STATISTICALLY EVALUATING WATER CONSUMPTION HISTORICALLY AND ACROSS MULTIPLE USERS IN VIRGINIA

Morgan Faye DiCarlo

Thesis submitted to the Faculty of Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Biological Systems Engineering

Julie E. Shortridge
Venkataramana R. Sridhar
Andrew Ellis

May 9, 2018
Blacksburg, VA

Key Words: Water Withdrawals, Trend Assessment, Water Systems Modeling
STATISTICALLY EVALUATING WATER CONSUMPTION HISTORICALLY AND ACROSS MULTIPLE USERS IN VIRGINIA

Morgan Faye DiCarlo

ABSTRACT

This study explores key aspects of water usage in Virginia via a broad-scale analysis of multiple water users through thirty years of time-series records from the Virginia Water Use Data System. A full spectrum of users is considered, including water used for energy, industrial, agricultural and municipal applications. The extent of the relationship between the volume of water used and drivers like economic and climatic conditions are not well defined in humid environments like Virginia. Mann-Kendall testing is applied to identify water use trends through time both statewide and at the county level. A panel regression is employed to identify relationships between water use and explanatory variables of climatic and economic conditions, both spatially and temporally. Key trends include that industrial and energy sector water withdrawals per facility are significantly decreasing over time. Water used for agricultural applications was found to increase on warmer than average years and decrease on wetter than average years, indicating the panel regression methodology successfully demonstrated and quantified intuitive trends. Interestingly, municipal and industrial water usage had a statistically significant relationship with the GINI coefficient, or more unequal rainfall distribution, indicating intraseasonal variability may play an important role in water use trends that is not apparent using seasonal averages alone. Overall, this work contributes to the understanding of water use trends at the state level for Virginia, and better characterizes long-term trends and short-term variability in water withdrawal.
STATISTICALLY EVALUATING WATER CONSUMPTION HISTORICALLY AND ACROSS MULTIPLE USERS IN VIRGINIA

Morgan Faye DiCarlo

GENERAL AUDIENCE ABSTRACT

This work applies statistical methods to better understand water use trends through time and across the state of Virginia. The primary data source is a record spanning thirty years of water use, reported by more than 2,400 users in all counties of Virginia to the Department of Environmental Quality (VDEQ). This analysis includes a full spectrum of water sectors, including water used domestically (municipal), water used for manufacturing (industrial), agriculture applications and water used in the production of energy. The first objective is to determine if water use, normalized by changes in population and the number of users, is increasing or decreasing over time for each county in Virginia.

Once the trends through time are identified, the next objective is to better define the underlying factors (explanatory variables) which may drive changes in how much water is used. One potential factor includes changes in the economic conditions. For example, the economic recession in 2008 caused some decline in industrial production. Did this likewise cause a reduction in water used by industrial facilities? Particularly, the analysis considers how the annual average temperature, total annual precipitation, rainfall variability and the length of heatwaves that occur in a given year might impact the amount of water withdrawn in that year. This work addresses a knowledge gap about how water use is impacted by climate change in humid environments like Virginia.

This work aims to establish whether or not there is a significant relationship between time, climate, economic change and water use in Virginia. The trends identified in this study will support the management of water supplies in Virginia and the development a more informed state water resources plan.
ACKNOWLEDGEMENTS

With many thanks to my advisor, Dr. Julie Shortridge, for the incredible opportunity to pursue research at Virginia Tech. Facilitating me through a mid-year switch into her research group gave me a chance to pursue something I am really passionate about, for which I will always be grateful, and I have learned so much under her guidance. Thanks also to my committee members, Dr. Venkat Sridhar and Dr. Andrew Ellis, for their time and inputs toward the successful completion of this project. Many thanks to our collaborators with the Virginia Department of Environmental Quality, Scott Kudlas and Rob Bergholzer, for providing resources and insight.

To the entire Water Systems lab group, including Julia Reis, Mitchell Paoletti and Morgan (II) McCarthy, thank you for your support and the opportunities to present and practice. Likewise, I would not be here without the support of my friends and loved ones, including Kyle Jacobs, Suraj Gupta, Lauren Eastes, Ellie Weiner, Mariah Gnety, Elyce Buell, Eleftheria Agioutanti and Zoe Schmitt. Finally, thank you to my parents and sister, for absolutely everything.
TABLE OF CONTENTS

TABLE OF CONTENTS ........................................................................................................................................ v

LIST OF FIGURES ........................................................................................................................................... vii

LIST OF TABLES ................................................................................................................................................ vii

1. INTRODUCTION ........................................................................................................................................... 1

2. LITERATURE REVIEW ................................................................................................................................. 3
   2.1 Water Use Conditions in the United States ............................................................................................ 3
   2.2 Water Use By Sector .................................................................................................................................. 5
   2.3 State of Virginia’s Water Resources ..................................................................................................... 6
   2.4 Underlying Drivers of Water Use .......................................................................................................... 7
      2.4.1 Population ........................................................................................................................................ 7
      2.4.2 Climate ............................................................................................................................................. 8
      2.4.3 Economics ...................................................................................................................................... 10
   2.5 Analytical Approaches ............................................................................................................................ 10

3. OBJECTIVES ............................................................................................................................................... 12

4. METHODOLOGY ......................................................................................................................................... 12
   4.1 Data Sources and Processing ............................................................................................................... 12
      4.1.1 Virginia Water Use Data System (VWUDS) ............................................................................... 12
      4.1.2 Bureau of Economic Analysis (BEA) ....................................................................................... 14
      4.1.3 PRISM .......................................................................................................................................... 14
   4.2 Trend Assessment ..................................................................................................................................... 16
   4.3 Regression ............................................................................................................................................. 17

5. RESULTS AND DISCUSSION .................................................................................................................... 21
   5.1 Time Trends .......................................................................................................................................... 21
5.1.1 Statewide Trends ........................................................................................................21
5.1.2 County Level Trends ...............................................................................................23
5.2 Regression Model Structure and Accuracy .................................................................25
5.3 Regression ..................................................................................................................27
6. FUTURE WORK AND LIMITATIONS ..............................................................................32
7. CONCLUSIONS..............................................................................................................33
REFERENCES....................................................................................................................34
APPENDIX A- Data and Code Repository .........................................................................37
LIST OF FIGURES

Figure 1: Total Water withdrawals by use category in 2016 and averaged 2012-2016 .......................................................... 5
Figure 2: Actual Water Use versus Predicted Demand, VDEQ ........................................................... .7
Figure 3: ArcGIS ModelBuilder Flow for PRISM Data Processing ....................................................... .14
Figure 4: Histogram of log-transformed Water Use ............................................................................. .18
Figure 5: NOAA Climate Divisions of Virginia .................................................................................. 19
Figure 6: Statewide Water Use Trends ............................................................................................. 23
Figure 7: County Water Use Time Trends ........................................................................................... 24
Figure 8: Form II- Predicted vs Observed Water Use ...................................................................... 27
Figure 9: Regression estimates ........................................................................................................ 31

LIST OF TABLES

Table 1: Number of Counties and facilities in VWUDS by Sector ............................................................. 13
Table 2: Climate Statistics per Virginia NOAA Division ..................................................................... 19
Table 3: Tested Model Formulations .................................................................................................. 21
Table 4: Statewide Water Use Trends ................................................................................................ 22
Table 5: Agricultural Regression Model Fit .......................................................................................... 25
Table 6: Energy Regression Model Fit .................................................................................................. 25
Table 7: Industrial Regression Model Fit ............................................................................................. 25
Table 8: Municipal Regression Model Fit ............................................................................................ 25
Table 9: Water Use per Agricultural Facility: Regression Results ...................................................... 28
Table 10: Water Use per Energy Facility: Regression Results ............................................................ 29
Table 11: Industrial Water Use per Capita: Regression Results .......................................................... 29
Table 12: Municipal Water Use per Capita: Regression Results .......................................................... 30
1. INTRODUCTION

Virginia boasts plentiful water resources, however, the need for more stringent water management is rising as population increases and climate change grows more pronounced. Making reliable water use decisions that are relevant statewide is difficult because of wide variations in water use by sector and county.

