Water Demand Forecasting

October 2016

City Internal Auditor's Office

City of College Station

Water Demand Forecasting

Table of Contents

Appendices

Introduction

The City of College Station has the capacity to produce approximately 30 million gallons of water to serve residents, businesses, and other organizations in the community each day. These customers use the water, returning some part of it back to the City's water treatment facilities that discharges its effluent into the area's rivers, streams, or other environments.

The purpose of water demand forecasting is to make information available to public water suppliers as they conduct their business. Capital investments associated with public water supply systems are extremely expensive, costing millions — even hundreds of millions — of dollars. It thus behooves utility management to make continuing comparisons between current conditions and longer-term forecasts.

Why we conducted this review: An examination of the Water and Wastewater utilities was included in the fiscal year 2016 audit plan based on direction given by the Audit Committee. We conducted a preliminary risk assessment in May 2016 of the Water Services Department and found that a review of the City's water demand forecasting methods was warranted. This was largely based on two factors.

First, population growth has put enormous strain on the City's water and wastewater infrastructure. Over $17 million in expenditures for water capital projects is estimated to be spent in fiscal year 2017 on infrastructure expansion to increase water capacity as well as rehabilitation projects to maintain current infrastructure. In addition, wastewater capital projects are estimated to be over $20 million to fund sewer improvements, wastewater facility expansion, and various line rehabilitation projects. Although no rate increase is proposed for the Water Fund in fiscal year 2017, a rate increase of 8% was approved for the Wastewater Fund in the upcoming fiscal year.

Second, the last comprehensive water and sewer utility rate study commissioned by the City occurred in 2003. This study was completed by Black and Veatch, a global engineering and consulting company specializing in utility infrastructure development and management consulting. Black and Veatch also conducted a utility rate study in 2002. Prior to that, a study conducted by McCord Engineering was completed in 1987. All three of these reports contain water demand forecasts that act as a basis for their cost-of-service study, which functions as a key component for setting utility rates and planning for future growth and infrastructure needs. However, we found that these reports relied on simple forecasting techniques. Although the forecasting methods used in these reports may be reasonable in certain circumstances, many larger utilities use more sophisticated methods that focus on the dynamics of residential, commercial, industrial, and public customers — issues that ultimately relate to the form and growth of a community or region.

Purpose and Objectives

Purpose: The purpose of this review is to evaluate the City's methods — both past and present — for determining future water demand. In addition, this report contains forecasting models that were developed using America Water Works Association accepted methodologies. These approaches are demonstrated in this report in order to best evaluate how widely accepted models compare to City water demand forecasting approaches.

Objectives: This report answers the following questions:

- How effective are the City's methodologies for forecasting water demand?
- How do the City's methodologies for forecasting water demand compare with the most widely accepted approaches promulgated by the American Water Works Association?
- How do weather and demographics impact water demand, and can a better understanding of these impacts be used to better inform policy decisions?

Scope and Methodology

The Office of the City Internal Auditor conducted this review of the City's water demand forecasting methods pursuant to Article III Section 30 of the College Station City Charter, which outlines the City Internal Auditor's primary duties. This examination was conducted in accordance with government auditing standards (except for the completion of an external peer review),[1] which are promulgated by the Comptroller General of the United States.

Although we conducted interviews with City staff and other relevant professionals and researched professional literature and peer reviewed articles; our primary source of criteria in conducting this review comes from the American Water Works Association.

The American Water Works Association (AWWA, http://www.awwa.org/) is the largest nonprofit, scientific and educational association dedicated to managing and treating water. The AWWA has approximately 50,000 members and is a trusted source of industry standards that establishes minimum requirements for materials, equipment, and practices used in water treatment and supply. These standards are used by thousands of manufacturers, distributors, and water treatment facilities worldwide.

The AWWA publishes manuals and journals with peer reviewed scientific research that are widely accepted in the industry. In developing our criteria for forecasting water demand we relied on the AWWA's most recently revised *Forecasting Urban Water Demand* and *Principals of Water Rates, Fees,*

[1] Government auditing standards require audit organizations to undergo an external peer review every three years.

and Charges manuals. The *Forecasting Urban Water Demand* manual was used to construct the water demand forecasting models that we created for comparative purposes.[2]

We examined the City's very-short-, short-, and long-term water demand forecasting approaches as well as their past use of consultants when these consultants provided water demand forecasts. In evaluating whether the City is utilizing an optimal approach to water demand forecasting, we (1) compared the methods utilized by Water Services to methodologies and accepted principles presented in AWWA materials and (2) evaluated the accuracy of the City's forecasts by comparing actual to forecasted amounts.

In this review, we also sought to discern if there were forecasting approaches that could possibly yield stronger and more accurate forecasts than those currently in use. To accomplish this objective, we identified accepted AWWA approaches that best lent themselves to the data that was readily available. Based on our research and data collection efforts, we were able to develop per capita, aggregated time-series extrapolation, sectoral time-series extrapolation, and multiple regression[3] (predictive) models.

We used historical population forecasts and estimates obtained from the City's comprehensive plan to develop a per capita water demand forecasting model. We also conducted a comparative analysis using census data to evaluate the accurateness. Based on this review, we found that the City tends to slightly under project City population on average by approximately 2%. We determined that this was an acceptable error rate for our purposes. The daily amount of water pumped to be put into production towards water consumption was obtained from Water Services. This pumped water data set contained records from January 1996 to July 2016.

Using this pumped water data set, we identified the peak day water demand for 1996 through 2015. We also obtained peak day water demand data for the years 1985 through 1995 from City staff. After combining this data we were able to estimate an aggregate time-series trend line to forecast peak day water demand growth.

For our sectoral time-series extrapolation model, we obtained monthly billed water consumption from 2008 to 2015 by utility customer from the City's utility billing information system. Based on when a meter was read, we created an algorithm to adjust the billed consumption to more accurately account for the month in which the consumption actually occurred. From the Brazos Central Appraisal District, we obtained the state property category descriptions for each property in the City. A relation by property address was created to combine these two data sources.

Finally, we examined the impacts of a number of demographic variables such as price and population, as well as those that account for conservation efforts and inflation using principles set forth by the AWWA. We also wanted to know what impacts weather had on water demand in our local area through the use of multiple regression analysis. To further ensure the reasonableness of our methods, we consulted with a distinguished ecologist, who is knowledgeable in climate science and an expert in developing multiple regression models, to examine and critique our analyses.

---

[2] See Appendices D and E.

[3] Regressions are statistical models that mathematically relates one variable to another, where the magnitude of one of the variables (the *dependent* variable) is determined by the second variable (the *independent* or explanatory variable). In the case of multiple regression, more than one explanatory variable is used to predict the magnitude of change in the dependent variable.

Water Demand Forecasting Background

Accepted Forecasting Methods Range from Simple to Complex

According to the American Water Works Association (AWWA), water utilities across the nation utilize a wide range of forecasting models. These methods can range from simple informal forecasts, in which decision makers judge that the future will act just like their recollection of the past, to complex formal models requiring many variables, large amounts of data, and a significant commitment of resources. The most common approaches fall within one of the following categories: (1) subjective methods (2) per capita or other unit-use coefficient approaches, (3) time-series extrapolation, and (4) regression models.

