices developed in this article offer two key advantages over prior efforts. First, the indices use publicly available data that are easy to collect. We illustrate techniques that can be used by others as better and newer data become available. Second, the additive indices are decomposable into their constituent components, which simplifies understanding of the individual factors driving water use at a localized level. This facilitates a comprehensive appreciation of pressures on water resources, which can assist with the formulation of better policies to reduce water use.

2. Indicators of Water Use

The United States used an estimated 1.343 trillion liters of water per day in 2010, according to the USGS, yet this level of water use represents a 13% decrease from consumption patterns observed in 2005 [18]. Reasons for this declining trend can vary by location, however, because the local economy, demographic shifts, conditions of the housing stock, and general growth patterns affect the level of water use [19]. Water used for agricultural purposes (e.g., irrigation and livestock) decreased from 488 to 446 billion liters per day, representing an 8% decline between 2005 and 2010, while industrial uses of water also declined by 8% (60.9 to 56 billion liters per day) during the same period [20]. Residential water use also declined from 670 to 522 liters per household per day (22% decline) and from 261 to 219 liters per capita per day (16% decrease) between 1999 and 2016 [21]. Despite the overall decrease in water use and consumption, it remains vitally important that we understand the composition of water use across these activities since it is likely that this trend manifests differently across geographic space with distinct regional variation.

The value of water and the need to understand pressures on water systems has prompted considerable attention in the scientific community. Water use is a multifaceted issue that involves population and demographic characteristics, natural features, and the climatic characteristics of place [6]. Although the uses of water vary by country and regions within countries [18,20] human uses may be categorized into three types: agricultural, industrial, and municipal [6,22]. To construct a more nuanced, multivariate index that characterizes the diverse range of human pressures on water resources, data were collected for each of these three domains of water use to coincide with our study year of 2010. To these three domains, climatic variables were added to account for the relationship between climatic conditions and water use [23,24]. Table 1 contains the suite of variables considered

as potential index inputs with associated sources (see also Table S1 in Supplementary Material for data dictionary). Several variables are standardized and placed in percentage terms to account for variations in establishment presence and housing types across counties of varying sizes. It is important to note that this is not an exhaustive list of all possible uses of water. Rather, this list is meant to summarize major uses of water in each of the three activity areas (agriculture, industry, and municipal) based on information derived from the literature.

Table 1. Data description and sources.

Variable Name	Description	Year and Source
Agricultural		
Irrigated Land	Irrigated agricultural land as a percent of total land area	2012 Census of Agriculture, USDA [25]
Corn Production	Corn production in tons per square mile	2012 Census of Agriculture, USDA [25]
Cotton Production	Cotton production in bales per square mile	2012 Census of Agriculture, USDA [25]
Soybean Production	Soy production in bushels per square mile	2012 Census of Agriculture, USDA [25]
Winter Wheat Production	Winter wheat production in bushels per square mile	2012 Census of Agriculture, USDA [25]
Industry		
Agricultural	Agricultural establishments as a percent of total establishments	2010 County Business Patterns, US Census Bureau [26]
Mining	Mining establishments as a percent of total establishments	2010 County Business Patterns, US Census Bureau [26]
Manufacturing	Manufacturing establishments as a percent of total establishments	2010 County Business Patterns, US Census Bureau [26]
Utilities and Energy	Energy generating establishments as a percent of total establishments	2010 County Business Patterns, US Census Bureau [26]
Residential		
4 Bedrooms or More	Percent of total housing units with 4 bedrooms or more	2008–2012 American Community Survey, IPUMS National Historical Geographical Information Systems (NHGIS) [27]
Median Home Value	Median house value (2012, USD\$)	2008–2012 American Community Survey, IPUMS National Historical Geographical Information Systems (NHGIS) [27]
Single Family Homes	Percent of total housing units that is single-family residential	2008–2012 American Community Survey, IPUMS National Historical Geographical Information Systems (NHGIS) [27]
Climate		
Precipitation	Average annual precipitation in mm	PRISM [28]
Temperature	Average annual temperature in degrees Celsius	PRISM [28]

2.1. Agricultural Water Use

While uses of water vary within and across countries, global statistics of water use indicate that agriculture accounts for 70% of water use on average [29]. Irrigation for crop growth is one of the major users of water for agriculture [30] and accordingly, studies use the amount of irrigated land as a proxy for agricultural water use [7]. Irrigation needs vary by the location and type of crop grown and accordingly, crop coefficients that represent the “soil, climatic, environmental and management factors” that impact crop water needs have been developed [31]. Thus, the crop profile of regions can help understand varying water uses.

To capture agricultural activities in counties, data were collected from the 2012 Census of Agriculture from the United States Department of Agriculture (USDA). Variables obtained from this database included the amount of irrigated land area and the production of wheat, cotton, corn, and soybeans; these crops are noted to be water intensive [30]. The Census of Agriculture is released every 5-years and, while there is a slight temporal mismatch between these data and the 2010 data collected as indicators of industrial and residential water demand, these data are useful because they illustrate where water-intensive crops are more likely to be grown.

2.2. Industrial Water Use

Industry is another activity that uses substantial amounts of water. Major industrial users of water are power generation, mining, livestock and aquaculture, and manufacturing [18,32,33]. Older manufacturing activities including those related to metals, chemicals, food, and petroleum [34] are well-known users of water, as is mining for solid minerals [18,35]. Water is used in industry as a solvent and in finished products; it is also used for washing and cooling [6].

New economy manufacturing activities oriented towards computer technology also use a lot of water. For example, manufacturing semiconductors, which requires the use of ultrapure water, can consume millions of liters of water per day in one factory alone [36,37]. Manufacturers of everyday household products such as jeans and cars also demand a large amount of water. For example, the amount of water used to produce one car requires an automotive manufacturer to use ~148,000 liters of water and to produce one pair of jeans, a textile manufacturer will use ~11,000 liters [38].

To proxy for industrial uses of water, information concerning the number of businesses by county was obtained from the 2010 County Business Patterns of the U.S. Census Bureau for agriculture, forestry, fishing, and hunting (NAICS 11), mining (NAICS 21), utilities (NAICS 22), and manufacturing industries (NAICS 31–33). Establishments in these industries are grouped by the North American Industrial Classification Industry (NAICS), which categorizes establishments by production process [39]. While there is overlap with the USDA data described in the previous section, business level data about agriculture was collected to capture farming and related activities not otherwise captured in the USDA data, including ranching, hatcheries, vegetable growing, and orchards [39]. The utilities sector includes power generation activities for multiple types of power (electric, hydroelectric, fossil fuel, natural gas, and nuclear) [39].

2.3. Residential Water Use

Population growth [7,40] and population density [41] are a popular means of characterizing human demands on water systems. In global studies of water scarcity, rising population levels increase competition for scarce water resources [42]. While this is certainly true in the developing world, in the developed world, where water saving fixtures and appliances are available, as are the means for water recycling and reuse, the link between people and water use is more complex. In this context, the characteristics of people, particularly their incomes matters, as do the characteristics of housing [9]. Studies use housing density to analyze water use because lower household densities are associated with more outdoor water use [43]. Higher housing densities present fewer opportunities for water use in this context. Thus, in the developed world, the relationship between population density or housing density and water use is complex.

From the large volume of research conducted on water use, we know that the three largest building factors affecting residential water use are type and size of the dwelling as well as the age of construction. There is widespread agreement that the size of single-family homes (square footage, number of rooms, and number of bedrooms) predicts higher water use [44–47]. In particular, single-family homes use the most water at the household level because they are correlated with larger lot size, more square footage, and household size [48]. Zhou [49] and [44] found that the number of bathrooms in a house is a contributor to water use.

Household income is the most studied sociodemographic variable related to residential water consumption and the one for which there is the highest agreement in terms of linkages with water use. As income increases in individual households, studies find total and per capita water use increase [50–54]. This is because household income is positively correlated with many other factors that are associated with higher water use, including single family residences (SFR) house size, lot size, and water appliances, so that whatever wealthier households may gain in newer plumbing, water-saving devices, or greater awareness of water conservation strategies, is more than offset by water-intensive lifestyles [9]. When income measures are not available, studies have used property values [45,55], education [56], occupation [44], and ethnicity [50] as proxies for affluence.

To account for the aforementioned factors that are associated with higher residential water use, three variables are drawn from the 2008–2012 American Community Survey of the U.S. Census Bureau compiled by the IPUMS National Historical Geographical Information System [27] where our study year of 2010 is the midpoint: the percentage of single-family homes in a county, the percentage of homes with four or more bedrooms, and median home value. The single-family home variable is designed to capture the relationship between single-family homes and greater water use [48]. The

bedroom variable is intended to capture the likelihood that a home will have more bathrooms in the absence of data availability on bathrooms in households at the county level. Median home value is used to capture the link between home value and water use [45,57].

2.4. Climatic Variables

Water demand for agriculture and urban/residential use is modulated by climate [6]. Aside from the agricultural, industrial, and residential mix within individual counties, their climatic conditions are also important to capture because studies have highlighted that seasonal climatic conditions account for substantial variation in outdoor water use [50]. Many agricultural crops and plants in urban areas (e.g., lawns) require access to sufficient soil moisture to survive and grow. Soil moisture depends principally on the balance of water inputs (precipitation, irrigation) and outputs (evapotranspiration, runoff and percolation) to the soil system [58]. In regions and seasons where precipitation exceeds evapotranspiration, plants typically have access to sufficient water without the need for irrigation (provided precipitation is not too heterogeneously distributed in time and runoff is not excessive). However, in regions where precipitation cannot supply sufficient water, irrigation may be required, and water demand is likely to be more fully coupled to thermal, moisture, and radiative climate. In many regions of the contiguous U.S., summer months represent peak irrigation requirements due to more heterogeneous precipitation, higher evapotranspiration (due to higher temperatures and solar radiative forcing), and greater plant growth (whose underlying power source, photosynthesis, consumes additional water).