The existing body of literature on water use in the United States focuses largely on problems in the arid southwest and California (Balling et al., 2008; Gutzler and Nims, 2005; Lee et al., 2015). Water use has previously been shown to respond to seasonal variabilities in weather, but the extent of this relationship is not well understood, particularly in humid climates like Virginia. In the short term, high demand during dry, hot periods increases the risk of scarcity. In the long term, precipitation events that are trending toward less frequent, more intensive occurrences, among other climate changes, will introduce more variability and uncertainty to the water supply (Orrock, 2016). A better understanding of the relationship between water use and climate is needed for both short-term water management, in the event of dry spells and drought, and for long term planning under climate change. Additionally, current estimates for Virginia forecast a 32% increase in water demand, which seems unrealistic given the demand for water has decreased over the last decade (Figure 2: VDEQ, 2015). Better characterization of trends and variability through time is needed as historical water use can help frame baseline predictions of future water demand.

There are many potential underlying factors that may influence water use through time. There is a consensus in existing literature that water use is statistically related to climate variations in urban environments (Polebitski et al., 2011), but that the degree of sensitivity varies in different regions across the country (Lee et al., 2015). However, there is insufficient information on the relationship between climate and the full spectrum of water users, including the agricultural and industrial sectors. Consideration of the full range of water use types is essential for basin-level planning. Previous statistical analyses on water use are typically either cross-sectional, considering multiple locations, or time-series, based on years of historical data. Avoiding bias from omitting potential explanatory variables and identifying broadly applicable trends requires robust statistical assessment. Rarely has a water use study integrated both spatial and temporal data, conducting an assessment through time and across multiple sites, and this study is the first to do so for Virginia’s water resources.
This study inquires if water withdrawals are correlated with climatic and economic changes in a realistic, full spectrum of users. Water use is separated into four use sectors, agricultural, energy, industrial and municipal, which function differently and therefore are expected to respond uniquely to explanatory factors. This study also aims to identify the extent of variations among different sectors, e.g. if agricultural users are more sensitive to climatic changes than other sectors. A panel regression is applied to assess multiple users through time.

This study will advance the existing literature to better address conditions in Virginia in three key areas. First, this study will analyze at scales greater than the typical household/census tract level, where water demand is influenced by different factors. Secondly, where the current body of knowledge focuses on the urban sector, this project includes all water user types, including industrial and agricultural. Finally, the statistical method of panel regression implemented as a more robust way of identifying influencing factors than cross sectional analyses and finding more generalizable relationships than a time-series analysis.
2. LITERATURE REVIEW

2.1 Water Use Conditions in the United States

Climate variations, increasing demand and more frequent weather extremes place mounting stress on water resource systems globally. Overconsumption of water is correlated to economic growth, an alarming trend that could place half the world in water scarcity by 2030 (Chika Urama et al., 2016). The United States faces many water management issues such as drought, aging infrastructure and rapid urbanization. Additionally, the United States has the highest consumptive use of water per capita in the world, averaging around 89 gallons per person per day in 2010 (Maupin et al., 2014). Consumptive use is that which is removed from the supply without return to the same source system. Many water use projections employ a simplistic statistical means to estimate water use based on population projections, and thereby indicate water demand is steadily on the rise (UNEP, 1998). Yet, about 355,000 million gallons per day of water was used in the United States during 2010, a substantial thirteen percent decline from 2005 (US Census Bureau, 2017). There is little consensus around the patterns of water use over time, and the significance of water use influencers such as climate variability and economic growth.

In the United States, the focus of water supply research lies west of the 100th meridian due to historically arid climate conditions. From 2012 to 2014, California faced the most exceptional water stress conditions compared to the last century and severe drought persists in Southern California today (Griffin and Anchukaitis, 2014). Subsequent wildfires have plagued the region and overdrawing from groundwater sources is a direct threat to ecological stability. High water footprint crops, like cow feed, alfalfa and almonds, contribute to California’s status as the largest agricultural producer in the U.S. and stresses the already limited water supply (USDA, 2017). Also in the Southwest, climate projections show Phoenix, Arizona, will become warmer and drier over the next decade and that population is on the rise (Balling et al., 2008).

In humid regions, water is comparatively abundant. On the eastern seaboard, Virginia is serviced by nine different river systems, stores of surface and groundwater and rainy seasons. Despite being a relatively wet region, the east coast is not without water resource pressures, and these demands are mounting. Population growth, increasing severity of seasonal droughts and consumptive industrial usage are potential water management challenges facing the eastern United States. Virginia’s water resource management plan lacks specificity and actionable tactics to address these developing issues (Orrock, 2016).
The existing legal and social systems of water rights in the eastern U.S. are structured around perceived water abundance, and often fall short when scarcity issues arise (Caccese, 2016). Surface water is governed by the riparian doctrine, where property owners bordering a water source have withdrawal access. When streamflow is low, the potential for conflict emerges. Also, there are not clearly defined stipulations prioritizing water for human consumption purposes during times of shortages. Many of the fastest growing cities in the United States are in the east and southeast, further contributing to water supply strain. Charlotte, North Carolina, Jacksonville, Florida and Washington, DC are among the fifteen cities with the fastest numerical increases in population (US Census Bureau, 2017). Virginia will be the tenth largest state by 2045, with population projections surpassing ten million (WCC-UVA).

Additionally, the influx of people is associated with changes in industrial activity. Some industries are shifting toward more water conservative practices, and power production facilities are beginning to reuse the water used for cooling purposes to cut down on withdrawals. However, many of the types of manufacturing processes in Virginia are still highly water consumptive. The largest industrial user in Virginia is Hopewell Plant manufacturing, one of the world’s largest production facilities for nylon polymer components and ammonium sulfate fertilizers (VDEQ, 2015). Other hefty consumers include papermills, mining quarries and chemical development plants (VDEQ, 2015). Energy production is an additional strain on the water supply, with the largest water use sector in Virginia being power facilities (VDEQ, 2015).

Eastern states are not immune to seasonal dry spells, and the risk of drought conditions grows as the buffer between water availability and water consumption in urban areas narrows (Sohn 2011). In 2002, a drought placed several large domestic systems, including Charlottesville, Virginia, within 60 days of failure, meaning these water districts were less than two months from running out of water for human consumption (VDEQ, 2015). An eastern water predicament is imminent, with states like Pennsylvania experiencing more frequent drought emergencies in the last three decades than in the previous three centuries (Caccese, 2016). With ligation and shortages on the rise, more informed resource management is needed to prevent a west-coast level water crisis occurring in the east.
2.2 Water Use By Sector

The United States Geological Survey (USGS) defines eight categories of water use: thermoelectric power, irrigation, public supply, manufacturing, aquaculture, livestock, mining, and domestic. The power-energy domain accounts for forty-five percent of water use in the United States, largely withdrawn from saline sources (Maupin et al., 2014). Agricultural purposes like irrigation, aquaculture and livestock account for thirty-seven percent of water use across the United States. The western U.S. usage is dominated by the agricultural sector. Municipal needs account for thirteen percent of the water used. Industry, including the mining and industrial categories, attributed to five percent of the total water usage in the United States (Maupin et al., 2014).

In Virginia, the breakdown of water use by sector differs from the western US, with domestic/public water needs attributing to approximately sixty-two percent of total water withdrawals, followed by industry (30%), agriculture (3%) and remaining sources as seen in Figure 1 (VDEQ, 2017).

![Figure 1: Total Water withdrawals by use category in 2016 and averaged 2012-2016. Figure provided in the VDEQ 2017 Water Resources report](image)

Virginia’s water use is not primarily agriculturally driven, and therefore is fundamentally different from studies that have observed water supply in the west. There is a need for better understanding of sectors beyond agricultural, including industrial and municipal, as consideration of water use across all sectors is requisite for accurate basin-level planning.
2.3 State of Virginia’s Water Resources

The Virginia Water Use Data System (VWUDS) has records of over thirty years of sectoral water use and is useful to identify patterns and variables explaining water use. Factors underlying water use behaviors include population, land use changes, mean annual temperature, growing season precipitation and income (Melillo et al., 2014). Virginia’s water supply is plentiful, with natural surface and groundwater sources and averaging more than forty inches of rainfall annually (Daly, 2008). Yet, this state faces unique management challenges. According to the Virginia Department of Environmental Quality, the state lacks data for approximately half of surface water users. Other users are not obliged to obtain a permit under grandfather land clauses and other exemptions (VDEQ, 2015). Another challenge is the wide variations in use by county, since the state spans from extremely rural areas to urban centers such as Arlington (Orrock, 2016). Additionally, there is the ‘crowding out’ phenomenon, where industrial demand substantially affects the water availability for other sectors. This phenomenon is problematic under drought conditions, as Virginia has no legal restrictions on water prioritization for domestic needs, and industry commandeers high quality, low cost sources (Orrock, 2016).