Subjective Methods. Judgment-based forecasting methods vary widely, ranging from the informed opinion of utility management to highly structured scenario-building methods. Presumably, utilities that do not use a formal forecasting method rely on the informed opinion of management to make decisions. For many small utilities that are experiencing slowly changing conditions, this may be sufficient. Larger utilities and those facing more rapid changes in their service areas would most likely benefit from more elaborate methods.

Per capita and other unit-use coefficient approaches. In its simplest form a per capita model multiplies estimated water use per person by the projected population. This method relies on the ability of analysts to identify reasonable numbers for gallons per capita per day and accuracy of population forecasts that are typically produced by other agencies.[4] Larger urban water systems tend to develop sectoral demand forecasts on a per customer basis, calculating unit water use coefficients for customer by categories such as residential, commercial, industrial, and public.

A variant of the unit-use coefficient approach is to calibrate the demand forecast to the land use plan in the utility service area. Residential, commercial, and industrial land uses are estimated to consume certain amounts of water per acre per year. It is important to note that the effectiveness of water demand forecasts based on land use is greatest in those areas with strict land use regulations, comprehensive land use planning, and a stable industrial structure. The long-range water demand forecast utilized by City Water Services uses this variant unit-use coefficient approach.[5]

Time-series extrapolation. Time-series extrapolation encompasses a variety of techniques, including, simple time trends, exponential smoothing, and autoregressive integrated moving-averages models to project historical water use trends into the future. These models rely on the assertion that future changes in water use can be predicted based on historical changes in water use (ignoring all other possible influences). These methods can provide reasonably accurate forecasts as long as the future is essentially similar to the past. The strength of extrapolation models is that the only data required are the historical data on the variable being forecasted. However, all single variable forecasting methods share a major limitation — they do not account for changes such as population shifts, conservation programs, or price increases.

---

[4] In the City of College Station, the Department of Planning and Development Services produces population forecasts based on certificates of occupancy.

[5] The most recent long-term water demand forecasting model was conducted by the consulting and engineering firm Freese and Nichols in 2014.

Regression Models. The essential feature of these statistical models is the use of a set of driver or explanatory variables to describe why water use has changed historically and to forecast future values. The models directly incorporate anticipated changes in driver variables such as customer income levels, water rates, conservation programs, weather factors, and technology advancements. Because per capita and unit-use coefficient forecasting methods ignore socioeconomic factors, properly designed regression models tend to yield more accurate forecasts.

If trends in water prices, personal income, ownership of water using appliances, population, urban density, and other factors are to be used together in a forecasting model, regression modeling is appropriate. The challenge arises because data for these driver variables must be readily available and obtained (or forecasted) first, before water use forecasts can be developed. This makes the entire forecasting effort far more complex.

Several Factors Should be Considered in Selecting a Forecasting Method

Water utilities should consider the following when selecting and evaluating a water demand forecasting method: purpose of the forecast, data availability, requirements for accuracy of the forecast, how well the forecasting model can be explained to stakeholders in the water planning process, and the ease of updating the model.

Forecasting purpose. The choice of methodology, including the forecast horizon, is directly linked with the intended purpose for the forecast results. The basic application areas for water demand forecasts include: (1) sizing system capacity and raw water supply, (2) sizing the staging treatment and distribution system improvements, (3) water rate setting, revenue forecasting, and budgeting, (4) program tracking and evaluation, and (5) system operations management and optimization.

Capacity issues and raw water supply usually relate to long-term forecast horizons that range from one to several decades. Rate setting and sizing and staging treatment and distribution system improvements in a water system usually involve a medium-term forecast horizon of several years to a decade. In the short term, a few months to a few years, the forecast focus is on budgeting, program tracking and evaluation, and revenue forecasting. Finally, managing and optimizing system operations, such as pumping and maintenance schedules, involve very-short-term forecasts — periods of hours, days, or weeks. See Table 1 below.

Table 1: Water Demand Forecasting Types and Applications

Forecast Type	Forecast Horizon	Applications
Long-Term	Decades, 10-50 years	Sizing system capacity, raw water supply
Medium-Term	Years to a decade, 7-10 years	Sizing, staging treatment and distribution system improvements, investments, setting water rates
Short-Term	Years, 1-2 years	Budgeting, program evaluation, revenue forecasting
Very-Short-Term	Hours, days, weeks	Optimizing, managing system operations, pumping

Customer Disaggregation. Forecast accuracy can often be improved, regardless of the choice of method, by segmenting utility customers into relatively homogeneous groups such as single family residential, multifamily residential, commercial, industrial, or governmental customers. The choice of segments depends on the characteristics of the utility service area and may include additional categories such as high- and low-valued housing areas.

For smaller public water supply systems, relatively simple forecast methods suffice, and not just because of costs. With smaller numbers of customers, disaggregating water use by categories is more likely to result in excessive volatility within each category. Simpler forecasting methods, such as the per capita water demand forecasting approach with no disaggregation, are appropriate in this case. As the water system grows in size, however, customer water use disaggregates become more predictable. Developing a sectoral water demand forecast, which focuses on movements of water use by major customer categories, can result in gains in accuracy and explicability.

Sectoral water demand forecasts also provide better benchmarks for tracking water demand in the near term. Maintaining a sectoral water demand forecast is good business practice when data availability and system financial resources allow it and heterogeneous groups of customers make it worthwhile.

Data Availability. The availability of data is often a primary constraint in developing forecasting models. In general, several years of data are needed to develop medium- to long-term water demand forecasting models. This requirement for time or historical depth of the data is closely related to the importance and unpredictability of weather on urban water use. The historical data must be of sufficient length or time depth to allow unusual weather effects — such as droughts or exceptionally wet, rainy, and cool periods — to be included or accounted for across the historical record.

At the same time, new forces can emerge in the community, causing changes in water use patterns. Examples can include “densification” of settlement patterns, or construction of substantially larger houses with more bathrooms and water using appliances on larger landscaped lots. Conversely, new constructions, especially of townhouses and condominiums, may be built around natural areas with no cultivated landscape. Carefully examining community trends helps analysts determine how many years of data are required and which community patterns are relevant in developing a forecast.

Model Accuracy. The ultimate test of any forecast is how close it came to predicting what actually happened. This suggests that utilities undertake periodic comparisons between previous forecasts and realized values for water demand and utility revenues. With water demand forecasts, however, this is complicated by the importance and inherent unpredictability of transitory weather events. As a result, water demand models should be closely monitored. This essentially means developing a comprehensive water demand regression (probably on a monthly or seasonal basis), and studying the performance of predicted parameters (such as water use rates of residential households) when making allowance for specific weather conditions.

Water Demand is influenced by a Number of Factors

A number of factors have a significant impact on water demand, including population, employment, economic cycles, technology, weather and climate, price, and conservation programs.

Population, Employment, and Technology. Population growth is often the major trend factor in water use. Business cycle factors affect water use because fluctuations in industrial and commercial production translate into commensurate changes in water demand. In addition, water consumption will increase, other things being equal, when family income rises. Technological change can also affect water use over time. For example, widespread installation of garbage disposals and automatic dishwashers in homes may increase domestic water use.

Weather and Climate.[6] Seasonal weather (such as summer high temperatures) and component water use are generated primarily by the local climate (humid, subtropical temperatures). Summer peaking demand is typical. Higher summer demand levels are related to water use for outdoor activities, including lawn watering and gardening, and to the use of evaporative coolers. Seasonal demand patterns are important in planning the capacity of water treatment and distribution systems. Short-term patterns are also critical for scheduling maintenance times for water services infrastructure.