Here, we focus on precipitation inputs and evapotranspiration outputs to the local water balance for irrigated locations during summer. In addition to plant characteristics and relative amounts of water in the soil and atmosphere, evapotranspiration is driven by solar radiation and is strongly modulated by air temperature (which is positively correlated with solar radiation). Therefore, we include air temperature as a proxy for evapotranspiration; a more direct approach based on the Penman-Monteith equation could employ the “MODIS/Terra Net Evapotranspiration 8-day” [59] data set available from the US Geological Survey. Alternatively, latent heat flux dynamically downscaled by a regional climate model that was evaluated against a suite of multi-scale annual and seasonal observational products across the contiguous US would provide a third option [60,61]. Broad regional patterns of air temperature in the US have a latitudinal gradient; evapotranspiration, however, presents a stronger east-west variation between the dry western and humid eastern summertime US climates. However, in our view, there is no single data set that stands out from the others, and a number of systematic errors may bias each modelled evapotranspiration product in unknown ways. Accounting for the full range of potential differences and characterizing the source of these differences is a useful endeavor but is beyond the scope of this particular paper. Nevertheless, one caveat of our approach that uses air temperature as a surrogate for evapotranspiration is that it may overestimate variation between northern and southern US counties and underestimate corresponding east-west variation.

We use the June–August-average air temperature and precipitation total from 2010 at 4 km resolution from the observationally-based Parameter-elevation Relationships on Independent Slopes Model (PRISM) data set [62], and subsequently spatially average all 4 km by 4 km precipitation and temperature data grids located within each county. The PRISM data set is based on quality-controlled data from more than 10,000 observation stations and uses spatial regression modeling that considers location, elevation, coastal proximity, topographic facet orientation, vertical atmospheric layer, topographic position, and orographic effectiveness of the terrain to assess gridded values. This product represents a notable improvement over similar data sets, particularly in mountainous terrain [62]. Importantly, because PRISM is a spatially continuous dataset, it offers key advantages for the present application that distinct meteorological stations do not provide.

3. Index Construction Methodology

To provide a multivariate characterization of demand-based pressures on water systems, the data discussed above were integrated into four indices (see Table S2 in Supplementary Materials for county-level values of indices calculated in this study). These indices are compared to estimates of water withdrawals in the year 2010 from the United States Geological Survey, a popular means of characterizing water use [63–65]. Data on water withdrawals are presented according to the broad categories of public supply, domestic, industrial, irrigation, livestock, aquaculture, mining, and thermoelectric, but we consolidate these categories into groups that more closely match those provided by the Food and Agriculture Organization (FAO) discussed later.

The methodology for constructing these indices and associated formulae are presented below. These indices are computed for U.S. counties, which is a means of administratively dividing states into smaller units. Aside from the availability of data at the county level, counties have important administrative responsibilities in the United States, including education and law enforcement [66], and thus represent units of analysis that have administrative power over people and resources. Counties are also more geographically stable units for analysis than are other spatial subdivisions in the United States that are based on population thresholds such as Census blocks or tracts, which change size and shape as the population of places expands or contracts [67]. In this study, data were compiled for 3108 counties in the continental United States.

3.1. Unweighted Decile Index

The first index developed is an unweighted sum of deciles computed from the fourteen variables that characterize agricultural, industrial and residential pressures on water resources. To accomplish this, the variables summarized in Table 1 were placed into deciles (D_i) and assigned a value of 0 through 9, where a 0 corresponds with counties that fall within the lowest 10% of all counties and a 9 indicates counties that fall within the highest 10% of all counties. In this index low values identify places with lower water use while high index values indicate higher water use. For example, placement of an observation in the lowest 10% indicates it contains a very low value relative to other observations for a given variable of interest. Alternatively, placement of an observation into the top decile indicates that observation contains a very high value for a particular variable, relative to other observations in the dataset.

In the case of average precipitation, however, this schema is reversed and counties with more precipitation are placed into lower deciles. Unlike the other variables in Table 1, higher precipitation typically reduces water use via the provision of soil moisture to agricultural and residential landscapes that would otherwise need to be provided by irrigation. Once a decile value is computed for each of the fourteen variables, the decile values are added and standardized (see Equation (1)) so the index takes on values between 0 and 100.

$$\text{Decile Index} = \left(\frac{\text{DecileSum} - \text{Minimum DecileSum}}{\text{Maximum DecileSum} - \text{Minimum DecileSum}} \right) \times 100 \quad (1)$$

where:

DecileSum is the additive total of decile ranks across all variables in each county;

Minimum DecileSum is the lowest value of the total of decile ranks across all counties;

Maximum DecileSum is the highest value of the total of decile ranks across all counties

In this version of the index, no weights are applied. To address the issue of weighting, information about the distribution of water use between sectors is incorporated into indices discussed later in this paper (see Section 3.3).

3.2. Principal Components Analysis Index

Principal components analysis (PCA) provides a means of constructing a weighted index of these same variables. PCA is a data reduction technique that produces linear combinations of the

original variables, called components, which are uncorrelated combinations of variables that explain the greatest amount of variation in the dataset. This method also determines the strength and direction of the relationship between the variables and components, or component loadings, and uses these loadings to calculate component scores, which indicate the value for each county on each component. We employed several heuristics to determine an appropriate number of components to extract and eventually rotate. First, we included a random variable in the analysis and if this variable exhibited a moderate to high loading (± 0.50 or greater), that particular component and any subsequent components were considered noise in the dataset and were removed. Next, we visually inspected the output and retained components with an eigenvalue greater than 1. Finally, we ran a parallel analysis of 1000 randomly generated datasets and identified components with an eigenvalue greater than mean eigenvalues produced from the randomly generated datasets [68].

When these criteria are applied to our dataset summarized in Table 1, the PCA model identified five components for extraction and rotation. These five components were rotated orthogonally using the Varimax method which maximizes the sum of the variances of the squared loadings and, in effect, produces either large loadings or loadings near zero with few intermediate values and simplifies interpretation of the PCA results. The five components accounted for over 59% of the total variation in the data.

Component 1 illustrates corn and soy production per square mile across U.S. counties and, to a lesser extent, the impact of average annual precipitation. Each is positively correlated with the component indicating that corn and soy production increases or decreases alongside average annual precipitation. Average annual summer temperature ($^{\circ}\text{C}$) and median house value are captured by component 2 and show an inverse relationship suggesting that as the annual average temperature increases, the median home value decreases. Wheat production per square mile is highlighted in component 3 and shows a positive relationship with the percentage of single-family housing units. Cotton production per square mile is positively correlated with the percentage of irrigated agricultural land in component 4 while agricultural establishments and energy-producing establishments as a percentage of total establishments are captured in component 5.

To incorporate some type of weighting into a composite index of water demand, we followed previous research [69,70] and used the results of the PCA analysis from Table 2 to create an unstandardized index of demand (*USID*) for each county i with the following notation:

$$USID_i = \sum P_1 f_{1i} + P_2 f_{2i} + P_3 f_{3i} \dots P_j f_{ij} \quad (2)$$

where:

P_j is the amount of variance explained by component j divided by the cumulative variance explained by the model; and,

f_{ij} is the component score on each component j for each county i .

This index was standardized per the notation below so that it can be compared to the unweighted decile index described above. This standardized version of the PCA weighted index takes on values between 0 and 100 where lower numbers correspond to counties that have fewer demands on water resources and counties with high index values have higher demands on water resources.

$$PCA\ Index = \left(\frac{County\ USID - Minimum\ USID}{Maximum\ USID - Minimum\ USID} \right) \times 100 \quad (3)$$

where:

County USID is the non-standardized value of the PCA weighted index for each county;

Minimum USID is the lowest non-standardized value of the PCA weighted index across all counties;

Maximum USID is the highest non-standardized value of the PCA weighted index across all counties.

Table 2. Results of principal components analysis (PCA).

Variables	Corn and Soy Production	Temperature & Home Value	Wheat and Single Families	Cotton and Irrigation	Establishment Type
Housing units with 4 bedrooms or more (%)	0.288	−0.463	0.416	−0.123	−0.279
Median house value	−0.224	−0.640	−0.187	0.092	−0.203
Single-family housing units (%)	0.478	−0.118	0.519	−0.213	−0.005
Agricultural establishments (%)	−0.049	−0.019	−0.161	0.167	0.749
Mining establishments (%)	−0.273	0.361	0.450	−0.185	0.140
Manufacturing establishments (%)	0.387	−0.009	−0.372	−0.256	0.251
Energy establishments (%)	−0.119	0.106	0.229	−0.017	0.631
Irrigated agricultural land (%)	0.155	−0.137	0.425	0.652	0.145
Tons of corn per square mile	0.890	−0.030	0.081	0.070	−0.093
Bales of cotton per square mile	−0.016	0.175	−0.078	0.806	0.074
Bushels of soy per square mile	0.859	0.053	0.073	0.127	−0.098
Bushels of wheat per square mile	0.052	−0.017	0.644	0.170	0.041
Average summer temperature (°C)	−0.079	0.766	−0.109	0.240	−0.100
Average annual precipitation (mm)	0.500	0.448	−0.281	−0.144	−0.209
Variance explained (%)	17.3	11.6	11.4	10.0	9.0
Cumulative variance (%)	17.3	28.8	40.3	50.3	59.3

3.3. Composite Principal Components Analysis Index

A drawback of the unweighted decile and PCA-derived indices described above is they do not reflect the variation in water use across various sectors. While there is variation in this distribution of uses across counties [22], an index that incorporates this allocation of use, is perhaps more insightful than one based on use-blind weights, such as those derived from the PCA above. To harness the analytical power of PCA and incorporate weights that reflect the distribution of water use across agriculture, industry, and households, a variant on the PCA-based index is proposed. To construct this index, three separate PCA models (Table 3) were run on the following groups of variables:

1. Agriculture: irrigated land as a percentage of total land area; tons of corn per square mile; bales of cotton per square mile; bushels of soy per square mile; and, bushels of winter wheat per square mile
2. Industrial: agricultural establishments as a percent of total establishments; mining establishments as a percent of total establishments; manufacturing establishments as a percent of total establishments; and, energy generating establishments as a percent of total establishments
3. Residential: percentage of housing units with four bedrooms or more; single-family housing units as a percent of total housing units; and, median home value

Table 3. Communalities of PCA models.