Virginia has increasing water system challenges from changing climate patterns and variations in seasonal patterns. The drought of 1999-2002 in Virginia exemplifies the capacity to improve state planning. In the summer of 2002, wildfire occurrences were at record-breaking high and streamflow reached an all-time low (VDEQ, 2015). More recently, in September 2015, low precipitation in the winter season led to dry summer conditions such that the Virginia section of the Roanoke River basin was placed under Drought Watch (VDEQ, 2015).

Estimates for Virginia forecast a thirty-two percent increase in water demand from 2010 to 2040, which seems unrealistic given the actual demand for water has decreased over the last decade (Orrock, 2016). The disparity between observed and predicted water usage in the state of Virginia is the central research question for this analysis (Figure 2).
Figure 2: Actual Water Use versus Predicted Demand, from the VDEQ 2015 Water Resources Plan. The dashed red line is expected water demand increase, based on population growth estimate. The blue line indicates actual observed water use.

There is an important distinction between water withdrawn then reintroduced into the system versus water which is consumed. About sixty-two percent of consumptive water withdrawals in Virginia are for human use, primarily by municipalities or regional water authorities (Orrock, 2016). Industrial usage for manufacturing accounted for about 31 percent of consumptive withdrawals, with the largest withdrawals made by paper and chemical manufacturers. The remainder of withdrawals were for agriculture, irrigation, commercial, and mining activities. Future supply estimates in the state water resources plan do not fully account for the return of water to the water supply.

2.4 Underlying Drivers of Water Use

2.4.1 Population

It is well-established that as the population grows, the municipal demand for water will rise. According to the 2010 census, Virginia’s population surpassed eight million, an increase of roughly one million people since 2000 (Census Bureau, 2017). The population continues to expand, with the US Census Bureau expecting six percent growth into 2017 (Census Bureau, 2017). This will
undoubtedly place some additional stress on the water supply. However, the extent of the relationship between water use and population in Virginia is not well known. Many models project water demand simply as an expansion of population (UNEP, 1998). Little work has been completed to create a refined portrayal of water demand, including quantifying demand management initiatives in their capacity to decouple water use from population rise. For example, water planners have limited information on how much new conservation initiatives, pricing structures or water efficient appliances have changed the municipal demand. Considering the disparity in Figure 2, these techniques may in fact be contributing to lower than expected water use (VDEQ, 2015). There is a substantial body of research forecasting domestic demand, particularly in urban areas (Lee et al., 2015; Sohn 2011). Based on population projections alone, the overall water withdrawals in the U.S. will grow by three percent over the next five decades (Melillo et al., 2014).

There is research demonstrating the expected trend that municipal per capita use will decline as water efficient technologies become more commonplace. Urban counties and higher income classes are correlated with increased water efficiency technologies (Sankarasubramanian et al., 2017). For example, the California Water Foundation funded a behavioral water conservation study with a tool called WaterSmart, which allowed residents to input usage data and then displayed comparisons with others in the vicinity. The pilot showed when participants compared their water consumption to neighborhood averages, their individual usage usually decreased by five percent (CWF, 2014). This type of behavioral study has not been replicated in Virginia. It is generally expected that domestic water use will decline as efficient appliances become more readily available, pricing structures change and conservation practices are promoted. However, there is not yet empirical evidence to confirm this in humid regions.

2.4.2 Climate

Previous studies suggest that water use in the municipal and agricultural sectors is influenced by climate (e.g. Polebitski and Palmer, 2010; Balling et al., 2008). Overall shifts, such as an annual average two-inch increase in precipitation over the continental U.S. in the last century, as well as intraseasonal changes, like increasing lengths of dry spells, influence the availability of water (Melillo et al., 2014). Industrial and energy water usage is potentially sensitive to temperature, precipitation and intraseasonal variations as well. When applying climate projections to historical USGS data, the demand for water withdrawals is expected to increase by an average
of thirty percent by 2060, with ninety percent of the continental U.S. facing a demand increase (Brown et al., 2013; Melillo et al., 2014). With climate change considerations included, Virginia could see an increase in water demand of up to twenty-five percent in the next forty years (Brown et al., 2013). These projections were determined via coupling a water supply model with a demand model, based on population estimates and historical water efficiency trends, and assessing model outputs as driven by general climate model (GCM) scenarios (Brown et al., 2013).

The vast majority of research pertaining to water use and climate focuses on dry, arid regions. Evapotranspiration rates, directly related to annual temperature, was the most significant influencer on water used for the agricultural sector in arid regions (Brown et al., 2013). Residential water consumption in Phoenix is significantly related to variations in temperature and precipitation (Balling et al., 2008). Similarly, most variations in the summer water demand in Albuquerque, New Mexico were attributed to changes in climate, with precipitation being the strongest influencer (Gutzler and Nims, 2005). The impacts of climate variability on water use in other climatic conditions besides dry environments are not well defined. Water use was more sensitive to maximum temperature, precipitation levels and vegetation cover in Phoenix than in Portland, a preliminary indication that humid environments respond differently to water stressors than arid environments (Lee et al., 2015). Further investigation is needed as to the role of climate in the eastern states’ water use patterns, and the extent to which water use is changing in response to the changing climate. Mean annual temperature, evapotranspiration rates and intensity of long-term droughts are all markedly increasing over time in the Southeastern United States, and such some changes in water use are anticipated (Melillo et al., 2014).

In addition to precipitation and temperature on an annual scale, climate variability from season to season influences water usage. Even if the overall temperature balances to a normal average for the year, a short, consecutive period of high temperatures could potentially damage summer crops. Heatwaves may cause a spike response in agricultural water usage or more industrial demand for cooling purposes. The distribution of precipitation over the course of a year could also play a significant role in water demand (Rajah et al., 2014). Since population, mean annual temperature, and intraseasonal variations are all on the rise in the Southeastern U.S., the likelihood that the water resources of Virginia are vulnerable to climate change impacts is high (Melillo et al., 2014).
2.4.3 Economics

Prosperous economies are linked with a higher water usage footprint, with the gross domestic product of a country and water consumption being directly related (Chika Urama et al., 2016). Water efficient technologies, including root-targeted irrigation and less water intensive manufacturing processes, are becoming more affordable and therefore are more accessible to every sector (Melillo et al., 2014). Developing a better understanding of how water use is correlated with economic conditions benefits demand management. Periods of economic troughs that trigger declines in industrial production could likewise reduce industrial water consumption. In Oregon, properties with higher physical value, including amenities like lawns, gardens and pools, were found to have significantly higher water use when compared with less affluent neighborhoods (House-Peters et al., 2010). In Phoenix, Arizona, more affluent census tracts are much more sensitive to drought evidencing interplay between the factors of climate and economics on influencing water demand, (Balling et al., 2008). While this spatial correlation between high-income areas and high water use has been well-established, potential temporal correlations (e.g., does water use increase in economically productive years) have not been evaluated.

2.5 Analytical Approaches

Most prior studies have assessed water use/climate/economic relationships at the household or census tract level. Few studies have done statewide assessments, yet greater scales are where water demand is most likely to be influenced by different factors. Additionally, many studies focus on only one sector, such as urban water use, while there is little cross-sectoral work (Sohn, 2011). However, water use behaviors by sector are changing, with withdrawals in “industry and at thermoelectric plants having steadily dropped per unit of output” over the past forty-five years, and the extent of these changes in the full spectrum of water use sectors is not yet well understood (Brown et al., 2013).

While previous statistical analyses on water use have identified relationships between water use and potential explanatory variables, many suffer from methodological limitations. Some are cross-sectional analyses of aggregated data from multiple sites, but fixed at one point in time (e.g. Brown et al., 2013; Lee et al., 2015). Cross-sectional studies cannot provide insight into causation due to the potential presence of unobserved confounding variables. Since these studies do not analyze behavior over time, the timing of the data window may not be representative of long-term
trends. Other studies are based on years of historical records from a single area (House-Peters et al., 2010; Balling et al., 2008). This type of analytical approach, called time-series, can be problematic in assessing environmental data, which are often not strictly normally distributed (Yue et al., 2002). Forecasting future trends based of a singular time-period in one location may not represent trends state or region wide. Avoiding bias from omitting potential explanatory variables and identifying broadly applicable trends demands a more robust statistical assessment.