Price. Price effects are important for short-, medium-, and long-term forecasts. Both water use and utility revenue are directly affected by water rate changes. In the short term of a few months, rate hikes can cause consumers to change their behavior. These changes can include taking shorter showers or reducing car washing and lawn watering. In the longer term, if a noticeable rate hike keeps pace with inflation, consumers may adapt through their selection of water using durable goods, favoring appliances with lower water use ratings and possibly innovative landscaping designed to cut back on water use.

Efficiency and Conservation Programs. Water efficiency and conservation programs typically couple an appeal to civic virtue with information on how to use less water. Crisis programs resulting from drought or other supply interruptions have generated large, albeit temporary, reductions in water use. Programs designed to permanently change individual behavior are capable of generating long-term reductions in water usage. Conservation program effects must be thoughtfully included in the water demand forecasting model to minimize errors in projected water use and revenue.

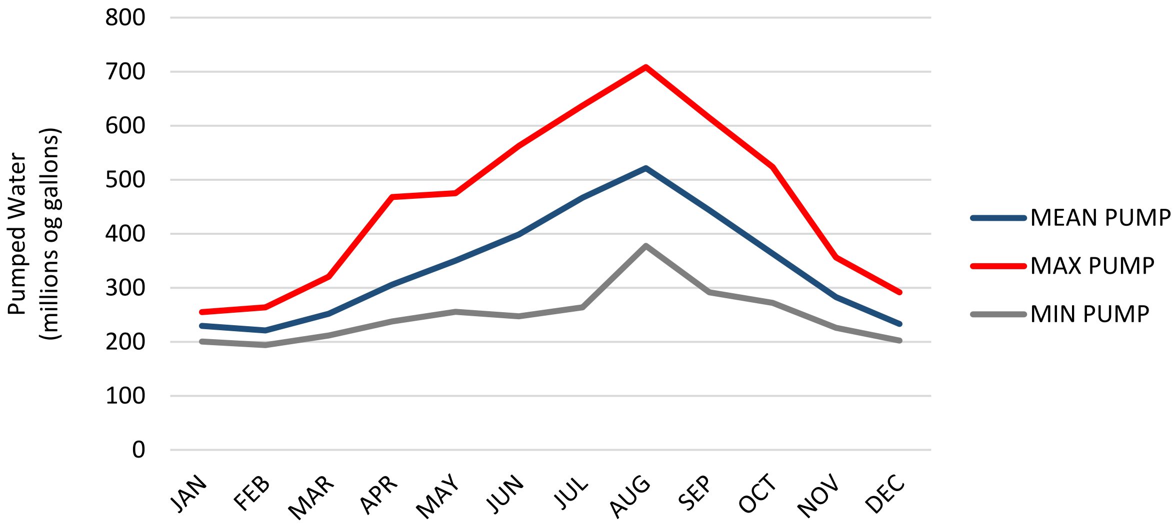
[6] Weather is the daily or monthly changes in temperature, precipitation, relative humidity, whereas, climate is more general, meaning year-to-year or region to region.

Findings and Analysis

Environmental Effects on Water Demand

In College Station, water demand typically peaks during the summer months, and peak summer demand usually occurs in August. However, what may go unmentioned is the increased variability in summer month demands. As we can see from Figure 1, the average water demand range (maximum minus minimum) during summer months (June-August) is almost 5 times larger than the average winter months (December-February) water demand range.

Figure 1: Water Demand Patterns by Month (2000-2015)



To evaluate this variability, we generated the following three predictive multiple regression models based on weather and pumped water data: 1) a daily model for each month, 2) a general monthly model, and 3) a monthly model for each month. For each of these models, we used daily pumped water and weather data from 2000 through 2015. The summary statistics for variables used in model development for each of these models are shown in Table 2 below.

Table 2: Variable Summary

Daily Models (N=5,844)					Monthly Models (N=192)				
Variable	**Mean**	**SE**	**Max**	**Min**	**Variable**	**Mean**	**SE**	**Max**	**Min**
PUMP (millions of gallons)	11.14	0.05	26.24	3.28	$PUMP_M$ (millions of gallons)	339.03	8.56	708.45	193.93
PRECIP (inches)	0.11	0.01	5.28	0.00	$PRECIP_M$ (inches)	3.31	0.19	12.89	0.00
T_{MAX} (°F)	79.92	0.19	112.00	31.00	T_{MMAX} (°F)	79.85	0.93	103.84	54.74
T_{MIN} (°F)	59.00	0.19	81.00	17.00	T_{MMIN} (°F)	58.93	0.92	78.03	35.58
T_{AVG} (°F)	69.46	0.19	93.50	25.50	T_{MAVG} (°F)	69.39	0.92	90.94	45.16
LAST2	0.37	0.01	1.00	0.00	$LAST2_M$	0.38	0.01	0.77	0.00
$DAYS_{SINCE}$	4.88	0.09	56.00	0.00	$DAYS_{MSINCE}$	13.56	0.54	4.90	53.55
$WEEK_{FREQ}$	1.69	0.02	7.00	0.00	$MONTH_{FREQ}$	7.35	0.25	18.00	0.00
WW_{INDEX}	53.41	0.26	93.50	0.00	MW_{INDEX}	53.05	1.04	85.98	22.41

A full explanation of each regression model, as well as the description for each variable used, can be seen in Appendix A. Below, in Table 3, are the results for the general model and the monthly models based on a monthly time-step.

Table 3: Weather Model Results – Monthly Time-Step[7]

Model	Coeff.	Variable	Partial R^2	P-value	Model R^2	F-value
General	-12.856				**0.9031**	**346.88**
	0.0112	MW_{INDEX}	0.7851	0.0001		
	0.0194	$MONTH_{FREQ}$	0.0565	0.0001		
	0.0073	*YEAR*	0.0518	0.0001		
	0.0018	$DAYS_{MSINCE}$	0.0070	0.0004		
	-0.0035	$PRECIP_{M}$	0.0027	0.0247		
February	2.0893				**0.2591**	**4.90**
	0.0039	T_{MMAX}	0.2591	0.0440		
March	2.0948				**0.6640**	**27.67**
	0.0067	MW_{INDEX}	0.6640	0.0001		
April	2.4007				**0.4824**	**13.05**
	0.0067	$DAYS_{MSINCE}$	0.4824	0.0028		
May	2.2151				**0.7996**	**25.93**
	0.0062	$DAYS_{MSINCE}$	0.5752	0.0019		
	0.0041	MW_{INDEX}	0.2244	0.0021		
June	0.0100				**0.9482**	**118.96**
	0.0281	T_{MMAX}	0.9249	0.0000		
	-0.0071	$PRECIP_{M}$	0.0233	0.0311		
July	-0.0672				**0.6883**	**30.91**
	0.0287	T_{MMAX}	0.6883	0.0001		
August	-0.1946				**0.7047**	**15.51**
	0.0306	T_{MMAX}	0.5784	0.0006		
	-0.0041	$DAYS_{MSINCE}$	0.1263	0.0347		
September	-0.0104				**0.6019**	**21.16**
	0.0328	T_{MAVG}	0.6019	0.0004		
October	2.1817				**0.6826**	**30.10**
	0.0068	MW_{INDEX}	0.6826	0.0001		
November	2.5072				**0.5066**	**14.37**
	-0.0157	$PRECIP_{M}$	0.5066	0.0020		
December	1.9889				**0.2540**	**4.78**
	0.0072	T_{MAVG}	0.2540	0.0465		