Categories and Variables	Component 1	Component 2	Communality
	Agriculture		
Irrigated agricultural land (%)	0.151	0.805	0.670
Tons of corn per square mile	0.957	0.037	0.917
Bales of cotton per square mile	−0.103	0.689	0.485
Bushels of soy per square mile	0.946	0.080	0.902
Bushels of wheat per square mile	0.085	0.546	0.306
Variance explained (%)	37.0	28.6	
Cumulative variance (%)	37.0	65.6	
	Industrial		
Agricultural establishments (%)	0.844	0.131	0.730
Mining establishments (%)	0.190	−0.576	0.368
Manufacturing establishments (%)	0.108	0.817	0.680
Energy establishments (%)	0.643	−0.348	0.534
Variance explained (%)	29.3	28.5	
Cumulative variance (%)	29.3	57.8	
	Residential		
Housing units with 4 bedrooms or more (%)	0.822	0.264	0.745
Median house value	0.005	0.965	0.931
Single-family housing units (%)	0.834	−0.241	0.754
Variance explained (%)	45.7	35.3	
Cumulative variance (%)	45.7	81.0	

These independent PCA models were also rotated orthogonally using the Varimax method to produce uncorrelated components. Communalities describe the amount of variation in each variable explained by the PCA models and these values serve as weights in our composite PCA index. Our original variables were converted to z-scores to account for differences in measurement scale and we applied the communality weights using the following set of equations:

$$\text{Agriculture} = (z_{irr} \times 0.670) + (z_{corn} \times 0.917) + (z_{cotton} \times 0.485) + (z_{soy} \times 0.902) + (z_{wheat} \times 0.306) \quad (4)$$

$$\text{Industrial} = (z_{agest} \times 0.730) + (z_{minest} \times 0.368) + (z_{manest} \times 0.680) + (z_{enest} \times 0.534) \quad (5)$$

$$\text{Residential} = (z_{4bdrms} \times 0.745) + (z_{single} \times 0.754) + (z_{mhv} \times 0.931) \quad (6)$$

Next, each group of variables was combined into a single, composite measure and additional weights assigned based on information about water use across agricultural, industrial, and municipal activities. Two sets of weights are used at this stage to reflect variations in use between the agricultural,

industrial, and municipal sectors. These two sets of weights come from information about water use in the United States from the Aquastat database of the Food and Agriculture Organization [29] and the United States Geological Survey [18]. The FAO indicates uses of 36% agriculture, 51% industrial, 13% municipal and the USGS indicates uses of 27% agriculture, 60% industrial, and 13% municipal. As such, we used the weighted totals of the variable groups above and calculated two distinct composite PCA indices as follows:

$$FAO \text{ Weighted PCA} = (Agriculture \times 0.36) + (Industrial \times 0.51) + (Residential \times 0.13) \quad (7)$$

$$USGS \text{ Weighted PCA} = (Agriculture * 0.27) + (Industrial * 0.60) + (Residential * 0.13) \quad (8)$$

Then, we standardized each of the composite PCA indices to a range of 0–100 as follows:

$$USGS \text{ or FAO Weighted Index} = \left(\frac{County \text{ CompPCA} - Minimum \text{ CompPCA}}{Maximum \text{ CompPCA} - Minimum \text{ CompPCA}} \right) \times 100 \quad (9)$$

where:

County CompPCA is the non-standardized value of the USGS or FAO weighted index for each county;

Minimum CompPCA is the lowest non-standardized value of the USGS or FAO weighted index across all counties; and,

Maximum CompPCA is the highest non-standardized value of the USGS or FAO weighted index across all counties.

It is worth noting that the information reported by these two sources is different. To align the sectoral categories reported by the FAO with the use categories reported by the USGS, the USGS use categories were mapped to the FAO categories in the following manner: municipal use contains the domestic withdrawals component of public supply; agriculture contains withdrawals for irrigation, livestock, and aquaculture; industrial withdrawals are classified as all other uses not classified as municipal or agriculture. This latter category includes water withdrawn for use in mining and thermoelectric power generation. It is important to note that, in our analysis, information about domestic use is the only component of municipal use. Municipal use is defined by the FAO as “water withdrawn for the direct use by the population” and can contain industrial and agricultural uses [22]. In the USGS database, public supply is the closest analog to municipal use and contains “water delivered to users for domestic, commercial, and industrial purposes” [71]. Since it was not possible to decompose public supply withdrawals into its constituent parts, only the domestic portion is classified as “municipal” and the remainder allocated to industrial use. This is not considered a huge assumption since “domestic deliveries represent the largest single component of public supply withdrawals” [71].

4. Results

Figure 1 presents the four strategies for characterizing potential water use. Figure 1a displays the distribution of index values for the unweighted index. Figure 1b presents the distribution of index values for the principal components derived index. Figure 1c,d present the index values for the USGS and FAO weighted principal components indices respectively. Each of these figures displays the index of interest in terms of quintiles with breaks set at the 20th, 40th, 60th, and 80th percentiles. We present our results using quintiles for two reasons: quintiles contain approximately the same number of counties in each of the five groups and they simplify visualization and interpretation by using a standardized range of values for each version of the index. Lower quintiles represent lower water use while higher quintiles indicate higher water use.

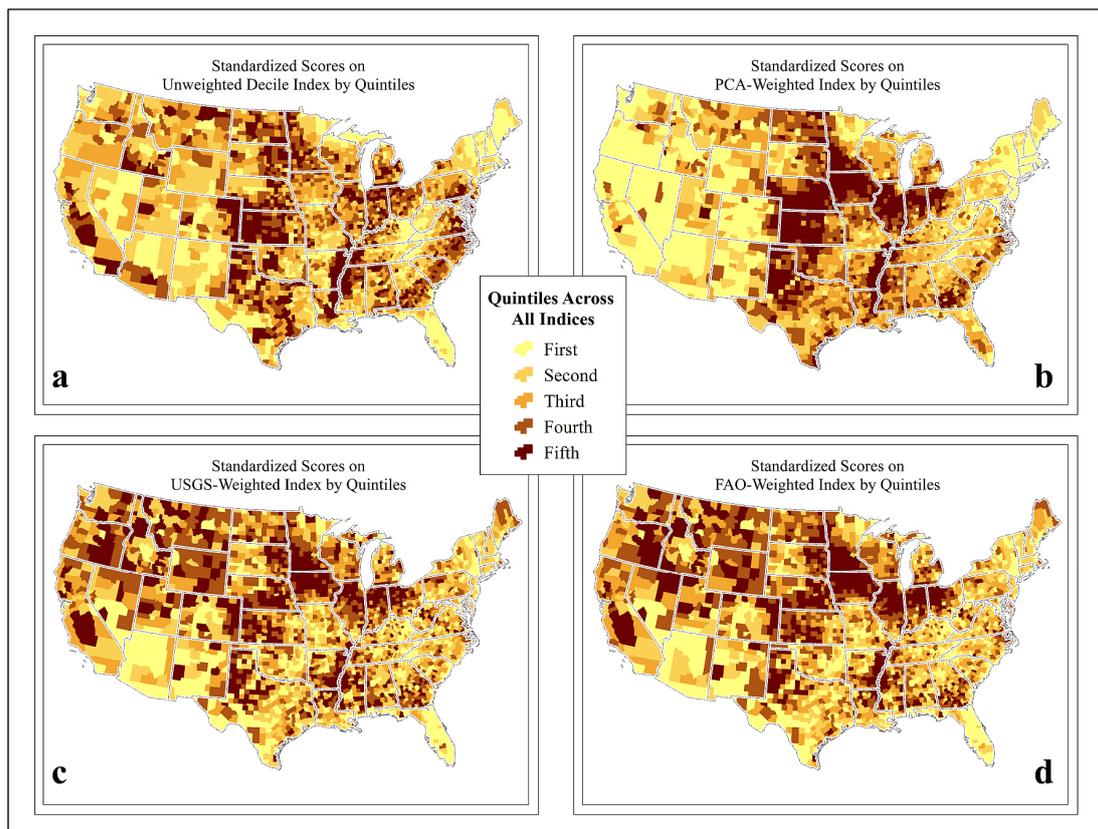


Figure 1. Comparison of indices by quintiles; (a) quintiles for unweighted decile index; (b) quintiles for PCA-weighted index; (c) quintiles for USGS-weighted index; (d) quintiles for FAO-weighted index.