Polebitski and Palmer implement regression models to evaluate how water demand fluctuates with climate, pricing structures and land use (2010). The study location is in the Puget Sound region of Washington State, a humid, northwestern area, and focuses only on urban sector water usage. The spatial scale is to the household level and the timescale is a thirty-year period of historical water use data. The study found that more rigorous water pricing policy decreases demand, indicating that economic conditions influence water use, but did not account for regional differences in pricing (Polebitski and Palmer, 2010). The study also contributes to the understanding of how climate scenarios influence water demand, including that increasing temperatures increases the summer demand. However, a larger spatial scale than the single-family tract would be better suited to developing statewide plans (Polebitski and Palmer, 2010).

Panel regression combines the assets of time-series and cross-sectional methods into a single technique, an approach based on variations in both time and space. Data from several locations and through time are pooled in search of qualities unique to each site and those shared among sites. Panel techniques can evaluate repeated temporal measurements across multiple users and allows for identification of generalizable relationships while accounting for potential biases from omitted variables (Lobell and Burke, 2010). There are many instances of panel analysis in econometric studies, where cross-sectional time series data can demonstrate differences between people in the study and observe how an individual has changed over the course of the study. Despite the potential, panel regression has rarely been applied to water and environmental studies.
3. OBJECTIVES

The objective of this work is to evaluate the relationships between water use, time and explanatory variables related to climate and economic conditions in Virginia. The primary research questions include:

- Are there statistically significant trends in water use through time in the state, and do these trends vary by water use sector?
- Are there variations by county in the trends of water use through time?
- Does climate have a significant relationship with water use in Virginia and does this relationship vary by water use sector?
- Does water use in Virginia have a statistically significant relationship with periods of economic transition, with a stronger relationship in any particular sector?

4. METHODOLOGY

4.1 Data Sources and Processing

4.1.1 Virginia Water Use Data System (VWUDS)

This broad-scale analysis of multiple water-withdrawing facilities in Virginia is through times series records from Virginia Department of Environmental Quality (VDEQ): VWUDS. VWUDS consists of monthly water withdrawals in million gallons per month, as well as the geographic location and use sector for each water-withdrawing facility. The dataset consists of approximately 1.3 million observations from 2,436 individual water-withdrawing facilities. The maximum length of record spans thirty-three years from 1982 to 2015 and the average record length is fifteen years. There are thirteen distinct sectors in VWUDS: agriculture, commercial, fossil power, hydropower, irrigation, manufacturing, mining, municipal, non-municipal, nuclear power, other and unknown. The data in VWUDS are largely user reported, where site managers report usage information directly into the VDEQ modeling system called VA Hydro. The VDEQ implements several QA/QC measures to account for potential user reporting errors. VA Hydro is designed to flag the site’s latitude/longitude if it is out of the Virginia state bounds and prompts the user to update before saving. Another control is on likely unit conversion errors, implemented by comparing all the sites’ five year and two year averages to the current year’s reporting totals to detect order of magnitude conflicts. There is also a check on missing values via a “signed and
submitted” feature. The water supply planners at VDEQ contact all sites that lack a signed submission for follow up, starting from the facilities which historically have high average use.

For the purposes of this study, the VWUDS water use sectors were subgrouped via Microsoft Access queries as: (1) Agriculture: containing irrigation and agriculture, (2) Energy: containing fossil power and nuclear power (3) Industrial: containing aquaculture, mining, manufacturing and industrial (4) Municipal: containing municipal and public water supply. The commercial, other, unknown and non-municipal entries were sorted based on a review of the facility names. Hydropower was not included in this analysis because the Virginia Department of Environmental Quality discontinued the requirement to report annual water withdrawals for those facilities around 2005, resulting in a large decrease in the volume of water use reported. Aquaculture, consisting of fisheries and hatcheries, is classified within the agriculture sector in VWUDS (VDEQ, 2015). However, a review of typical fishery practices in Virginia, such as at Coursey Springs, suggest that hatcheries consist largely of tanks with recirculating water and, oftentimes, catching, gutting and processing takes place on site. Therefore, aquaculture water use sites operate similarly to industrial sector proceedings, and were reclassified as such. The monthly time step of VWUDS was aggregated by water year, a hydrological period from October 1st of one year to September 30th of the next as defined by the USGS. Additionally, the data were spatially aggregated into county water year totals.

Table 1 shows sorted VWUDS information, for example, there are seventy-three counties that have some facilities withdrawing water for purposes classified as agricultural. There is a wide range of distribution, with Accomack County having 103 facilities reporting agricultural water use and Greene having one such facility. Yet, to give a glimpse of sample size, the average number of facilities in a given county reporting agricultural water use is eleven. The smallest sample size is for the energy sector, and therefore is susceptible to error.

<table>
<thead>
<tr>
<th>Water Use Sector</th>
<th>Number of Counties Withdrawing Water</th>
<th>Number of Facilities Withdrawing Water</th>
<th>Average Number of Facilities in County</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>73</td>
<td>804</td>
<td>11</td>
</tr>
<tr>
<td>Energy</td>
<td>19</td>
<td>21</td>
<td>1.1</td>
</tr>
<tr>
<td>Industrial</td>
<td>93</td>
<td>302</td>
<td>3.2</td>
</tr>
<tr>
<td>Municipal</td>
<td>103</td>
<td>1301</td>
<td>12.6</td>
</tr>
</tbody>
</table>
4.1.2 Bureau of Economic Analysis (BEA)

Population and personal income data were obtained via the public access Interactive Tables from the Bureau of Economic Analysis (BEA). The population data available in BEA, taken from the U.S. Census, are available at the county level and used to normalize the municipal water use data per capita. Per capita personal income, in dollars, was selected as a regression parameter over Gross Domestic Product (GDP), as it is available annually at the desired county-level spatial scale.

4.1.3 PRISM

Climate information from the PRISM Climate Group of Oregon State University is a reanalysis product of weather station and radar data, interpolated to create gridded spatial datasets (Daly, 2008). Daily PRISM records for temperature and precipitation are available from 1980 to the present.

![Figure 3: ArcGIS ModelBuilder Flow for PRISM Data Processing. A folder of daily PRISM files of temperature or precipitation values from 1983 to 2015 are input. These are trimmed to a Virginia buffer, and a counties layer with zonal statistics tool is used calculate averaged annual climate information in each county.](image)

In ArcGIS Model Builder, an iterative process was used to spatially aggregate the daily PRISM data (Figure 3). First, a state boundary of Virginia and a shapefile of Virginia counties were created in GIS. The raw inputs were PRISM raster files of mean temperature and precipitation every day from 1983 to 2015. Applying the iterate raster tool, these were accumulated into a full PRISM table. Information from the PRISM raster file is trimmed to the extent of a ten-kilometer state buffer of Virginia, to capture any cell touching Virginia since PRISM rasters have a resolution of four kilometers. Zonal statistics calculates the mean of the trimmed PRISM daily data within the bounds of each county for temperature and precipitation. The cleaned outputs are tables containing the spatially averaged precipitation and mean temperature on each day from 1983 to
2015 in each county. These were then temporally aggregated to water year in an R loop. This process aggregates the daily temperature into mean temperature and the daily precipitation to total precipitation in millimeters, over the water year in each county.

Intraseasonal fluctuations, such as periods without rainfall, also have the potential to drive changes in water use. Heatwaves could increase the water demand for crops or require more water for cooling purposes. Such dry or hot periods might not be reflected in a seasonal average measurement. The daily county-level records were used to calculate measurements of intraseasonal variability in temperature and precipitation, for each year and county. A calculation of maximum consecutive days of temperature above threshold is used to quantify periods where there is a heatwave. The threshold temperature value was taken at 27°C, the 95\textsuperscript{th} percentile value of PRISM mean daily temperature data for the whole state. A higher than 95\textsuperscript{th} percentile temperature value for more than two days is regarded as the threshold for human health impacts as defined by Anderson and Bell (2011). The heatwave metric is reported in days. For the length of the daily temperature time series, every day above the threshold temperature adds one to the day count, restarting the count each time a subsequent day does not exceed the threshold.