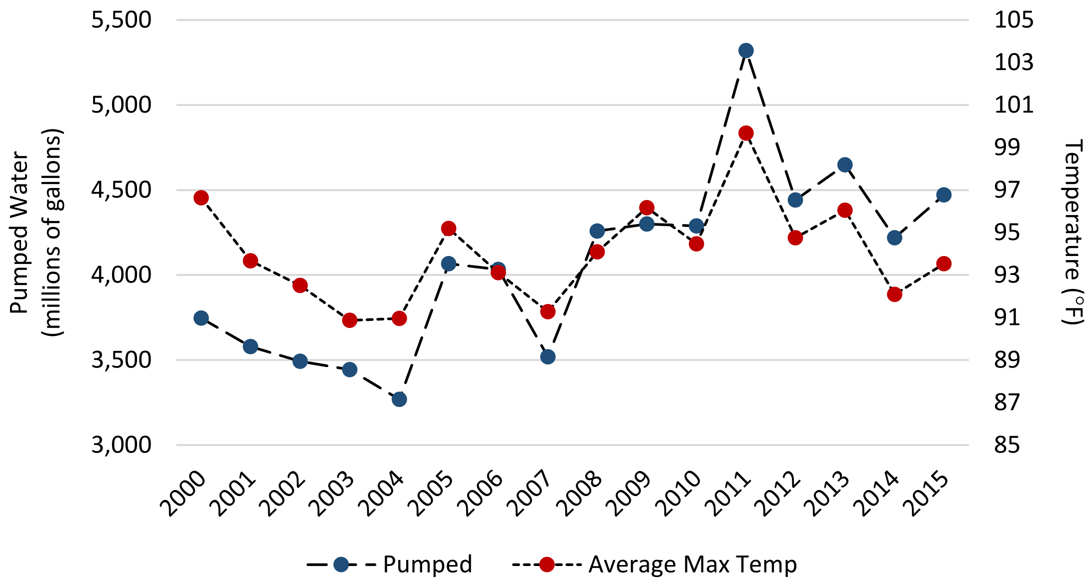
Weather Explains Much of the Summer Variability in Water Demand

As we can see from Table 3 above, water demand in different months is affected by different weather events. Specifically, during the months of June, July, August, and September a single temperature variable (either T_{MMAX} or T_{MAVG}) explains more than half of the variability in water demand. Moreover, in June 92.5% of water demand variability can be explained by maximum temperature (T_{MMAX}) alone. This is most likely

[7] January is absent from this table, because no significant weather variables could be identified; All monthly models were $\log_{10}$ transformed.

because it is a transition month from Spring to Summer. As we can see from Figure 2, average summer month (June – September) temperatures are seemingly significant drivers in yearly water demand (correlation coefficient of 0.77).

Figure 2: Average Maximum Summer Temperatures and Total Pumped Water



On another note, there is no weather forecasting model for January. This is because none of the variables we used to measure weather effects in other months were statistically significant in January. This is particularly useful to know when attempting to understand indoor versus outdoor water usage as we can reasonably assume that, in general, there is very little outdoor (and thus weather related) water demand that occurs during January.

To aid in future water demand planning, we have provided the table below summarizing the average value (between 2000 and 2015) of each weather variable listed in the general model for each month.

Table 4: Monthly Weather Variable Summary

Variable	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
T_{MMAX}	61.72	64.67	72.28	80.06	86.67	92.82	95.03	97.10	91.30	82.10	71.30	63.17
T_{MMIN}	40.78	43.96	51.13	59.26	66.71	73.15	74.70	75.07	69.96	60.26	49.98	42.24
MW_{INDEX}	37.66	38.79	45.09	54.88	58.05	62.83	68.76	70.57	63.08	54.23	44.12	38.56
$MONTH_{FREQ}$	8.31	8.13	8.38	6.44	7.56	7.38	5.94	5.63	6.56	7.38	8.25	8.31
$DAYS_{MSINCE}$	11.57	11.00	10.74	10.00	13.31	12.97	16.14	19.05	15.83	15.41	15.35	11.36
$PRECIP_{M}$	3.05	2.85	3.71	2.05	4.20	3.73	2.70	2.09	3.59	4.87	3.81	3.10

Other Factors May Affect Yearly Demand Predictions

It is important to note that, while weather is an important factor in predicting monthly water demand (mostly due to seasonality), other environmental factors (such as demographic changes and conservation efforts) may also impact a yearly water demand model. The results from a yearly causal demand model, including weather, are presented in Table 5 on the next page.

Table 5: Yearly Water Demand Environment Model

	Potential Models (1 – 5)				
Variable	**1**	**2**	**3**	**4**	**5**
$T_{SUMMERMAX}$	189.73***	161.50***	158.38***	155.50***	162.98***
POP		0.0269***	0.0202**	0.0130	0.0107
CONS			137.67	165.78	22.23
PRICE				701.42	
INFLATE					-1434.06*
Constant	-13656.97	-13492.12	-12636.04	-13247.17	-8642.72
Model R^2	0.7760	0.9431	0.9493	0.9520	0.9681
Model F-Stat	34.64	74.63	49.89	34.68	53.03
Model Sig	0.0002	<0.0000	<0.0000	0.0001	<0.0000

Note: * indicates significance above the 90% level, ** indicates significance above the 95% level, and *** indicates significance above the 99% level.

If the coefficients presented above were to be used for predictive purposes, then only Model 2 should be considered because it provides the best-fit model explaining 94.3 % (R^2 = 0.9431, p < 0.0001) of the variation in water pumped per year ($PUMP_Y$) with two significant variables $T_{SUMMERMAX}$ and POP, where $PUMP_Y = -13492.12 + 161.50 (T_{SUMMERMAX}) + 0.0269 (POP)$. In the other models, although the R^2 may be slightly higher, not all variables were significant, and therefore should not be used as a predictive model.

In the models above, the variables considered included: $T_{SUMMERMAX}$, an average of the maximum daily temperature during the months of June through September; POP, an estimated population of the City of College Station based on certificate of occupancy; CONS, an indicator variable representing years in which there was an active conservation program (starting in 2010; 1 = active, 0 = inactive); and the PRICE and INFLATE variables, both measure the lowest residential volumetric water rate; PRICE indicates the nominal price and INFLATE indicates the real price in 2016 dollars. Full model development is provided in Appendix B.

Table 5 presents the results of five separate regression models. Unlike previously discussed predictive weather models, the models presented in Table 5 were prepared for causal analysis. More commonly used in econometrics, causal analysis through the use of regression modeling attempts to determine whether a particular independent variable meaningfully affects the dependent variable. In the models above, the average summer maximum temperatures are very significant and impactful to annual water demands. However, when the price variables are included, population no longer seems significant.

This is most likely because the price variables and population have a very strong correlation (correlation coefficient above 0.80 or below -0.80) and thus a higher variance inflation factor. This is presumably due to time being a strong driver for both variables. Though this effects the significance of these two variables, it should not have an effect on coefficient estimation.