4.1. Geographic Trends in Metrics

A visual comparison of the panels presented in Figure 1 highlights differences in potential water use. The unweighted decile index (Figure 1a) indicates five regions with high water use. The first region is in and adjacent to the San Joaquin River Valley in California. Counties of the Northern and Central Plains states where soybeans and corn are grown extensively also are delineated as having high water use. The industrial Rustbelt, consisting of counties located in Illinois, Indiana, Michigan and Ohio, is a third region identified as having high water use. A fourth region with high water use, the Mississippi River Valley, includes counties in the states of Mississippi, Louisiana and Arkansas which are some of the more intensive cotton producing locations in the country [72]. The southeastern seaboard is a fifth region with high water use and includes counties in Georgia, South Carolina, North Carolina and Virginia, where the apparel and automobile industries are a strong presence [73] as well as low-wage manufacturing and military defense industry [74]. On the other hand, counties along the Northeastern seaboard and those located in Appalachia are classified as having lower water use. The same is true for most counties in Florida as well as counties surrounding Seattle and Portland in the Pacific Northwest.

The PCA-weighted version of the index presented in Figure 1b illustrates several similarities and differences with the unweighted decile index. For example, this index also identifies higher water use in the North and Central Plains states but with greater concentration and intensity. The Rustbelt region and the Southeastern United States (with the exception of Appalachia) are also shown to be areas of higher water use according to the PCA-weighted index but, again, with varying intensities. Unlike the unweighted decile index, the West Coast states of Washington, Oregon, and California are shown to be areas having lower water use. An exception to this pattern is the Central Valley of California (which includes the San Joaquin River Valley). As well, the Northeast portion of the country, particularly

New York, New Jersey and Pennsylvania, is shown to be an area with lower potential water use when compared against the unweighted decile index.

This picture of water use is quite different from the picture presented by the unweighted index. The reason for these differences is that the PCA index uses the proportion of cumulative variance explained by the individual components as weights and these vary across geographic space. For instance, our first component explains 17.3% of the cumulative variance and receives greater weight when calculating the index while our fifth component receives less weight because it explains only 9% of the cumulative variance. In other words, water-intensive crops such as corn and soy, which are the highest loading variables on our first component, contribute more to this index than the agricultural and energy generating establishments that define our fifth component.

Figure 1c,d present the geography of water use, as indicated by the USGS and FAO-Weighted indices respectively. Visually, these indices are quite similar to one another and they also share several similarities with the decile index. Like the decile index, the Central Valley farming region of California is highlighted as having high water use. The agriculturally intensive Plains states are also highlighted as having higher use for water. Pockets of higher use for water are noticeable in industry-intensive states including Illinois, Indiana, Michigan and Ohio. The Mississippi River Valley and the Southeastern seaboard are also pockets of higher water use. These results reflect a more nuanced weighting schema that accounts for distinct regional variation in agricultural, industrial, and residential characteristics across US counties. For this reason, it is not surprising that farming intensive regions (e.g., Central Valley of California or the Great Plains) show high potential water use considering that agricultural activities constitute a large share of overall water use. Likewise, industrial activity that remains in the traditional manufacturing belt continues to indicate potentially high use of water.

4.2. Metric Comparison

Given the similarities and differences highlighted, a series of contingency tables were tabulated (Tables 4 and A1, Tables A2–A4) to quantify the degree to which results vary across the different indices of potential water use. The rows and columns of the tables may be interpreted as the number of counties that were classified in a particular quintile for one index but changed quintile values in another index version. The diagonals represent counties that were classified in the same quintile for both indices. Table 4 compares the quintile assignments of the unweighted decile and PCA indices. This table shows some agreement in the lower decile. The unweighted decile index classifies 9% of observations in the first quintile, as does the PCA index. There is also some agreement in the assignment of counties to the fifth quintile; both indices classify 10% of their observations in this quintile. In quintiles 2–4 there is somewhat less agreement between the indices, as indicated by the off-diagonal values. To summarize similarities in quintile classification, the values along the diagonal are summed and added to 36%. This number means that 36% of all counties were assigned to the same quintiles by each index.

Table 4. Quintile classification comparison of the decile and PCA indices.

Group	PCA (1)	PCA (2)	PCA (3)	PCA (4)	PCA (5)
Decile (1)	9%	6%	4%	1%	0%
Decile (2)	6%	6%	4%	3%	1%
Decile (3)	3%	5%	5%	5%	4%
Decile (4)	1%	3%	5%	6%	5%
Decile (5)	1%	1%	2%	5%	10%

Table 5 summarizes the similarities and differences between the four indices computed in this study. The numbers in this table were computed by summing the diagonal elements of Tables 4 and A1, Tables A2–A4. This table also presents the similarities and differences with total withdrawals data

from the USGS 2010 estimates [18]. To compute this comparison, water withdrawals were assigned to quintiles (Table A4) and compared with the quintiles for the indices derived in this study.

Table 5. Quintile classification summary.

Group	Decile	PCA	USGS PCA	FAO PCA
Decile				
PCA	36%			
USGS PCA	39%	41%		
FAO PCA	41%	45%	83%	
USGS Withdrawal	23%	18%	20%	20%

Based on the information in this table, there are some similarities between the derived indices and the quintiles derived from USGS water withdrawal data. Of these indices, the weighted PCAs are most similar to one another; they classified 83% of all counties in the same quintile. Of these two indices, the FAO-weighted PCA has the most similarities with the other indices. It is 41% similar to the unweighted decile index and 45% similar to the unweighted PCA index. A comparison of the USGS water withdrawal data and the four indices indicates that they are dissimilar from one another; the indices classify counties into similar quintiles about 20% of the time. The unweighted decile index is the most similar of the derived indices to the USGS data with about 23% of counties classified similarly.

To expand on the analysis in Table 5, Figures 2–4 depict the geographic similarities and differences between the four indices developed in this paper via difference maps. To construct these maps, the quintile of one index is subtracted from the quintile of the comparison index. For example, Figure 2a compares the decile index to the principal components derived index. If the results of this difference are positive, the decile index has higher values than does the PCA index. These locations appear in green on the map. If the results of this difference are negative, meaning that the decile index produces lower index values than the PCA index, counties are displayed in brown. Gray counties in this figure are counties that were classified in the same quintile by each index.

Figure 2 summarizes the similarities and differences between the decile index and the other three indices. A comparison with the PCA index (Figure 2a) reveals that the decile index presents a picture of higher potential use in the West and along the East coast. Alternatively, the PCA index indicates higher use in the Plains states of Iowa and Nebraska, as well as portions of Southeast states including Texas, Florida, Louisiana and Arkansas. These differences are derived from how the indices are constructed. The PCA index emphasizes agricultural and industrial variables while the decile-based index affords equal consideration to all variables including residential variables. As described in the contingency tables, there is some agreement between the two indices; counties in gray are evident in the Plains states, as well as counties in several states including Illinois, Indiana, and New York.

A comparison of the decile and USGS-weighted PCA index (Figure 2b) conveys a somewhat different picture. The USGS weighted index categorizes the Western states as having higher potential use than does the decile index. This is particularly evident in Oregon, Washington, Montana, and Wyoming. The USGS weighted index also ascribes higher potential use to Plains states such as Nebraska and Iowa than does the decile index. The decile index, which ascribes equal weighting to all variables, indicates more potential for water use along the East Coast and in states such as Nebraska, Oklahoma, and Texas. A comparison of the FAO weighted index (Figure 2c), which ascribes slightly more weight to industry than does the USGS weighting scheme, presents a similar picture of differences in potential use between the two indices. The decile index ascribes higher potential use in the Plains states and the Southeastern region of the country. These differences are largely attributable to several agricultural variables in the decile index that are capturing these regions even though their mix of crops is different.



Figure 2. Quintile comparison of decile index to other derived indices; (a) unweighted decile index to PCA-weighted index; (b) unweighted decile index to USGS-weighted index; (c) unweighted decile index to FAO-weighted index.

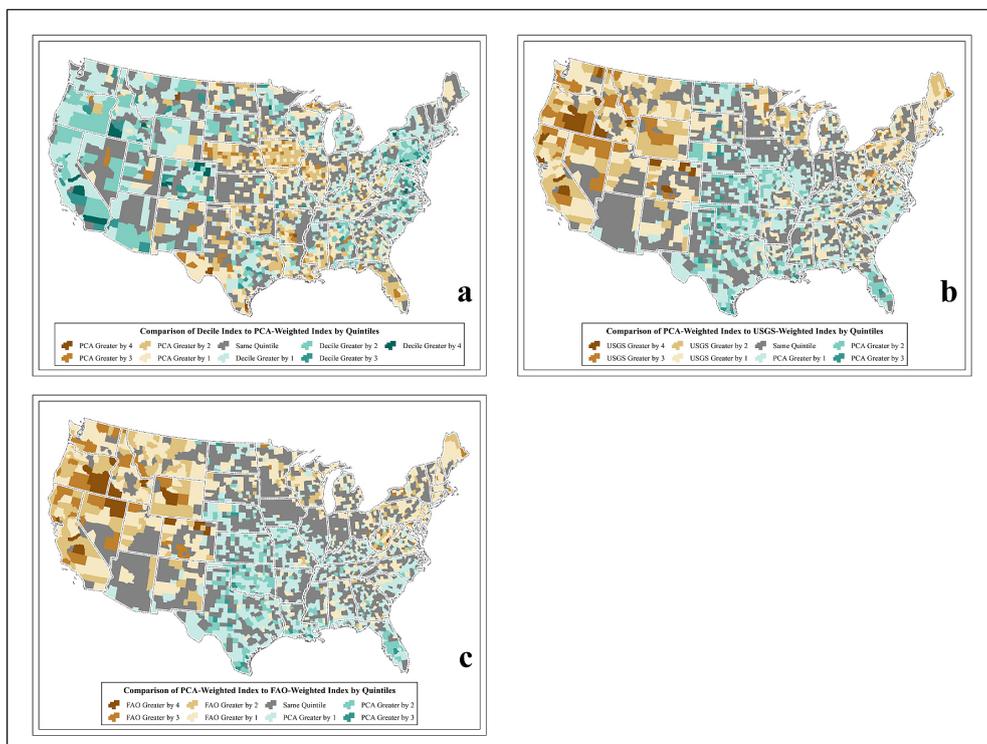


Figure 3. Quintile comparison of PCA index to other derived indices; (a) PCA-weighted index to unweighted decile index; (b) PCA-weighted index to USGS-weighted index; (c) PCA-weighted index to FAO-weighted index.