The GINI index is as a measure of inequality within a distribution, and commonly applied in economics to quantify income disparities. The GINI index can also be applied as a measure of rainfall variability (Rajah et al., 2014). GINI is reported as a ratio ranging from 0 to 1, where 0 represents the exact same amount of rainfall being evenly distributed over each day of the year, and 1 representing an extreme in which all the rainfall occurred on a single day. To estimate the GINI index, the first step is to form a ‘Lorenz’ curve. The daily precipitation PRISM data were sorted in order of increasing depth, summed cumulatively and taken as a percentage of the total annual precipitation (Rajah et al., 2014). As seen in Eq. 1, the GINI index is then calculated by doubling the area between a straight diagonal line, representing the zero case of rainfall distribution, and the Lorenz curve, normalized by the length of the time series (n). A precipitation value from the observed population of daily PRISM precipitation information is represented by \(y\) (Rajah et al., 2014). The GINI Index is dimensionless and nonparametric, meaning it is not dependent on whether or not the data are of normal distribution, and therefore offers a more robust measure of variability than typical metrics like standard deviation (Rajah et al., 2014).

\[
G = \frac{1}{n} \left( n + 1 + 2 \left( \frac{\sum_{i=1}^{n}(1 + n - i)y_i}{\sum_{i=1}^{n}y_i} \right) \right) \quad (1)
\]
4.2 Trend Assessment

The first objective of this analysis is to assess water use trends through time, accomplished via Mann-Kendall testing. Mann-Kendall (MK) tests are used to assess if there is a statistically significant upward or downward trend in variable of interest through time, where the null hypothesis is that no trend exists. MK is a non-parametric test that ranks observations chronologically, beneficial as the water use data are not necessarily from a normal distribution. There is evidence that issues of autocorrelation, where observations within the dataset are interdependent, impacts the ability to detect trends in hydrological data (Yue et al., 2002). Pre-whitening is a process that computes independent residuals from a lag-one autoregressive model, and the MK test is subsequently performed on the detrended, independent residuals (Bayazit, 2007). The ‘zyp’ package in R was applied to perform modified MK tests that included pre-whitening to reduce the effects of autocorrelation.

Separate MK tests were performed on each of the major water use sectors (energy, municipal, industrial, and agricultural), and on both total water usage and water usage normalized per facility/capita, resulting in a total of eight statewide MK tests. Total usage is valuable for understanding overall trends, and per facility testing provides insights into time trends irrespective of whether more water withdrawal sites are opening or population is growing. Unusual trending in agricultural water use (pre 1990) and municipal (pre 1988) were not included in the tests and was likely due to changes in reporting requirements. The input time-series data for the MK tests were:

- Agricultural Water Use, from 1990 to 2015
- Energy Water Use, from 1983 to 2015
- Industrial Water Use, from 1983 to 2015
- Municipal Water Use, from 1988 to 2015

After assessing the statewide trends through time, MK tests were looped over each county in Virginia to determine which counties are the strongest contributors to water use trends. Significance is generally taken at a p-value of 0.05, indicating that there is less than a five percent chance that no trend occurred in the county. At the county level, Bonferroni correction is used such that significance level of 0.05 is normalized by the number of counties in each sector, adjusting for multiple statistical tests being performed simultaneously on the data. For example, if there are 10 counties withdrawing water for the energy sector, the threshold for significance is taken at 0.005.
4.3 Regression

A panel regression was applied to identify relationships between water use and explanatory variables at a county-level scale. Panel regression combines the assets of time-series and cross-sectional methods into a single technique, an approach based on variations in both time and space. Data from multiple locations and through time are pooled in search of qualities unique to each site and those shared among sites. Panel techniques allow for identifying generalized relationships while accounting for potential bias from omitted variables (Gelman, 2006).

The general form of a panel regression equation is:

\[ \log(y_{it}) = \beta_0 + \beta_1 x_{it} + \varepsilon_{it} \]  

(2)

The response variable \( y \) in this study is volume of annual water used per facility or capita in each county \( i \) and year \( t \). Water use was log-transformed such that the data from each sector would better match a normal distribution, as the histogram demonstrates (Figure 4). Energy water use is not normally distributed under log-transform (Figure 4). There is a heavy tail at the high end of the distribution, attributed to the nuclear facility in Louisa operating with much higher water withdrawals compared to the other facilities in the energy sector. A log-transformed formulation was still applied to the energy model for consistency with the other sectors, and may have attributed to some uncertainty in the results, as addressed in future work.
Figure 4: Histogram of log-transformed Water Use. (a) Agriculture water use log-transformed distribution, containing 73 counties (b) Energy water use log-transformed distribution, containing 19 counties. (c) Industrial water use log-transformed distribution, containing 93 counties. (d) Municipal water use log-transformed distribution, 103 counties.

The explanatory variables, $x$, are potential drivers of water usage which are tested for statistical significance. The betas are regression parameters, $\beta_0$ being the intercept term and $\beta_1$ being a slope. The subscripts $i$, $t$ are indices corresponding to entity (county) and time, representing the two-dimensionality of panel regression, and $\varepsilon_{it}$ is the error term (Gelman, 2006). The intercept terms are county-specific, while a single slope term is used. The panel regression was performed with the R ‘lme4’ package.

Panel regression is useful to identify water use relationships with explanatory factors in Virginia because it allows for the water use to be evaluated both spatially and temporally, which lowers uncertainty and enhances the accuracy in predicting population parameters (Gelman, 2006). The relationship between climate and water use may vary by water use sector, for example, if drought years result in a more substantial difference in agricultural water use than industrial. To account for this, separate regression models were developed for each of the water use sectors. The inputs for each sector’s regression were merged datasets, containing the water use per facility or capita, county name,
year, mean water year temperature, total water year precipitation, intraseasonal variables and per capita income for each county. Water use per capita/per facility was selected over total use in order to account for any trends relating to a changing amount of users. For example, an increase in the number of facilities opening could cause a usage increase that is not a reflection of actual water use behavior at the facility level, but rather the number of water withdrawals.

Virginia’s climate is not uniform through the state, requiring the conversion of weather records to incorporate spatial variability. The National Oceanic and Atmospheric Administration’s climate divisions dataset defines 344 distinct regions of climatic homogeneity within the continental US, and six of which are within Virginia (NOAA; see Figure 5).

![NOAA Climate Divisions](image)

**Figure 5: NOAA Climate Divisions of Virginia**

**Table 2: Climate Statistics per Virginia NOAA Division**

<table>
<thead>
<tr>
<th>Climate Zone</th>
<th>Avg Annual Temp (C)</th>
<th>Total Precip (mm)</th>
<th>Std Dev Annual Precip</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tidewater</td>
<td>14.97</td>
<td>1180.42</td>
<td>191.53</td>
</tr>
<tr>
<td>E. Piedmont</td>
<td>14.10</td>
<td>1127.64</td>
<td>188.42</td>
</tr>
<tr>
<td>W. Piedmont</td>
<td>13.52</td>
<td>1162.66</td>
<td>204.11</td>
</tr>
<tr>
<td>Northern</td>
<td>12.52</td>
<td>1084.58</td>
<td>187.58</td>
</tr>
<tr>
<td>Central Mts</td>
<td>11.53</td>
<td>1079.98</td>
<td>169.61</td>
</tr>
<tr>
<td>SW Mts</td>
<td>11.65</td>
<td>1145.44</td>
<td>149.52</td>
</tr>
</tbody>
</table>

Both temperature and precipitation vary depending on the climate division (Table 1). The Tidewater zone of Virginia, near the coast, has on average highest temperatures and highest total precipitation annually in the state. Comparatively, the mountainous regions at higher elevations have lower mean temperatures annually. These distinct climate trends by region call for an analysis
that accounts for measurements as normalized by the expected value for the specific region, not the whole state. Economic variables also vary widely across the state. For example, in 2015, the per capita income in the northern Arlington County (86161 USD/person) was 160% higher than that of the southwestern Montgomery county (33184 USD/person).