Some Billing Strategies May Be Diminishing Conservation Impact

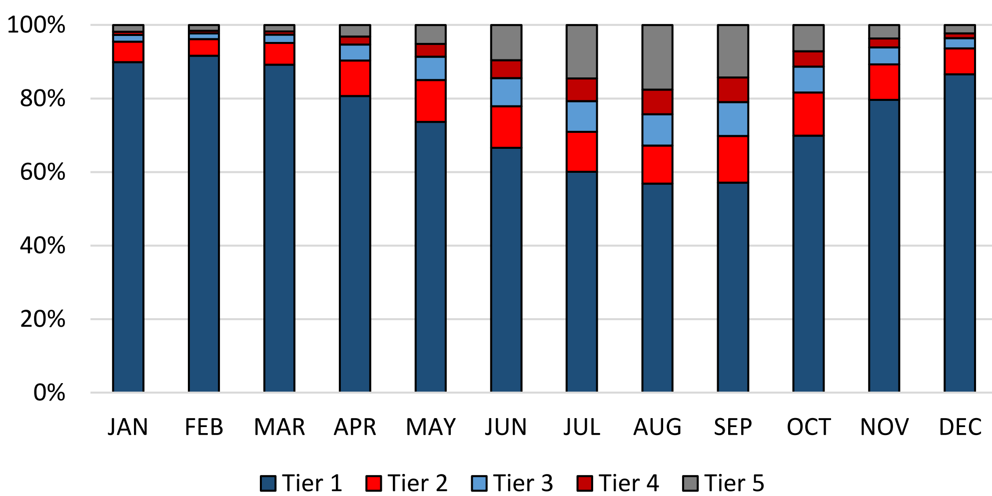
While conservation is not significant in our causal model, it may still be having an impact on water demand. For example, from 2010 to 2015, 874 high efficiency toilet rebates have been issued and 138 rain barrels have been installed. Over the course of these five years, this resulted in estimated savings of at least 5 million gallons of water. These conservation efforts should not be marginalized. However, to put this into perspective, Water Services has the capacity to produce 30 million gallons of water each day.

This lack of statistical significance of conservation (Table 5) may be partially attributed to fiscal procedures and strategies. For example, some College Station residents take advantage of the budget billing system. This system allows customers to pay the same amount of money every month for their utility bill, no matter what they actually consumed. Then each year, there is a "settle up" month in which any remaining balance is paid for that year's utility bills. This, along with an automatic bank draft service, allows customers to pay their utilities without ever having to look at their bill. These two fiscal procedures increase the risk of large water leaks going unnoticed and diminish the effects of conservation-oriented, block rates as well as other conservation-oriented programs.

Block Rates May Not Be Having Their Expected Impact on Conservation

In fiscal year 2008, the City changed its water rate structure to have five tiers of increasing block water rates for residential customers.[8] At each tier water is charged a different rate per thousand gallons.[9] According to the AWWA's Manual M-1 *Principals of Water Rates, Fees, and Charges*, increasing block rate structures are generally considered to be conservation-oriented. Also, the manual mentions that usage block sizes should correspond to the utility's usage patterns. We examined the percentage of customers that were in each tier on average throughout the year. The results are presented in Figure 3 below:

Figure 3: Average Monthly Water Usage by Tier

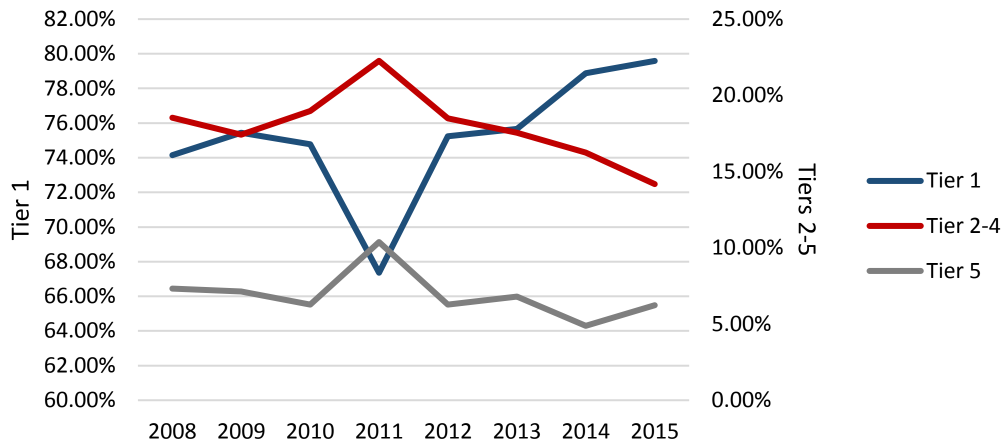


According to the Environmental Protection Agency, the average US family uses about 400 gallons of water per day. Over the course of a month, this is about 12 thousand gallons. As we can see from Figure 3, the largest tier of College Station residential customers is Tier 1 (0 – 10 thousand gallons per month). This is a little lower than the EPA's average usage. Moreover, it appears that this tier of customers has been growing in the past eight years (see Figure 4 on the next page).

[8] Tier 1 ranges from 1-10 thousand gallons, Tier 2 ranges from 11-15 thousand gallons, Tier 3 ranges from 16-20 thousand gallons, Tier 4 ranges from 21-25 thousand gallons, and Tier 5 includes all thousand gallons consumed at or above 26 thousand gallons.

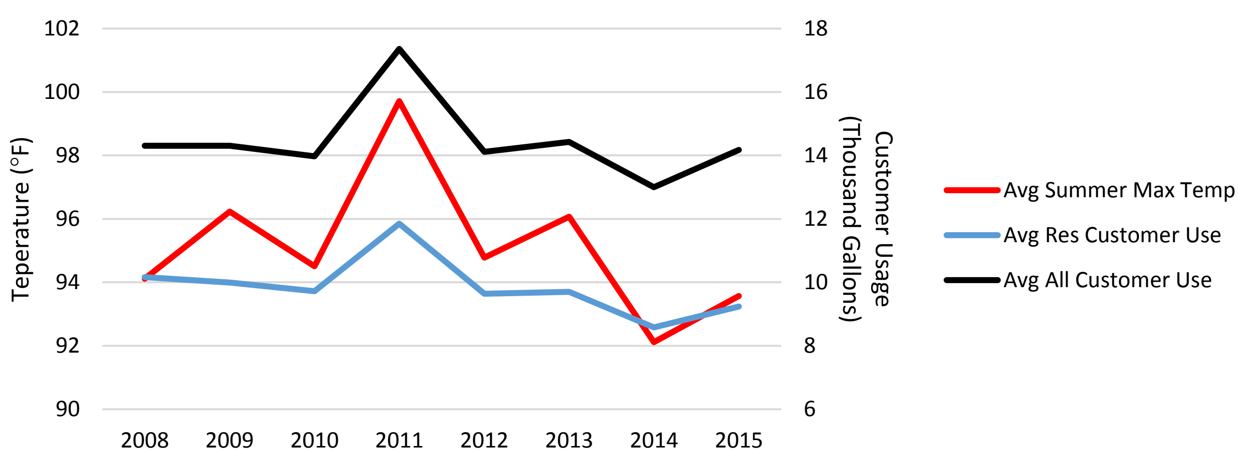
[9] Partial thousand gallons are not charged; for instance a customer can use an extra 999 gallons of water before the next thousand gallons is charged.