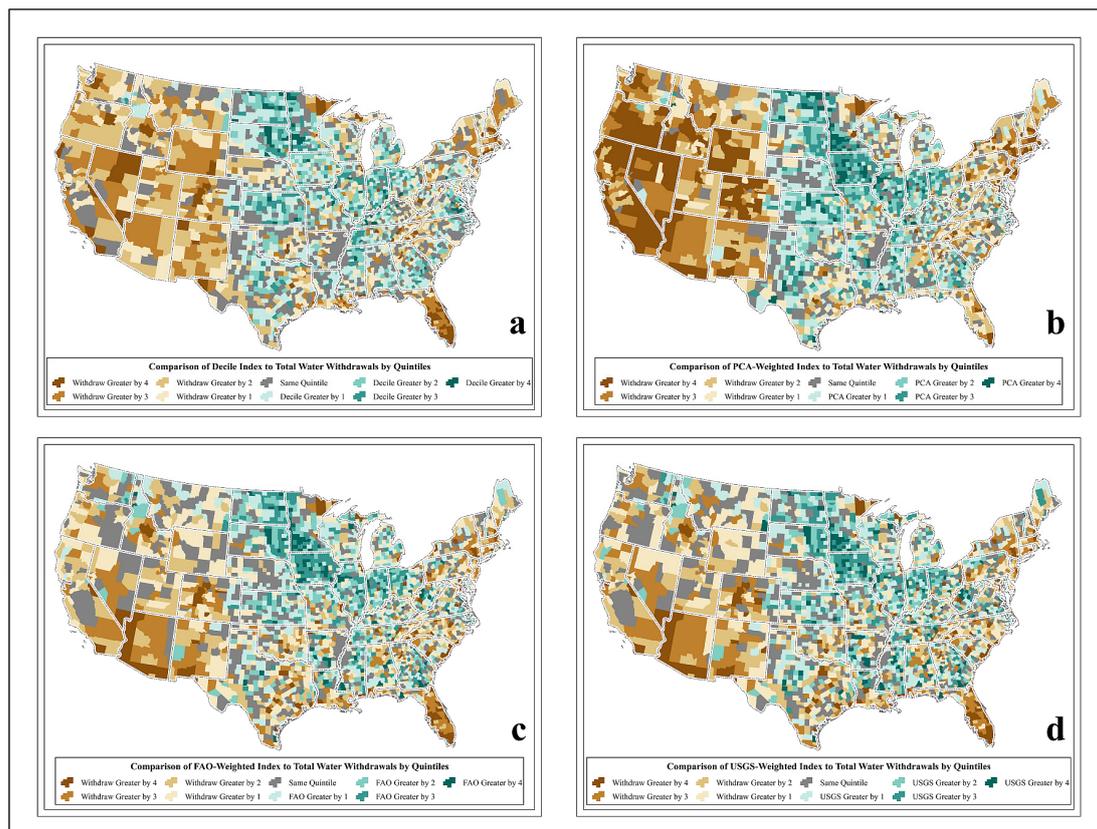


Figure 4. Quintile comparison of total water withdrawals to other derived indices; (a) total water withdrawals to unweighted decile index; (b) total water withdrawals to PCA-weighted index; (c) total water withdrawals to FAO-weighted index; (d) total water withdrawals to USGS-weighted index.

Unlike the decile index, which produced somewhat mixed patterns of similarities and differences, a comparison of the PCA index and the other three indices in Figure 3 produces more geographically pronounced trends. Figure 3a contains a difference map of the PCA and decile indices for comparative purposes only since this difference map for the two indices was discussed above. In terms of how the PCA index compares to the USGS (Figure 3b) and FAO weighted (Figure 3c) indices, the similarities and differences are quite similar. Both the USGS and FAO indices, which are weighted more towards industry, present the West and Northeast as having higher potential water uses than the PCA index. Alternatively, the PCA index classifies counties in the Plains states, the Midwest, and portions of the Southeast as having higher potential water use.

A geographic comparison of the quintile classifications highlights that those based on the USGS withdrawal data present counties in the West as having higher potential water use than do the indices derived in this paper (Figure 4). The same is true for counties in the Northeast and in Florida. These differences in classification are most stark in the comparison of the decile (Figure 4a) and PCA-derived (Figure 4b) indices to the quintiles based on withdrawal data. The differences between quintile classifications fade a bit once FAO and USGS weighting schemes are incorporated (Figure 4c,d). More gray counties, indicating classification similarities between two indices, appear in the West and Northeast. All four indices characterize the agriculturally intensive Plains states as having higher potential use than do the withdrawal data. However, these findings are made cautiously since publications by the USGS indicate the types of data used in their analysis, but the precise weighting schemes are not presented [18,75].

4.3. Index Decomposition

One of the means of understanding variations in classification results between indices—drivers of index values and how these map to quintile classifications—is the decomposition of each index into individual components. This moves beyond industry decomposition into municipal, industrial, and agricultural uses, which is possible via information from Aquastat [29], or, the decomposition of withdrawal data into different types of uses (i.e. irrigation, thermoelectric power, public supply). We use the unweighted decile index to demonstrate the value of this approach. Figure 5a–d presents a decomposition analysis of the unweighted decile index and three of its constituent variables in decile form.

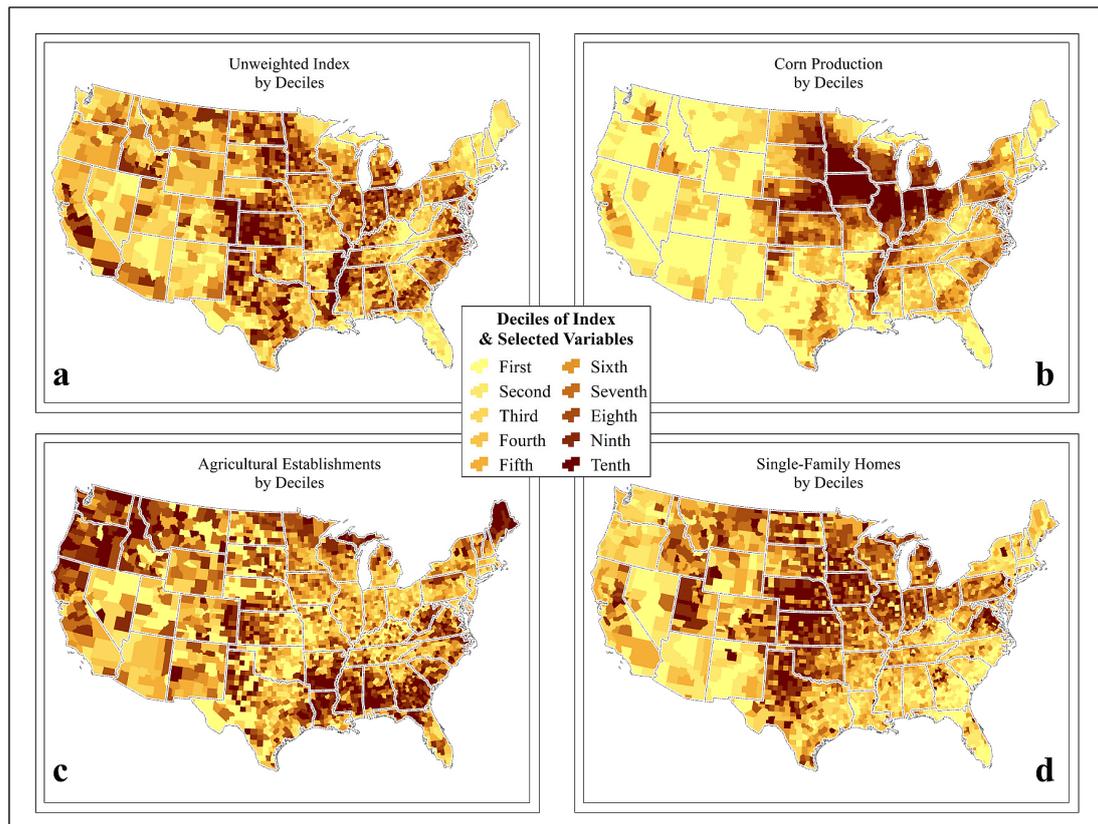


Figure 5. Decomposition analysis of decile index and leading variables; (a) unweighted decile index by decile groups; (b) tons of corn produced per square mile by decile; (c) agricultural establishments as a percent of total establishments by decile; (d) single-family homes as a percent of total housing units by decile.

The unweighted decile index takes into account each of the 14 indicators of potential water use and inherently captures geographic variability in agricultural, industrial and municipal pressures. These patterns are shown in Figure 5a where counties in the North and Central Plains, the Mississippi River Valley, the San Joaquin River Valley in California, the Rustbelt, and the Southeast are identified as areas with high potential water use. On the other hand, counties in Appalachia and the Northeast are depicted as regions with low water use.