Anomaly values represent how the measurements in a given county and year compare to the long-term average for that county, helping to account for this significant spatial variation in climate across the state. As seen in Eq. 3, anomaly values of temperature were calculated by subtracting the long-term average temperature (ranging over study length, from 1983 to present) for that county from the observed value and dividing by the long-term standard deviation, i.e. a z-score (Shortridge et al., 2016). For example, in a county with a long-term mean temperature of fourteen degrees and standard deviation of two degrees, a year with a mean temperature of fifteen degrees would have an anomaly value of 0.5, signifying a warmer than average year for that county. Negative values would represent lower than average conditions. Anomaly values were calculated for all explanatory variables: temperature, precipitation, GINI, heatwave and per capita income. The use of anomaly values also decreased the cross-correlation observed between explanatory variables. Mean annual temperature (mean.t) and consecutive days above heatwave threshold (htwave) were the two explanatory variables that exhibited the highest correlation, as both are rooted in the daily temperature data from PRISM. Using anomaly values was found to reduce the correlation coefficient between mean.t and htwave from 0.55 to 0.43.

\[
T_{AN} = \frac{T_{obs} - \bar{T}}{sd(T_{obs})}
\]  

(3)

Two possible regression formulations were tested to evaluate the importance of incorporating intraseasonal variables (Table 2). Anomaly values were used in both formulations as a means of accounting for the high spatial variability both in climate and economic conditions in Virginia. Year (Y) is included as an explanatory variable as a means of accounting for trends through time. The independent variables of climate temperature (T), precipitation (P), GINI Index (G), Heatwaves (H) and per capita income economics (E) are the proposed explanations for changes in water use, the statistical significance of which were assessed. Indices \(i\) and \(t\) relate to the two-dimensionality of panel regression, corresponding to county and time, respectively.
Table 3: Tested Model Formulations

<table>
<thead>
<tr>
<th>Form</th>
<th>Regression Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Form I</td>
<td>( \log(WU_{it}) = \alpha_i + \beta_1 T_{it} + \beta_2 P_{it} + \beta_3 Y_{it} + \varepsilon_{it} )</td>
</tr>
<tr>
<td>Form II</td>
<td>( \log(WU_{it}) = \alpha_i + \beta_1 T_{it} + \beta_2 P_{it} + \beta_3 Y_{it} + \beta_4 G_{it} + \beta_5 H_{it} + \beta_6 E_{it} + \varepsilon_{it} )</td>
</tr>
</tbody>
</table>

Two metrics were applied to evaluate model fit. Akaike Information Criterion (AIC) assesses relative model quality as compared to other models, with a lower AIC indicating better fit. AIC penalizes for the complexity of the model, favoring more parsimonious formulations. In contrast, with Normalized Root Mean Square Error (NRMSE), additional regression parameters can only decrease the NRMSE value or cause no change. NRMSE measures the difference between model predictions and the observed values, with lower NRMSE indicating a better model fit (Eq. 4).

\[
\frac{NRMSE}{sd} = 100 \frac{\text{mean}(\text{pred} - \text{obs})^2}{sd(\text{obs})} \tag{4}
\]

The final proposed regression equation, using Form II is:

\[
\log(WU_{it}) = \alpha_i + \beta_1 T_{it} + \beta_2 P_{it} + \beta_3 Y_{it} + \beta_4 G_{it} + \beta_5 H_{it} + \beta_6 E_{it} + \varepsilon_{it} \tag{5}
\]

5. RESULTS AND DISCUSSION

5.1 Time Trends

5.1.1 Statewide Trends

One objective of this analysis was to determine whether there are any water use trends through time in Virginia. At the statewide level, Mann-Kendall tests were implemented on each sector to identify overall trends. Two sets of MK tests, one for total usage and another for per-facility usage, were performed on each sector. P-values indicate the statistical significance of the trend (i.e., the probability that the observed trend is simply the result of random variation), while tau values indicate directionality. Total agricultural water usage has an upward trend over time, but this trend is not statistically significant, likely due to the large amount of year-to-year variability and shortened record length in agricultural water use (Table 3). Per-facility agricultural use is flatter than total agricultural use, indicating that an increasing number of agricultural water users drives
the upward trend in total use (Figure 6a). Unusual trending in agriculture (pre 1990) and municipal (pre 1988) usage, were not included in the statistics and are likely due to reporting changes, as indicated by red lines in Figure 6. There is a significant decrease in water usage over time in the industrial sector, both in total use and when normalized per facility (Figure 6c). When normalized per facility, the energy sector water usage has significantly decreased over time. The municipal water usage, both statewide and per capita, has not significantly changed over time (Figure 6d).

Table 4: Statewide Water Use Trends

<table>
<thead>
<tr>
<th>Mann-Kendall</th>
<th>Total Use</th>
<th>per Facility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sector</td>
<td>Tau</td>
<td>p-value</td>
</tr>
<tr>
<td>Agriculture</td>
<td>0.24</td>
<td>0.097</td>
</tr>
<tr>
<td>Energy</td>
<td>-0.11</td>
<td>0.37</td>
</tr>
<tr>
<td>Industrial</td>
<td>-0.63</td>
<td>3.9e-07*</td>
</tr>
<tr>
<td>Municipal</td>
<td>0.17</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Figure 2 shows a disparity between predicted and observed water use in Virginia. While demand was projected to increase by thirty-two percent by 2040, actual water use statewide from 2000 to 2010 has declined. The statewide MK analysis gives insight into the driving sectors behind the observed decline. Industrial and energy sector water usage significantly decreasing through time has contributed to the overall water use decrease observed in Figure 2 (Table 3). The water demand estimate for Figure 2 was potentially based largely on population growth, and does not adequately capture the effect of factors within the energy and industrial sectors like technological changes. The discrepancy between projected and actual trends in Figure 2 may be a factor of water efficient technologies becoming cheaper and more accessible to the energy and industrial sectors. Additionally, conservation policies implemented by the VDEQ to restrict energy facilities’ water withdrawals appear to have had significant impact.
**Figure 6: Statewide Water Use Trends over the study period of 1983-2015.** Orange lines are per facility/capita water use, and blue indicates total water use. Areas before red dashed lines are discluded years. (a) Agriculture sector displayed no significant trends through time. (b) Energy sector water use declining per facility over time. (c) Industrial water use declining per facility and in total. (d) Municipal sector displayed no significant trends over time, though the recent (2009 on) decline is of interest.

### 5.1.2 County Level Trends

Water use trends through time were also explored spatially by county. This was done by repeatedly performing MK testing over water withdrawal facilities as sorted by county. Hatching indicates statistically significant trends in total water withdrawals in the county (Figure 7). Significance is generally taken at a p-value of 0.05, indicating that there is less than a five percent chance that no trend occurred in the county. Bonferroni correction was employed such that significance level of 0.05 is normalized by the number of counties in each sector, adjusting for the multiple statistical tests that were performed simultaneously on the data. Dark blue counties with hatching have the strongest declining water usage through time. Dark red counties have the strongest increase in water use through time. These spatial depictions demonstrate which counties are the strongest players driving water use per facility/capita patterns through time. This confirms some trends, for example, that northern Virginia, a rapidly urbanizing area, has several counties...
with significantly increasing municipal water usage. However, these increases are offset by declining trends elsewhere. Interestingly, there are no individual counties with significant water use changes through time in the energy sector, but a sufficient majority of energy facilities have declining water usage to contribute to the overall significantly decreasing water use trend statewide.

The county level time trend analysis provides insight on hotspots of water use, highlighting particular counties with increasing trends through time. For example, although there was no definitive statewide trends in water use per agricultural facility, Caroline county demonstrated a strong, significant increase in total water use for agricultural purposes over time, and can be specifically targeted for conservation policies. This is particularly useful in the municipal sector, where counties including Lee and Washington can visualize and address their increasing domestic water use per capita over time (Figure 7d).

Figure 7: County Water Use Time Trends. Hatching indicates significance, taken and a p-value of 0.05 divided by the number of counties in each sector in a Bonferroni correction. Red values indicate a high increasing water use trend over time, and blue indicates a declining trend. White spaces indicate counties where there were no facilities using water in the sector.
5.2 Regression Model Structure and Accuracy

As seen in Table 2, two regression model formulations were evaluated. Form I consists of using the simpler climate parameters of mean annual temperature and total annual precipitation. Form II incorporates the more complex climatic explanatory variables of heatwaves and GINI Index, and per capita income. As seen in Tables 4-7, each potential model formulation was evaluated for AIC and root mean squared error as related to the standard deviation. AIC penalizes for each additional parameter. The lower AIC among two model formulations signifies better fit.

RMSE is expected to decrease or remain the same as more explanatory variables are added, however, as seen in Tables 6 and 7, there are instances in this analysis which RMSE increased slightly. This is due to the input parameter, water use, being log transformed. Converting the predictions from the model back to the original water use units by taking the exponent caused the RMSE to occasionally increase slightly when additional parameters were added. When accounting for the log transform, errors remain the same or decrease with each added parameters.