Figure 4: Residential Customer Tiers over Time[10]



Though this appears to suggest that block rates and conservation are having their intended effect on water usage, this may not be the case. As you can see from Figure 4 above, there was relatively little change in water usage from 2008 through 2010. However, when the 2011 drought happened, the higher water usage categories increased, especially Tier 5. This suggests that movement between water tiers has actually been driven due to weather patterns and not conservation efforts or rate structure. Due to the budget billing system and automatic bank draft service, we were unable to draw any conclusions about the rate structure from this analysis. However, if this wasn't the case, these patterns might suggest that the Tier 1 cap is too high to effectively act as a conservation-oriented price signal in this community.

Figure 5: Average Customer Water Usage and Average Summer Maximum Temperature



As we can see from Figure 5, there appears to be a strong correlation between average customer water usage and the average maximum summer (June – September) temperatures. In fact, the correlation between average customer use and average summer max temperature is above 0.9 for both customer sets. This would imply that, instead of conservation, weather is actually the strongest driving force in decreasing customer water usage, as we concluded from our multiple regression analysis. This supports

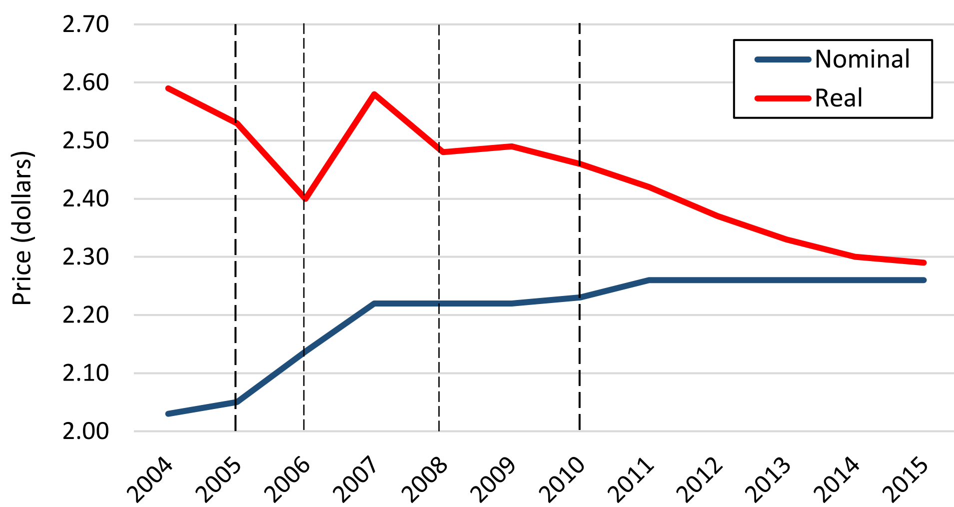
[10] Water Tiers 2 through 4 were combined to simplify the figure and because their trend lines were nearly the identical (correlation coefficient above 0.85).

the need for more thorough and complex analysis when trying to understand water consumption, as simple trend analysis can occasionally be misleading.

Inflation Must Be Considered to Understand the True Cost of Water

As we can see from Table 5, when price is included in a nominal sense in the yearly water demand regression model, it is not significant and has a positive coefficient. However, when price is included in the model as real 2016 dollars (INFLATE), then it is a significant variable (above the 90% level) and has a negative coefficient (Table 5). To understand this, it is important to note that real water price has been decreasing for the last twelve years. Essentially, this means that water consumption has taken up less and less of customer's purchasing power as time has gone on. This decrease in the real price of water may be part of the reason conservation effects have been limited — it has become increasingly cheaper to use more and more water over time. Figure 6 illustrates this further.

Figure 6: Lowest Volumetric Price of Water[11]



As we can see from Figure 6, the real and nominal price of water have been converging in the past ten years. This is due to declining US inflation rates. However, this trend may not continue into the future. If inflation rates increase without a change in water rates, effective price of water will increase. This may or may not change revenue streams, but an increase in real water prices may increase conservation incentives. Evidence of this can be seen when comparing Figure 4 and Figure 6. There was an increase in inflation rates between 2008 and 2009 that can be seen in Figure 6. In Figure 4, we can see a corresponding increase in the lower usage tiers, as well as a decrease in the highest usage tier.

Insignificant variables are still important. Though weather is the most significant factor in water demand, this does not mean that other factors should be ignored. According to *Forecasting Urban Water Demand*, price and conservation should both be significant demand influencers but, locally, they are not. This evidences a need for change in rate setting efforts, city-wide conservation efforts, or both. If water conservation is deemed a worthy objective by the College Station community, additional support should be given to Water Services in order to generate real change in usage.

[11] Volumetric Water rates could not be determined before 2004. Real prices are expressed in constant 2016 dollars. Dashed lines indicate years with water rate changes.

Very-Short- and Short-term Forecasting Evaluation

Generally, very-short-term forecasting is focused on optimizing water production by forecasting water demand for the current day. More specifically, this helps the water production operators stabilize production and increase energy efficiency. Water Services staff have an equation that they use to forecast daily water demand in this capacity. This equation is shown in the figure below:[12]

Figure 7: Water Services Daily Water Demand Forecast Formula

Previous Days Flow x **Weather Constant** x **City Activity** x **Day of Week** x **Rain Day Interval**

Previous Days Flow	Weather Constant	City Activity	Day of Week	Rain Day Interval
	Full Sunshine = 1.01	Normal = 1.0	Monday = 1.05	Day of = .98
	Partly Cloudy = 1.0	Home Football = 1.01	Tuesday = .98	1-4 days = 1.0
	Full Clouds = .99	Parents Weekend = 1.01	Wednesday = 1.01	4-12 days = 1.01
		Christmas Weekend = .99	Thursday = .98	>12 days = 1.02
		Thanksgiving Weekend = .99	Friday = 1.01	
			Saturday = .97	
			Sunday = .98	

Daily Forecasts Are Effective and in Alignment with AWWA Approaches

Water Services provided daily forecast data dating back to 2006. Not only did we find that these forecasts were accurate but we also found the methods utilized to be in general alignment with forecasting principles promulgated by the AWWA.

Water Services daily water demand equation makes intuitive sense. Many of the factors the AWWA advises to consider in forecasting water demand seem to be accounted for in this equation. For instance, the Previous Days Flow variable accounts for long-term time variant factors like seasonality and demographic changes. Furthermore, during most of the year, population does not change drastically from one day to the next. On the days that it does, the equation accounts for this in the City Activity variables. In addition, we were told that Monday, Wednesday, and Friday are high irrigation days traditionally. Therefore, it would then make sense that these days have higher "Day of Week" coefficients.

Most variables significantly affect daily water demand. In order to test if each of the variables included in the equation above were significantly affecting water demand, our office developed a predictive daily water demand model using multiple regression. The regression was developed using forward selection and a set of model comparison criteria to determine if each variable should remain in the model (see Appendix A, Table A-3).

[12] On Tues and Thurs, the Weather Constant coefficients follow this pattern: Full Sunshine = 1.0, Partly Cloudy – 0.99, and Full Clouds = 0.98 and the Rain Day Interval coefficients follow this pattern: day of = 0.98, 1-8 days = 1.0 and >8 days = 1.01.

In this way, we determined that all variables were significant influences on water demand except the Christmas, Thanksgiving, and Parents Weekend variables. Specifically, this means that average water demand is not significantly different during Christmas, Thanksgiving, or Parents Weekend than average water demand during the whole year, other things being equal. However, we excluded the Thursday, Saturday, and GAMEDAY variables from our predictive model because they were not very explanatory (explain less than 0.1% of variation). Table 6 provides the output of our final predictive regression model. A full explanation of variables and model development can be seen in Appendix C. Overall, our model shows that the amount of water pumped the previous day ($PUMP_{PREV}$) explains the largest proportion of the variation (92.5%), while other factors are much less important (all combined explaining only an additional 2%).