Three variables are identified from the unweighted decile index for further examination and are used to help explain the distribution of index values in Figure 5a. These variables are selected because they show the highest component loadings in our independent PCA models presented in Section 3.3 and are displayed in Figure 5b–d. For instance, corn production (Figure 5b) is highly concentrated in counties located in the North and Central Plains states as well as Illinois, Indiana, Ohio, Michigan, Minnesota, and Wisconsin. The number of agricultural establishments portrayed in Figure 5c shows

counties with high potential water use located in the Southeast region as well as the Pacific Northwest. Single-family homes, depicted in Figure 5d, represent heightened potential water use and counties positioned in higher deciles are located in/ around the North and Central Plains, the upper Midwest and Rustbelt, and several counties in the Rocky Mountain region. The contribution of various factors to potential water use across the United States is complicated, as demonstrated by the series of indices derived in this paper, but through a decomposition analysis, it is possible to begin to understand how, why, and where potential water use varies geographically.

5. Conclusions

Given the scarcity of freshwater resources across the globe, the goal of this paper was to derive and evaluate the classification performance of indices that could serve as proxies for water use. Classification similarities between the indices were compared against one another and against water withdrawal data from the USGS, which is a frequently used proxy for water use [63–65]. This comparison revealed varying degrees of similarity between the four indices in terms of the classification of counties into quintiles. County classifications from the PCA derived indices were more similar to one another but dissimilar from those based on the water withdrawal data. The unweighted decile-based index was most similar to the water withdrawal data. From a geographic perspective, there were also significant differences in the geography of potential water use depicted by each index. These differences reflect the variations in weighting schemes from one index to another. The differences in index values between the indices presented in this study and the USGS data are likely due to differences in the composition of the indices. While the indices constructed in the study use similar data sources as the USGS water withdrawal estimates, the composition of the indices and these data are quite different from one another.

Water withdrawals include consumptive uses of water, conveyance losses (i.e., water lost in transit), as well as water returned to surface and groundwater sources [76]. Ideally, a breakout of withdrawals into each of these components would be important for understanding consumptive uses, which are not returned to water sources, and leakages from water systems. The latter is particularly important to analyze given the volume of water lost from leaks annually [77,78]. That said, this level of granularity in data at the regional level is difficult to find. In the absence of this information, related efforts to derive information about water use is an ongoing challenge that attempts to strike a balance between coverage and precision [76]. Numerous means of estimating water use are available to the scientific community ranging from input-output based techniques to multivariate regression [76]. Given this diversity of approaches, estimates of water use are better referred to as representing “potential use” instead of actual use. Thus, the indices offered in this paper represent a simpler approach to representing potential water use at the regional level and offer four advantages over prior measures of use. First, the indices make use of publicly available data that are easy to collect and integrate. Second, the additive indices are decomposable into their constituent components, which makes it easier to pinpoint the individual drivers of potential use. Third, the indices integrate a targeted yet reasonably comprehensive list of variables across agricultural, industrial and residential sectors and include data about climatic conditions. Fourth, the localized resolution is likely to be more informative within the policy realm than are aggregate measures of use.

These indices are a first step in the direction of computing robust indices of water use at the regional level. The results of this analysis suggest four areas for additional research. The first line of research is with respect to a need for work comparing existing measures of water use to one another as well as across time. A variety of strategies are used to estimate water use because of the “legal, financial, and political constraints” associated with collecting precise information [76]. Comparisons of these measures in terms of their component variables and weighting schemes (as was done in this study) are critical to understanding differences in water use presented by various metrics; this includes where measures agree or disagree and how measures change over time.

The present study examined various sectoral weighting schemes based on national averages and did not pursue regionalized weighting schemes. However, a second extension to this study is the exploration of regionalized weighting schemes that reflect regional variations in water use [20]. This type of analysis would require a detailed examination of water withdrawal estimates to design potential weighting schemes. Third, the present study examined proxies for water use, which is one element of water stress. Given the time and effort dedicated to the development of water stress indices, similar comparative studies of water stress metrics could be undertaken to understand differences and the sources of variation in these metrics.

It is also important to note that this study developed indices for water use in the United States, which is one of the larger users of water globally. As discussed previously, uses of water but also the technologies and financial ability of households to purchase and adopt these technologies to mitigate water use is perhaps greater in the developed world than the developing world. This suggests that some of the drivers of water use in the developing world may differ from those discussed in the present study. Thus, an evaluation of the utility of these indices in the developing world, and necessary modifications to these indices is a fourth recommended area for future research.

Good proxies for water use at the regional level on an annual basis are needed to help in the management of scarce freshwater resources. To date, estimates of water use are varied and often difficult to replicate. Thus, regional metrics that are possible to replicate and estimate on an annual basis are needed to track water use at finer temporal and geographic scales. This study represented one step in this direction, but it is hoped that more will follow to enhance our knowledge and management of this vital natural resource.

Supplementary Materials: The supplementary material in the Word document contains Table S1, serves as the data dictionary for this dataset and is available at <http://www.mdpi.com/2071-1050/11/8/2292/s1>. The supplementary material in the Excel document (Table S2) contains county level indices that serve as proxies for water demand as computed in this paper and is available at <http://www.mdpi.com/2071-1050/11/8/2292/s2>.

Author Contributions: J.D.W. designed and conducted the principal components analysis and calculated the indices, was responsible for the cartographic work and contributed to the writing of the article. E.A.M. originated the study, collected the data, and led the writing of the article. E.S.K. provided the PRISM data related to climatic variables and assisted with the writing of the article. S.L.H. contributed to the writing of the article. M.G. contributed to the writing of the article. K.R. contributed to the writing of the article.

Funding: This research received funding from the National Science Foundation grant number 1444758 and supplement number 1444758.

Acknowledgments: The authors wish to thank the anonymous reviewers for helpful comments and suggestions that greatly improved the content of this research.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Quintiles of PCA index compared to quintiles of other indices.

Group	PCA (1)	PCA (2)	PCA (3)	PCA (4)	PCA (5)
Decile (1)	9%	6%	4%	1%	0%
Decile (2)	6%	6%	4%	3%	1%
Decile (3)	3%	5%	5%	5%	4%
Decile (4)	1%	3%	5%	6%	5%
Decile (5)	1%	1%	2%	5%	10%
USGS PCA (1)	9%	7%	3%	1%	0%
USGS PCA (2)	6%	6%	5%	2%	0%
USGS PCA (3)	3%	4%	5%	6%	1%
USGS PCA (4)	1%	2%	5%	7%	5%
USGS PCA (5)	1%	1%	2%	4%	13%

Table A1. Cont.

Group	PCA (1)	PCA (2)	PCA (3)	PCA (4)	PCA (5)
FAO PCA (1)	9%	7%	3%	1%	0%
FAO PCA (2)	6%	7%	6%	2%	0%
FAO PCA (3)	3%	4%	6%	6%	1%
FAO PCA (4)	1%	2%	4%	8%	4%
FAO PCA (5)	1%	0%	1%	4%	14%
Withdrawal (1)	3%	4%	5%	5%	3%
Withdrawal (2)	2%	4%	4%	5%	5%
Withdrawal (3)	4%	4%	4%	4%	4%
Withdrawal (4)	5%	4%	3%	4%	4%
Withdrawal (5)	6%	4%	4%	3%	4%

Table A2. Quintiles of USGS PCA index compared to quintiles of other indices.

Group	USGS PCA (1)	USGS PCA (2)	USGS PCA (3)	USGS PCA (4)	USGS PCA (5)
Decile (1)	12%	4%	2%	1%	1%
Decile (2)	5%	6%	4%	3%	2%
Decile (3)	2%	5%	6%	5%	3%
Decile (4)	1%	3%	5%	6%	5%
Decile (5)	0%	1%	3%	5%	9%
PCA (1)	9%	6%	3%	1%	1%
PCA (2)	7%	6%	4%	2%	1%
PCA (3)	3%	5%	5%	5%	2%
PCA (4)	1%	2%	6%	7%	4%
PCA (5)	0%	0%	1%	5%	13%
FAO PCA (1)	19%	1%	0%	0%	0%
FAO PCA (2)	1%	17%	2%	0%	0%
FAO PCA (3)	0%	2%	15%	3%	0%
FAO PCA (4)	0%	0%	3%	15%	3%
FAO PCA (5)	0%	0%	0%	3%	17%
Withdrawal (1)	4%	4%	4%	4%	5%
Withdrawal (2)	3%	4%	4%	4%	5%
Withdrawal (3)	4%	4%	5%	4%	3%
Withdrawal (4)	5%	4%	4%	4%	4%
Withdrawal (5)	4%	4%	3%	4%	4%

Table A3. Quintiles of FAO PCA index compared to quintiles of other indices.

Group	FAO PCA (1)	FAO PCA (2)	FAO PCA (3)	FAO PCA (4)	FAO PCA (5)
Decile (1)	13%	5%	2%	1%	1%
Decile (2)	5%	7%	5%	3%	2%
Decile (3)	2%	5%	6%	5%	3%
Decile (4)	1%	3%	5%	6%	5%
Decile (5)	0%	1%	3%	5%	10%
PCA (1)	9%	6%	3%	1%	1%
PCA (2)	7%	7%	4%	2%	0%
PCA (3)	3%	6%	6%	4%	1%
PCA (4)	1%	2%	6%	8%	4%
PCA (5)	0%	0%	1%	4%	14%

Table A3. Cont.

Group	FAO PCA (1)	FAO PCA (2)	FAO PCA (3)	FAO PCA (4)	FAO PCA (5)
USGS PCA (1)	19%	1%	0%	0%	0%
USGS PCA (2)	1%	17%	2%	0%	0%
USGS PCA (3)	0%	2%	15%	3%	0%
USGS PCA (4)	0%	0%	3%	15%	3%
USGS PCA (5)	0%	0%	0%	3%	17%
Withdrawal (1)	4%	4%	4%	5%	4%
Withdrawal (2)	3%	4%	4%	4%	5%
Withdrawal (3)	4%	4%	4%	4%	4%
Withdrawal (4)	5%	4%	4%	4%	4%
Withdrawal (5)	4%	4%	3%	4%	4%

Table A4. Quintiles of USGS water withdrawals compared to quintiles of other indices.