Table 5: Agricultural Regression Model Fit

<table>
<thead>
<tr>
<th></th>
<th>Form I</th>
<th>Form II</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>4139.76</td>
<td>4139.72</td>
</tr>
<tr>
<td>RMSE/sd</td>
<td>0.824</td>
<td>0.822</td>
</tr>
</tbody>
</table>

Table 6: Energy Regression Model Fit

<table>
<thead>
<tr>
<th></th>
<th>Form I</th>
<th>Form II</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>1406.89</td>
<td>1398.26</td>
</tr>
<tr>
<td>RMSE/sd</td>
<td>0.411</td>
<td>0.395</td>
</tr>
</tbody>
</table>

Table 7: Industrial Regression Model Fit

<table>
<thead>
<tr>
<th></th>
<th>Form I</th>
<th>Form II</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>8365.55</td>
<td>8365.71</td>
</tr>
<tr>
<td>RMSE/sd</td>
<td>0.366</td>
<td>0.374</td>
</tr>
</tbody>
</table>

Table 8: Municipal Regression Model Fit

<table>
<thead>
<tr>
<th></th>
<th>Form I</th>
<th>Form II</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>5181.44</td>
<td>5181.45</td>
</tr>
<tr>
<td>RMSE/sd</td>
<td>0.443</td>
<td>0.444</td>
</tr>
</tbody>
</table>
The Form II equation was applied to each sector, with four individual regression being conducted with water use in sector as the response variable. The predictions from the Form II model of each sector were evaluated as compared to observed water use values in that sector (Figure 8). Although the agricultural model is under predicting at high observed values, this is indicative that the model is behaving conservatively (Figure 8a). The energy sector follows the best-fit line closely, with one cluster of high values that may be correlated with single facility displaying different water use behaviors than others grouped in the sector (Figure 8b). The errors in the energy model may also related to the relatively small sample size, with only nineteen energy facilities in Virginia. The industrial sector displayed fairly good model fit, with nearly equal distribution of over and under predicted values following the best fit line (Figure 8c). The municipal sector has better fit for lower water use values, and under-predicts for higher observed water use values (Figure 8d). However, there are two clusters of poor model fit at higher water use values, correlated with two facilities that are likely using water differently than other users in the sector (discussed in Future Work). The clustering data structure throughout Figure 8 is associated with the intercept term for panel regression (Equation 5). Under panel regression, the intercept term, alpha, is not fixed and varies by county, such that the clustering displayed in Figure 8 is usually associated with the behavior of a single county.
Figure 8: Form II- Predicted vs Observed Water Use. Observed water use values, directly from VWUDS are on the horizontal axis, and predicted values output by Form II of the study’s regression model are the y-axis. Black lines indicate best fit.

5.3 Regression

A panel regression was formulated with water usage per facility (or capita) as the independent variable. In the MK test analysis, quite a few time trends were identified in the water usage data, including that industrial and energy water use is decreasing over time. This trending was accounted for in the regression by including year as an explanatory variable in the panel regression. Another method, detrending the data by using residual values from a linear regression of water use on years, was evaluated but resulted in worse model fit.

The explanatory variables relating to climate are annual average temperature, annual total precipitation, GINI index of days without rainfall and a heatwave parameter of consecutive days above a threshold of 27 °C. Finally, per capita income was included as a variable to assess economic conditions. The explanatory variables were converted into anomaly values to account for the high spatial variation in climate and economics across Virginia. A higher anomaly means that a given year is warmer than on average for a particular spatial point.
This study aims to identify whether climatic and economic conditions have a significant relationship with water use in Virginia and how those relationships vary by water use sector. In the agricultural, industrial and municipal sectors, a significant relationship between water use and a climate parameter was indicated via panel regression. The estimate value indicates whether there is a positive or negative trend within the given margin of error. The null hypothesis is when the regression parameter is zero the variable has no relationship with water use. On warmer than average years, significantly more water was used for agricultural purposes (Table 10). Similarly, on wetter than average years, less water was used in the agriculture sector (Table 10). The demonstration of these intuitive trends quantitatively is indicative that the panel regression approach is working effectively. The highest water-withdrawing county in the agricultural sector was Chesapeake, drawing more than 200 MG in 1994. The model formulation did not change which variables displayed significance.

Table 9: Water Use per Agricultural Facility: Regression Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Form I</th>
<th></th>
<th></th>
<th>Form II</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Standard Error</td>
<td>pvalue</td>
<td>Estimate</td>
<td>Standard Error</td>
<td>pvalue</td>
</tr>
<tr>
<td>Temperature</td>
<td>0.162</td>
<td>0.037</td>
<td>1.17e-05*</td>
<td>0.121</td>
<td>0.042</td>
<td>0.004*</td>
</tr>
<tr>
<td>Precipitation</td>
<td>-0.126</td>
<td>0.038</td>
<td>1.01e-03*</td>
<td>-0.110</td>
<td>0.041</td>
<td>0.007*</td>
</tr>
<tr>
<td>GINI Index</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.023</td>
<td>0.041</td>
<td>0.567</td>
</tr>
<tr>
<td>Heatwave</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.067</td>
<td>0.037</td>
<td>0.067</td>
</tr>
<tr>
<td>Economics</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.187</td>
<td>0.123</td>
<td>0.129</td>
</tr>
<tr>
<td>Year</td>
<td>0.0064</td>
<td>0.0046</td>
<td>0.166</td>
<td>-0.013</td>
<td>0.013</td>
<td>0.330</td>
</tr>
</tbody>
</table>

Interestingly, during years where the per capita income increased, less water tended to be withdrawn by energy facilities (Table 11). The highest water using county, Louisa, had yields over 800,000 MG. Louisa is home to the only nuclear energy facility in Virginia, and its’ high withdrawals indicate it is a controlling player in water use trends for the energy sector. This facility is significantly larger user than the other facilities in sector, also contributing to the heavy tail at the top of the energy water use histogram (Figure 4). In addition to concerns about the higher water withdrawing facility, there is a considerably lower sample size for the energy sector than the other sectors considered in this analysis, with VWUDS data containing only nineteen Virginia counties using water for the energy productions and only twenty-one total energy users reporting (Table 1).
Table 10: Water Use per Energy Facility: Regression Results

<table>
<thead>
<tr>
<th>Energy</th>
<th>Form I</th>
<th>Form II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parameter</td>
<td>Estimate</td>
</tr>
<tr>
<td></td>
<td>Temperature</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>Precipitation</td>
<td>-0.027</td>
</tr>
<tr>
<td></td>
<td>GINI Index</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Heatwave</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Economics</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Year</td>
<td>0.0105</td>
</tr>
</tbody>
</table>

Industrial water use per facility and municipal water use per capita were only statistically significant with the GINI Index (Tables 10, 11). For both these sectors, water use increases in a high GINI year, indicating that when there is a more unequal distribution of rainfall, more water was used. This suggests that at both the domestic and commercial level, unequal rainfall distribution, like dry spells, causes increasing water demand and strains the supply. The highest water use county for the industrial sector is Hopewell and for the municipal sector is New Kent.

Table 11: Industrial Water Use per Capita: Regression Results

<table>
<thead>
<tr>
<th>Form I</th>
<th>Form II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>Estimate</td>
</tr>
<tr>
<td>Temperature</td>
<td>0.020</td>
</tr>
<tr>
<td>Precipitation</td>
<td>0.0006</td>
</tr>
<tr>
<td>GINI Index</td>
<td>-</td>
</tr>
<tr>
<td>Heatwave</td>
<td>-</td>
</tr>
<tr>
<td>Economics</td>
<td>-</td>
</tr>
<tr>
<td>Year</td>
<td>-0.0357</td>
</tr>
</tbody>
</table>
Table 12: Municipal Water Use per Capita: Regression Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Form I</th>
<th>Form II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Standard Error</td>
</tr>
<tr>
<td>Temperature</td>
<td>0.006</td>
<td>0.012</td>
</tr>
<tr>
<td>Precipitation</td>
<td>-0.0025</td>
<td>0.012</td>
</tr>
<tr>
<td>GINI Index</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Heatwave</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Economics</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Year</td>
<td>-0.002</td>
<td>0.001</td>
</tr>
</tbody>
</table>

As seen in Tables 4-7, AIC values among models from the same sector were all similar. Adding the intraseasonal variables in Form II proved beneficial for the municipal and industrial sectors, as it revealed a statistically significant relationship with GINI, which was not apparent at the annual-scale. The additional parameter of per capita income revealed a statistically significant relationship between per capita income and energy water use, another advantage of the Form II model (Table 11). Form II, including the six regression parameters as anomaly values, was ultimately selected as the final formulation for this analysis as the additional parameters revealed several statistically significant relationships and using anomalies contributed to reducing the correlation between the climatic variables.