Table 6: Daily Water Demand Model – Replica Output Summary

Variable	Partial R^2	P-value	Variation Inflation Factor	Model R^2	F-value
$PUMP_{PREV}$	0.9256	<0.0001	2.25	0.9454	72649.9
MON	0.0038	<0.0001	1.15		316.1
WED	0.0042	<0.0001	1.15		370.6
FRI	0.0061	<0.0001	1.15		586.0
T_{MAX}	0.0030	<0.0001	2.03		300.6
SUN	0.0014	<0.0001	1.15		148.9
$DAYS_{SINCE}$	0.0013	<0.0001	1.22		143.5

Model R^2 = 0.9324; Mean of Squared Error = 1.175; 5844 Observations (days/month x 12 months x 16 years)

Water Services' daily forecasts are accurate. We compared the accuracy of Water Service's daily forecasts to our replica regression model (Table 6) as well as a simple model that forecasts current day water demand to be the same as the previous day's water demand. The results can be seen in Table 7:

Table 7: Very-Short-Term (Daily) Forecasting Comparison

Forecast Method	**Avg. Error Percentage**	**Standard Deviation**
Water Services Predictions	- 0.22%	7.80%
Previous Day Predictions	- 0.08%	8.51%
Daily Water Demand Model – Replica	- 0.34%	8.66%

As we can see, Water Services predictions have the second lowest average error percentage and the lowest standard deviation. This means that, assuming a normal distribution, 68% of Water Services' daily demand predictions are within ± 7.80% of the mean error percentage.

It is important to note, however, that the previous day's water demand accounts for 92.56% percent of the current day's water demand. This is supported by the almost nonexistent inaccuracy of the Previous Day Predictions model (average of -0.08% error). Though this simple forecast is extremely accurate, it fails to illustrate how and to what extent other variables impact water demand. This supports our finding that weather is the largest determinant in water demand variability, since the previous day variable takes into account the effects of seasonal weather factors. Due to this, it is not surprising that Water Service's daily forecasts are so accurate.

Subjective Forecasting Methodologies Present a Risk

Water Services very-short-term forecasting methodology's lower average error percentage and standard deviation are most likely due to prediction alterations made by operators due to their experience. We believe that despite the equation used by Water Services (Figure 6), much of their daily demand knowledge is subjective. According to staff, throughout the day water production operators monitor the weather and City radio channels for water demand influencers (rain, heat advisories, water main breaks, fire flow tests, etc.). Once these events are identified, SCADA[13] parameters are reset.

Since we were not provided with the calculations for each of Water Service's predictions, we attempted to identify each predictive coefficient used based on our historic data and the equation. Of the 2,922 daily observations we had (2008 through 2015), we were able to correctly identify all coefficients about 44% of the time. This suggests that, while Water Services is using their equation, they often use professional judgement and experience when predicting daily demand.

According to AWWA literature, subjective forecasting methods can be effective in some circumstances. More importantly, these methods have served Water Services well in predicting daily water demand. Currently, understanding the effects of these variables is largely based on years of experience with the City's wells and pumps system and demand influencers. This type of experience is invaluable and largely contributes to the accuracy of Water Services daily predictions. However, these subjective methods present a risk to continued accuracy if the department experiences employee turnover, and thereby loses this institutional knowledge, especially as the City's water demand patterns diversify and change as the City grows and becomes more urbanized.

Short-term Forecasting: There is a Lack of Connection to Water Demand Data

Generally, short-term forecasting is used for budgeting, revenue forecasting, and program evaluation. Since demand forecasting is often the first step in setting rates, we evaluated Water Service's water revenue forecasting as completed by the City's Fiscal Services Department.

According to City staff, water revenue projections are mostly based on historic billed water data, taking into account extreme weather conditions. Actual water usage is then monitored to evaluate the accuracy of the forecasts. We reviewed the accuracy of these revenue forecasts for the past four years, focusing in on water customer revenues.[14] The results are presented in Table 8.

Table 8: Revenue Forecasting Evaluation (percentage error)

Year	Residential	Commercial	Total
2012	-7.09%	-13.00%	-9.50%
2013	-4.13%	-12.71%	-7.68%
2014	10.54%	1.89%	6.84%
2015	3.20%	-1.39%	1.25%
Average:	0.63%	-6.30%	-2.27%
Stnd. Dev.	7.90%	7.69%	7.68%

[13] The City's SCADA system controls various equipment and monitors water transportation, distribution, and treatment.

[14] Excluding revenues from commercial/sale of effluent, other operating, investment earnings, and other non-operation categories.

As we can see over the past four years,[15] revenue forecasts have often varied from the actual revenue received (Table 8). However, these error rates have not significantly affected Water Services as a department. This may change if conservation efforts take a stronger effect in the future, the water system grows in size, the City becomes more urbanized, or land uses become less homogeneous. Additionally, over- or under-estimating revenues increases risk. For example, if revenues are over-estimated, the department may not have enough money to meet its debt obligations. On the other hand, if revenues are under-estimated, the department may increase its debt obligations without necessity. These risks do not seem to have actualized in the past, but Water Services should be aware of them — especially since information systems between departments are not currently integrated, resulting in revenue forecasts that do not take advantage of data Water Services utilizes for their long-term water demand forecasts.

Medium-term Forecasting Evaluation

Medium-term forecasting is used for sizing and staging treatment and distribution system improvements, deciding when and how to invest, and to set water rates. Over the course of our review, we found three consultant cost-of-service studies conducted in the last thirty years. One was completed by McCord Engineering in 1987, and the other two were completed by Black & Veatch, Inc. in 2002 and 2003. Due to lack of data and the antiquity of a 1987 cost-of-service study, we will focus on the Black & Veatch studies.

Consultant Cost-of-Service Studies do not Adequately Forecast Water Demand

According to the AWWA's M-1 *Principals of Water Rates, Fees, and Charges*, a cost-of-service study often has these typical objectives: (1) Fair and Equitable Cost Recovery, (2) Revenue Stability and Predictability, (3) Promotion of Conservation, (4) Simplicity in Understanding and Execution, and (5) Legality and Defendability.

The first step in this process is the Revenue Requirement Analysis, which compares aggregate costs to utility revenues to determine adequacy of existing rates. In order to determine revenue, a water utility must forecast water demand into the future as accurately as possible. We examined the water demand forecasting aspect of each Black & Veatch cost-of-service study (2002 and 2003). The results of our evaluation are presented below.

The Test period is too short. The consultant studies conducted in both 2002 and 2003 used only the previous fiscal year as their test period. This is considered common practice when developing a cost-of-service study, but it may not be the most accurate way to forecast water demand. Only analyzing a single year's consumption and customer growth fails to account for any abnormal weather or demographic events. For example, the 2011 drought caused about a 25% growth in average daily demand for that year. If this year alone was used as a test period, demand predictions would be extremely high when weather conditions normalized. On the other hand, land annexation can cause large, irregular increases in number of water customers, which could skew growth patterns.

[15] A four year period was chosen due to changes in categorization of revenue in FY 2011.