Group	Withdrawal (1)	Withdrawal (2)	Withdrawal (3)	Withdrawal (4)	Withdrawal (5)
Decile (1)	6%	3%	3%	4%	3%
Decile (2)	5%	4%	4%	4%	4%
Decile (3)	4%	4%	4%	3%	4%
Decile (4)	3%	5%	4%	4%	4%
Decile (5)	2%	4%	4%	5%	4%
PCA (1)	3%	2%	4%	5%	6%
PCA (2)	4%	4%	4%	4%	4%
PCA (3)	5%	4%	4%	3%	4%
PCA (4)	5%	5%	4%	4%	3%
PCA (5)	3%	5%	4%	4%	4%
USGS PCA (1)	4%	3%	4%	5%	4%
USGS PCA (2)	4%	4%	4%	4%	4%
USGS PCA (3)	4%	4%	5%	4%	3%
USGS PCA (4)	4%	4%	4%	4%	4%
USGS PCA (5)	5%	5%	3%	4%	4%
FAO PCA (1)	4%	3%	4%	5%	4%
FAO PCA (2)	4%	4%	4%	4%	4%
FAO PCA (3)	4%	4%	4%	4%	3%
FAO PCA (4)	5%	4%	4%	4%	4%
FAO PCA (5)	4%	5%	4%	4%	4%

References

- USGS. The World's Water. U.S. Department of Interior, U.S. Geological Survey. 2016. Available online: <http://water.usgs.gov/edu/earthwherewater.html> (accessed on 9 September 2018).
- Mekonnen, M.M.; Hoekstra, A.Y. Four billion people facing severe water scarcity. *Sci. Adv.* **2016**, *2*, e1500323. [[CrossRef](#)] [[PubMed](#)]
- United Nations (UN). *Coping with Water Scarcity*; UN: New York, NY, USA, 2006.
- Schwarzenbach, R.P.; Egli, T.; Hofstetter, T.B.; Von Gunten, U.; Wehrli, B. Global Water Pollution and Human Health. *Annu. Rev. Environ. Resour.* **2010**, *35*, 109–136. [[CrossRef](#)]
- Marvel, K.; Bonfils, C. Identifying external influences on global precipitation. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19301–19306. [[CrossRef](#)] [[PubMed](#)]
- Shiklomanov, I.A. World water resources. In *A New Appraisal and Assessment for the 21st Century*; UNESCO: New York, NY, USA, 1998.
- Vörösmarty, C.; Green, P.; Salisbury, J.; Lammers, R.B. Global Global Water Resources: Vulnerability from Climate Change and Population Growth. *Science* **2000**, *289*, 284–288. [[CrossRef](#)]
- Zhang, Q.; Seto, K.C. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote. Sens. Environ.* **2011**, *115*, 2320–2329. [[CrossRef](#)]

9. Harlan, S.L.; Yabiku, S.T.; Larsen, L.; Brazel, A.J. Household Water Consumption in an Arid City: Affluence, Affordance, and Attitudes. *Soc. Nat. Resour.* **2009**, *22*, 691–709. [[CrossRef](#)]
10. Feldman, D.L. *Water*; Polity Press: Cambridge, UK, 2012.
11. Gleick, P.H. The water of the United States: Freshwater availability and use. In *A Twenty-First Century Water Policy*; Christian-Smith, J., Gliick, P.H., Eds.; Oxford University Press: New York, NY, USA, 2012.
12. Falkenmark, M. The massive water scarcity now threatening Africa: Why isn't it being addressed? *Ambio* **1989**, *18*, 112–118.
13. Gleick, P.H. Basic Water Requirements for Human Activities: Meeting Basic Needs. *Water Int.* **1996**, *21*, 83–92. [[CrossRef](#)]
14. Raskin, P.; Gleick, P.; Kirshen, P.; Pontius, G.; Strzepek, K. Comprehensive Assessment of the Freshwater Resources of the World. *Water Futures: Assessment of Long-Range Patterns and Problems*. Stockholm Environment Institute. 1997. Available online: <https://www.popline.org/node/527101> (accessed on 25 January 2019).
15. Ohisson, L. Water conflicts and social resource scarcity. *Phys. Chem. Earth B Hydrol. Oceans Atmos.* **2000**, *25*, 213–220. [[CrossRef](#)]
16. Devineni, N.; Lall, U.; Etienne, E.; Shi, D.; Xi, C. America's water risk: Current demand and climate variability. *Geophys. Res. Lett.* **2015**, *42*, 2285–2293. [[CrossRef](#)]
17. Etienne, E.; Devineni, N.; Khanbilvardi, R.; Lall, U. Development of a Demand Sensitive Drought Index and its application for agriculture over the conterminous United States. *J. Hydrol.* **2016**, *534*, 219–229. [[CrossRef](#)]
18. Maupin, M.A.; Kenny, J.F.; Hutson, S.S.; Lovelace, J.K.; Barber, N.L.; Linsey, K.S. *Estimated End Use of Water in the United States in 2010*; U.S. Department of the Interior, U.S. Geological Survey (USGS): Reston, VA, USA, 2014.
19. Coomes, P.; Rockaway, T.; Rivard, J.; Kornstein, B. *North American Residential Water Usage Trends since 1992*; Project #4031; Water Research Foundation: Denver, CO, USA, 2010.
20. Dieter, C.A.; Maupin, M.A.; Caldwell, R.R.; Harris, M.A.; Ivahnenko, T.I.; Lovelace, J.K.; Barber, N.L.; Linsey, K.S. *Estimated End Use of Water in the United States in 2015*. U.S. Department of the Interior, U.S. Geological Survey (USGS): Reston, VA, USA, 2015.
21. DeOreo, W.; Mayer, P.; Dzieglelewski, B.; Kiefer, J. *Residential End Uses of Water, Version 2*; Project #4309; Water Research Foundation: Denver, CO, USA, 2016.
22. Aquastat. Aquastat Glossary: Definition of Municipal Water Withdrawal. 2019. Available online: <http://www.fao.org/nr/water/aquastat/data/glossary/search.html?submitBtn=-1andtermId=4251> (accessed on 12 January 2019).
23. Destouni, G.; Jaramillo, F.; Prieto, C. Hydroclimatic shifts driven by human water use for food and energy production. *Nat. Clim. Chang.* **2013**, *3*, 213. [[CrossRef](#)]
24. Ray, D.K.; Gerber, J.S.; Macdonald, G.K.; West, P.C. Climate variation explains a third of global crop yield variability. *Nat. Commun.* **2015**, *6*, 5989. [[CrossRef](#)] [[PubMed](#)]
25. U.S. Department of Agriculture (USDA). *2012 Census of Agriculture*; United States Department of Agriculture, National Agricultural Statistics Service: Washington, DC, USA, 2012.
26. U.S. Census Bureau. *2010 County Business Patterns*; United States Department of Commerce, Economics and Statistics Administration: Washington, DC, USA, 2010.
27. Manson, S.; Schroeder, J.; Van Riper, D.; Ruggles, S. *IPUMS National Historical Geographic Information System: Version 13.0 [Database]*; University of Minnesota: Minneapolis, MN, USA, 2018.
28. Oregon State University. PRISM Climate Data. Available online: <http://prism.oregonstate.edu> (accessed on 13 December 2017).
29. Food and Agriculture Organization (FAO) of the United Nations. *AQUASTAT Main Database*; Food and Agriculture Organization (FAO) of the United Nations: Rome, Italy, 2016.
30. West, P.C.; Gerber, J.S.; Engstrom, P.M.; Mueller, N.D.; Brauman, K.A.; Carlson, K.M.; Cassidy, E.S.; Johnston, M.; Macdonald, G.K.; Ray, D.K.; et al. Leverage points for improving global food security and the environment. *Science* **2014**, *345*, 325–328. [[CrossRef](#)]
31. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *FAO* **1998**, *300*, D05109.