Figure 9 is a representation of the regression estimates for the six explanatory variables across all sectors, with standard error bars indicated in bold. The regression estimates plot also provides insight into the relationships that did not prove statistically significant. For example, it is interesting that per capita income was positively associated with water use in both the agriculture and the industrial sectors (Figure 9). Likewise, there is a positive association between heatwaves and increasing agricultural water use. Some of the estimates, such as the temperature parameter for both industry and energy, have quite wide margins of error, as addressed in future work.
Figure 6: Regression estimates for the explanatory factors as related to water use. The horizontal axis indicates the estimated slope values for a given regression parameter. Points indicate the slope estimate that the regression model predicted, while lines indicate a 95% confidence interval. The green line is associated with agriculture water use as the response variable, red with energy, blue with industrial and black with municipal. Significant trends are where the entire confidence interval is contained on either the positive or negative side.
6. FUTURE WORK AND LIMITATIONS

There are several areas for further refinement in this work and opportunities to continue this study. First, the VWUDS dataset was clustered into four sectors of agriculture, energy, industrial and municipal based on a review of facility names. The regression models would fit better when the input response variable, water use, behaves similarly for all the observations. A successful example of the sorting based on facility names is the agricultural sector, where farms, golf courses and orchards were grouped together since the volume and applications of water use is very similar among these sorts of facilities. However, in future analyses, a more thorough investigation into each facility’s function would contribute to better model fit. For example, an approach where facility managers are surveyed about water use operations could help precisely sort the sectors. Figure 8 shows two clusters of higher water use that the municipal sector model did not accurately capture. The largest using county, New Kent, has high water withdrawals respective to the county’s population. Upon further investigation, New Kent has a municipal water facility which treats and ships water to other counties, withdrawing water more water than what is actually used by the New Kent population. This is indicative that per capita usage may not have been the best method for normalizing municipal water use by county, as it does not reflect water being transported across county lines. Other limitations include the small number of reporting energy facilities and counties which have only one facility reporting to VWUDS (Table 1).

More sectors, beyond agriculture, were thought to have been sensitive to mean annual temperature and total annual precipitation. Additionally, some of the parameters identified as significant, have wide standards of error, such as per capita income in the energy sector (Figure 11). If the monthly VWUDS data was aggregated to a seasonal scale instead of water year, more refined results could be achieved. In addition to working on a seasonal scale, a more refined analysis could be achieved by assessing water usage at each facility, rather than county totals. In understanding water use trends on the finer temporal and spatial scales, future work includes:

- Adjusting the temporal scale: aggregating seasonally, rather than yearly
- Refining the spatial scale: aggregating per facility, rather than by county
- Conducting the study on consumptive water use as opposed to water withdrawals
- Quantifying more potential explanatory variables, like conservation policies
- Developing better water demand projections, accounting for industrial and energy sector influences, which contribute to declining water use trends in Virginia
7. CONCLUSIONS

This statistical assessment of water use trends across multiple users and through extensive historical records from VWUDS identified trends through time and helped define the extent of the relationships between water use and climatic/economic conditions. The current body of knowledge focuses on the urban sector and this project includes all water use sectors, including industrial and energy. This led to the identification of key trends to influence state planning, including that both industrial and energy usage are statistically decreasing over time in Virginia. The statistical method of panel regression is uniquely implemented and is working effectively as indicated by the demonstration of intuitive trends, such as more water required for agricultural purposes on warmer than average years. Other trends identified through the regression analysis include a positive relationship between a higher GINI Index, or more unequal rainfall distribution, and water use in both the municipal and industrial sectors. Interestingly, this analysis demonstrated that on years with higher per capita income, less water was used for energy purposes. There are concerns with uncertainty in this trend due to the small sample size of energy facilities in Virginia. Overall, this assessment identified variables which have a statistically significant relationship with water use, and contributed to a better understanding of trends that can inform effective and reliable of water resource planning.
REFERENCES


Weldon Cooper Center for Public Service, Rector and Visitors of the University of Virginia (WCC-UVA). Retrieved 2/28/2018 at https://demographics.coopercenter.org/


APPENDIX A

Data and R Code Repository

Data and source code, in R programming language, for the analyses in this study are published online in an Open Science Framework (OSF) data repository available at: https://osf.io/uwzgs/

The site contains the following:

Folder 01: Water Use Data

This folder contains VWUDS usage data aggregated to water year, code and outputs normalized per facility:

Total Water Use:
- Crosstab_Usagebycountyag.txt: Monthly agricultural water use by county
- Crosstab_Usagebycountyen.txt: Monthly energy water use by county
- Crosstab_Usagebycountyin.txt: Monthly industrial water use by county
- Crosstab_Usagebycountymun.txt: Monthly municipal water use by county
- agcountywateryear.csv: Cleaned agriculture water use by county, aggregated to wtryear
- ecountywateryear.csv: Cleaned energy water use by county, aggregated to water year
- indcountywateryear.csv: Cleaned industrial water use by county, aggregated to water year
- mcountywateryear.csv: Cleaned municipal water use by county, aggregated to water year

Users per county:
- Userbycountyag.txt: Cleaned agricultural water use
- Userbycounty_energy.txt: Cleaned energy water use
- Usersbycountyind.txt: Cleaned industrial water use
- UsersbyCounty_Municipal.txt: Cleaned municipal water use
- Statewide summary and per facility.R: Script to calculate statewide averages and to normalize per users
- bysectorstatewideuse.RData- supporting RData file for above script

Water User Per Facility/capita by sector:
- agperfaciluseeachcounty14.csv: Cleaned agricultural water use per facility
- ecountywateryear12_18_0rm.csv: Cleaned energy water use per facility
- indcountywateryear12_18_0srm.csv: Cleaned industrial water use per facility
- mperpopuseeachcounty12180srm.csv: Cleaned municipal water use per capita
Folder 02: Climate Data (PRISM)

This folder contains data from Oregon State U. PRISM group:

- PRISM_wtryr.R: Script to process daily weather data into seasonal-scale measurements and form variables representing variability in weather conditions for each county
- Temperature:
  - Tempprismlongtable.txt: Daily values of temperature in degrees C
  - Tempwateryear_.csv: Temperature data aggregated to annual means. Subfiles have certain years removed to reflect the shorter year ranges used for agriculture, municipal sectors
- Precipitation
  - Pptprism.CORRECT.longtable.txt: Daily values of precipitation in mm
  - waterppt_.csv: Precipitation data aggregated to annual totals. Subfiles have years removed to reflect the shorter year ranges used for agriculture, municipal sectors

Folder 03: BEA Data

This folder contains per capita income for each county of Virginia.

Folder 04: Mann Kendall Testing

This folder contains the code for the Mann Kendall analysis:

- MannKendallTestingALL.R: Script to determine the trends through time both statewide and in each individual county
- GIS Mann Kendall figure.ppt: Figures generated in GIS showing the county time trends

Folder 05: Regression

This folder contains the following files and subfolders:

- Agriculture Sector
  - Agriculture_regression_results.R: Script to run panel regression for the agriculture sector in each tested formulation, outputs regression parameters, pvalues
  - Agregressionformatfull.csv: Contains agricultural water use and explanatory data (temperature, precipitation, GINI, heatwave, per capita income in regression ready, formatted table)
- Energy Sector
  - Energy_regression_results.R: Script to run panel regression for the energy sector in each tested formulation, outputs regression parameters, pvalues
  - Eregressionformatfull.csv: Contains energy water use and explanatory data (temperature, precipitation, GINI, heatwave, per capita income in regression ready, formatted table)
- **Industrial Sector**
  - **Ind_regression_results.R**: Script to run panel regression for the industrial sector in each tested formulation, outputs regression parameters, pvalues
  - **Indregressionformatfull.csv**: Contains industrial water use and explanatory data (temperature, precipitation, GINI, heatwave, per capita income in regression ready, formatted table)

- **Municipal Sector**
  - **mun_regression_results.R**: Script to run panel regression for the municipal sector in each tested formulation, outputs regression parameters, pvalues
  - **Munregressionformatfull.csv**: Contains municipal water use and explanatory data (temperature, precipitation, GINI, heatwave, per capita income in regression ready, formatted table)