32. Renzetti, S. Economic Instruments and Canadian Industrial Water Use. *Can. Resour. J. Rev. Can. Ressour. Hydr.* **2005**, *30*, 21–30. [[CrossRef](#)]
33. Blackhurst, B.M.; Hendrickson, C.; Vidal, J.S. Direct and Indirect Water Withdrawals for U.S. Industrial Sectors. *Environ. Sci. Technol.* **2010**, *44*, 2126–2130. [[CrossRef](#)] [[PubMed](#)]
34. Kollar, K.; Brewer, R. The Impact of Industrial Water Use on Public Water Supplies. *J. Am. Work. Assoc.* **1977**, *69*, 468–473. [[CrossRef](#)]
35. Vella, H. Managing Water Consumption in Mining. Available online: <https://www.mining-technology.com/features/feature-managing-water-consumption-mining-global-shortage/> (accessed on 23 December 2019).
36. Hesseldahl, A. Taiwan’s Dry Chips. 2002. Available online: <https://www.forbes.com/2002/05/13/0513drought.html#3b6b67659a43> (accessed on 13 May 2002).
37. Lin, W.-S.; Lee, M.; Huang, Y.-C.; Den, W. Identifying water recycling strategy using multivariate statistical analysis for high-tech industries in Taiwan. *Resour. Conserv. Recycl.* **2015**, *94*, 35–42. [[CrossRef](#)]
38. Schelmetic, T. Down the Drain: Industry Water Use. 2012. Available online: <https://news.thomasnet.com/imt/2012/04/10/down-the-drain-industry-water-use> (accessed on 23 December 2018).
39. U.S. Census Bureau. North American Industrial Classification System. What is NAICS and How is it used? 2018. Available online: <https://www.census.gov/eos/www/naics/faqs/faqs.html#q1> (accessed on 1 October 2018).
40. Falkenmark, M. Rapid Population Growth and Water Scarcity: The Predicament of Tomorrow’s Africa. *Popul. Dev. Rev.* **1990**, *16*, 81–94. [[CrossRef](#)]
41. March, H.; Saurí, D. The Suburbanization of Water Scarcity in the Barcelona Metropolitan Region: Sociodemographic and Urban Changes Influencing Domestic Water Consumption. *Prof. Geogr.* **2010**, *62*, 32–45. [[CrossRef](#)]
42. Population Action International. Why Population Matters to Water Resources. 2011. Available online: <https://pai.org/wp-content/uploads/2012/04/PAI-1293-WATER-4PG.pdf> (accessed on 5 October 2018).
43. Jorgensen, B.S.; Graymore, M.; O’Toole, K. Household water use behavior: An integrated model. *J. Environ. Manag.* **2009**, *91*, 227–236. [[CrossRef](#)]
44. Clarke, G.P.; Kashti, A.; McDonald, A.; Williamson, P. Estimating Small Area Demand for Water: A New Methodology. *Environ. J.* **1997**, *11*, 186–192. [[CrossRef](#)]
45. Dandy, G.; Nguyen, T.; Davies, C. Estimating Residential Water Demand in the Presence of Free Allowances. *Land Econ.* **1997**, *73*, 125–139. [[CrossRef](#)]
46. Grafton, R.Q.; Ward, M.B.; To, H.; Kompas, T. Determinants of residential water consumption: Evidence and analysis from a 10-country household survey. *Resour. Res.* **2011**, *47*, 1–14. [[CrossRef](#)]
47. Hong, C.Y.; Chang, H. Uncovering the influence of household sociodemographic and behavioral characteristics on summer water consumption in the Portland Metropolitan Area. *Int. J. Geospat. Environ. Res.* **2014**, *1*, 2–25.
48. Dziegielewski, B.; Sharma, S.C.; Bik, T.; Margono, H.; Yang, X. *Analysis of Water Use Trends in the United States: 1950–1995—Special Report*; Southern Illinois University: Carbondale, IL, USA, 2002; p. 28.
49. Zhou, S. Three Essays on Clean Energy Technology Diffusion and Policy Innovations. Ph.D. Thesis, Georgia Institute of Technology, Atlanta, GA, USA, 2016.
50. Balling, R.C.; Gober, P.; Jones, N. Sensitivity of residential water consumption to variations in climate: An intraurban analysis of Phoenix, Arizona. *Resour. Res.* **2008**, *44*, 1–11. [[CrossRef](#)]
51. Ouyang, Y.; Wentz, E.A.; Ruddell, B.L.; Harlan, S.L. A multi-scale analysis of single-family residential water use in the Phoenix metropolitan area. *JAWRA J. Am. Water Resour. Assoc.* **2014**, *50*, 448–467. [[CrossRef](#)]
52. Hamilton, L.C. Saving water: A causal model of household conservation. *Sociol. Perspect.* **1983**, *26*, 355–374. [[CrossRef](#)]
53. Hanak, E.; Browne, M.K. Linking Housing Growth to Water Supply: New Planning Frontiers in the American West. *J. Am. Plan. Assoc.* **2006**, *72*, 154–166. [[CrossRef](#)]
54. Jorgensen, B.S.; Martin, J.F.; Pearce, M.W.; Willis, E.M. Predicting household water consumption with individual-level variables. *Environ. Behav.* **2014**, *46*, 872–897. [[CrossRef](#)]
55. Stoker, P.; Rothfeder, R. Drivers of urban water use. *Sustain. Cities Soc.* **2014**, *12*, 1–8. [[CrossRef](#)]
56. Pintar, K.D.M.; Waltner-Toews, D.; Charron, D.; Pollari, F.; Fazil, A.; McEwen, S.A.; Nesbitt, A.; Majowicz, S. Water consumption habits of a south-western Ontario community. *J. Heal.* **2009**, *7*, 276–292. [[CrossRef](#)]

57. Shearer, H. Using Geographic Information Systems to explore the determinants of household water consumption and response to the Queensland Government demand-side policy measures imposed during the drought of 2006–2008. In Proceedings of the 5th State of Australian Cities National Conference, Melbourne, Australia, 29 November–2 December 2011; Volume 29.
58. Nouri, H.; Glenn, E.P.; Beecham, S.; Boroujeni, S.C.; Sutton, P.; Alaghmand, S.; Noori, B.; Nagler, P. Comparing Three Approaches of Evapotranspiration Estimation in Mixed Urban Vegetation: Field-Based, Remote Sensing-Based and Observational-Based Methods. *Remote. Sens.* **2016**, *8*, 492. [CrossRef]
59. Running, S.; Mu, Q.; Zhao, M. *MOD16A2 MODIS/Terra Net Evapotranspiration 8-Day L4 Global 500m SIN Grid V006*; NASA EOSDIS Land Processes DAAC, Earth Resources Observation and Science (EROS) Center: Sioux Falls, SD, USA, 2017. [CrossRef]
60. Krayenhoff, E.S.; Moustou, M.; Broadbent, A.M.; Gupta, V.; Georgescu, M. Diurnal interaction between urban expansion, climate change and adaptation in US cities. *Nat. Clim. Chang.* **2018**, *8*, 1097–1103. [CrossRef]
61. Wang, M.; Wagner, M.; Miguez-Macho, G.; Kamarianakis, Y.; Mahalov, A.; Moustou, M.; Miller, J.; Vanloocke, A.; Bagley, J.E.; Bernacchi, C.J.; et al. On the long-term hydroclimatic sustainability of perennial bioenergy crop expansion over the United States. *J. Clim.* **2017**, *30*, 2535–2557. [CrossRef]
62. Daly, C.; Halbleib, M.; Smith, J.I.; Gibson, W.P.; Doggett, M.K.; Taylor, G.H.; Curtis, J.; Pasteris, P.P. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Clim.* **2008**, *28*, 2031–2064. [CrossRef]
63. Brown, T.C.; Foti, R.; Ramirez, J.A. Projected freshwater withdrawals in the United States under a changing climate. *Resour. Res.* **2013**, *49*, 1259–1276. [CrossRef]
64. Shi, D.; Devineni, N.; Lall, U.; Piñero, E. America’s Water Risk: Water Stress and Climate Variability. Columbia Water Center White Paper. 2013. Available online: http://water.columbia.edu/files/2013/09/GB_CWC_whitepaper_climate-water-stress_final.pdf (accessed on 13 January 2019).
65. Worland, S.C.; Steinschneider, S.; Hornberger, G.M. Drivers of Variability in Public-Supply Water Use Across the Contiguous United States. *Resour. Res.* **2018**, *54*, 1868–1889. [CrossRef]
66. National Geographic. Encyclopedia: County. 2018. Available online: <https://www.nationalgeographic.org/encyclopedia/county/> (accessed on 5 November 2018).
67. U.S. Census. Census Tracts. Geographic Products Branch. 2018. Available online: <https://www2.census.gov/geo/pdfs/education/CensusTracts.pdf> (accessed on 20 June 2018).
68. O’Connor, B.P. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer’s MAP test. *Behav. Res. Methods Instrum. Comput.* **2000**, *32*, 396–402. [CrossRef]
69. Antony, G.M.; Rao, K.V. A composite index to explain variations in poverty, health, nutritional status and standard of living: Use of multivariate statistical methods. *Public Health* **2007**, *121*, 578–587. [CrossRef]
70. Krishnan, V. Constructing an area-based socioeconomic index: A principal components analysis approach. In Edmonton, Alberta: Early Childhood Development Mapping Project. 2010. Available online: https://pdfs.semanticscholar.org/b115/2cdc9e9ca217434db492a9e1c4ba519640ce.pdf?_ga=2.150972726.1288394831.1555074745-945321398.1555074745 (accessed on 9 April 2018).
71. USGS. Water Use in the United States: Public Supply Water Use. 2018. Available online: <https://water.usgs.gov/watuse/wups.html> (accessed on 12 January 2019).
72. National Cotton Council. Frequently Asked Questions: Where is cotton grown in the U.S.? 2019. Available online: <http://www.cotton.org/edu/faq/> (accessed on 3 December 2018).
73. Feser, E.; Sweeney, S.; Renski, H. A Descriptive Analysis of Discrete U.S. Industrial Complexes. *J. Reg. Sci.* **2005**, *45*, 395–419. [CrossRef]
74. Glickman, N.J.; Glasmeier, A.K. The international economy and the American South. In *Deindustrialization and Regional Economic Transformation*; Routledge: London, UK, 2017; pp. 60–79.
75. Hutson, S.S.; USGS. Guidelines for Preparation of State Water-Use Estimates for 2005. U.S. Department of the Interior, U.S. Geological Survey. Techniques and Methods Book 4, Chapter E1; 2005. Available online: <https://pubs.usgs.gov/tm/2007/tm4e1/pdf/tm4-e1.pdf> (accessed on 3 December 2018).
76. National Research Council. *Estimating Water Use in the United States*; The National Academies Press: Washington, DC, USA, 2002.

77. Carrington, D. Water Companies Losing Vast Amounts through Leakage, as Drought Fears Rise. Available online: <https://www.theguardian.com/environment/2017/may/11/water-companies-losing-vast-amounts-through-leakage-raising-drought-fears> (accessed on 11 May 2017).
78. Schaper, D. As Infrastructure Crumbles, Trillions of Gallons of Water Lost. NPR. Available online: <https://www.npr.org/2014/10/29/359875321/as-infrastructure-crumbles-trillions-of-gallons-of-water-lost> (accessed on 29 October 2014).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).