Data and analysis are lacking for a sectoral forecast. Under the consultant's methodology, the largest factors in water demand forecasting are customer growth and per capita usage. Both Black & Veatch reports state that City staff provided them with an estimated customer growth of approximately 3% per year. Though this may be true, it is not a metric that lends itself to a sectoral forecast. A key advantage to the sectoral forecast is the ability to separately identify customer growth patterns, thereby creating a more accurate water demand forecast. Assuming a singular growth rate for all customer classifications nullifies this benefit.

Moreover, the consultant forecasted water demand (and revenues) in both 2002 and 2003. However, we found no evidence that the water demands forecasted in 2002 were ever compared to the actuals in 2003 to check for accuracy. Instead, water demand was completely re-forecasted. We were unable to compare the Black & Veatch forecasts to billed actuals. However, we were able to estimate a 0.99% error percentage between the 2003 forecast in the 2002 report and the 2003 actual from the 2003 report. This is an acceptable error rate, however, error rates should continue to be evaluated as one moves further and further into a forecast.

Revenue classifications May Not Continue to Be Effective as Forecasting Categories

Through developing our sectoral forecast (see Appendix E), we discovered that different consumption patterns exist within revenue customer categories. It is not our intention to suggest that an average consumption trend should be developed for every customer. However going forward, revenue categories may not adequately distinguish between these consumption patterns.

We developed a sectoral forecast model based on both the revenue classifications and on state tax board property types. When the accuracy of these two models were compared through the use of a backcast, they do not appear to have significantly different error rates. See Table 9 for further illustration:

Table 9: Sectoral Backcast Comparison

	State Property Types		Revenue Classifications	
Year	Metered	Total	Metered	Total
2008	0.19%	-2.32%	-1.07%	-3.48%
2009	0.88%	-1.28%	0.00%	-2.09%
2010	2.39%	0.90%	1.88%	0.43%
2011	-16.77%	-17.15%	-16.90%	-17.27%
2012	2.56%	1.19%	2.75%	1.37%
2013	1.92%	-1.70%	2.43%	-1.24%
2014	8.15%	10.20%	9.02%	11.03%
2015	5.26%	5.46%	6.40%	6.53%
Mean:	0.57%	-0.59%	0.56%	-0.59%
SD:	7.46%	7.90%	7.78%	8.28%

The similarity in results of these backcasts can be explained by the historic homogeneity of City property types. Specifically, single family homes make up a little over 60% of all metered water locations and consume almost half of all metered water in the City.[16] Also, single family homes consume water over time similarly to one another even if at different levels of consumption. Simply, this means that uniformity in water customers has inadvertently made forecasting water demand easier.

[16] See Appendix E, Table E-2

There is evidence of City diversification. When we examined trends in consumption over time, we discovered that the percentage of total water consumption driven by single family properties has been decreasing. Meanwhile, the percentage of total water consumption driven by apartments and commercial properties has been increasing, especially in recent years. This shift in consumption coincides with the City's Comprehensive Plan, amended in 2015, which anticipates higher density development.

As the City continues to diversify, a more detailed sectoral forecast should be used to predict water demand. The sectoral forecasting method presented in Appendix E is just one of many ways a more comprehensive forecast can be performed. However, future sectoral forecasting would require more extensive data collection. Specifically, these factors should be considered before deciding on a sectoral forecasting methodology: 1) forecasting categories, 2) water consumption trends, and 3) future water rate strategies.

Departmental cooperation is necessary for effective data use. Also, it is important to note that any changes in forecasting method must be supported by both Water Services and Utility Billing. When selecting forecasting categories, water rate strategies and customer categorization should both be taken into account, otherwise, data cannot easily be used by both departments. For instance, customers can be classified as well as possible when forecasting in the long- or medium-term. However, if demand forecasting categories do not translate to rate strategies, this forecast cannot be easily used in budgeting, revenue forecasting, or rate setting efforts.

Long-term Forecasting Evaluation

In general, long-term forecasting is used for sizing system capacity and raw water supply. In the City, this type of forecasting is completed about every five years as part of a Water Master Plan update. We reviewed the water demand forecasting aspects of the Water Master Plan in both 2010 and 2016. Each plan was developed by a separate consultant: in 2010 HDR Engineering, Inc. and in 2016 Freese & Nichols Engineering, Inc. Water Services also periodically forecasts peak day water demand in the long term. The results of each of these evaluations are presented below:

Some Consultant Long-term Forecasting Methods Are Unclear

HDR and Freese & Nichols both used a variant unit-use approach to calibrate water demand forecasts. Specifically, land uses assigned to each parcel by Planning and Development Services were used to assign water usage factors and estimate water build out needs based on the comprehensive plan.

For both methodologies, the actual forecasting approach is fairly simple. First you project the unit of interest and then multiply times the per unit usage. For example, in a per capita model, the unit of interest is population. Population is then projected forward and multiplied by the per capita usage to estimate future water demands.[17] However in the variant approach, the more complex part is developing the unit of interest. In the 2010 Master Plan, this was the gallons per acre, while in 2016, this was the living unit equivalent (LUE).

[17] See Appendix D for an example of a simple per capita forecasting model.

While gallons per acre is a self-explanatory metric, LUE is more complex to describe. According to the 2016 Water Master Plan, 2014 LUEs were assigned with the following criteria: 1) single family parcels received 1 LUE, 2) duplex parcels received 2 LUEs, 3) commercial parcels received 6 LUEs per acre, and 4) apartment parcels received the number of LUEs equivalent to their number of units. A population density was then estimated based on population data received from City staff.

Unit of interest forecasting methods are unclear. We conducted a thorough review of the 2016 Freese & Nichols report and examined the analysis conducted by these consultants documented in EXCEL spreadsheets. We also conducted multiple interviews with City staff who worked closely with Freese & Nichols in developing their water demand forecast. Though it appears that forecasts were developed using professional judgment, we were unable to identify any stated criteria for LUE projections. Moreover, it is not made apparent in the HDR Master Plan how parcel land uses were combined with billed water demand.

Forecasts are purposefully inflated. When reviewing the 2010 Water Master Plan, we noted that the consultant reviewed five years of water billing records. However, within the plan, the consultants stated that they chose only one specific year to base their consumption patterns on due to it being "the driest and highest water demand year." This purposeful inflation can then be seen further when we review the accuracy of the consultant's demand forecasts.

Table 10: Forecast Accuracy Evaluation (Avg Day Demand)

Year	HDR Forecast	Freese & Nichols Forecast	Actuals	Error Percentage
2008	14.25	N/A	11.63	22.53%
2013	15.71	N/A	12.73	23.41%
2014	N/A	13.31	11.56	15.14%

We also reviewed the accuracy of the Freese & Nichols forecast, which showed signs of purposeful inflation as well. Similarly to the 2010 Water Master Plan, the 2016 Water Master Plan states that historical data was used as a basis for projecting water demands. Moreover, the actual usage factors utilized were slightly exaggerated for each consumption class. This inflation is further illustrated in Figure 9. However, while water demand forecasts used for rate setting and revenue projections should be as accurate as possible, this may not be true when forecasting demand for system capacity and supply.

Inflated long-term forecasts may be necessary in some circumstances. Underestimating water demand can inconvenience continued City development and growth if the water system cannot meet regulatory standards. For example, the Texas Commission on Environmental Quality may order a public water supply system to stop operations if it was constructed without prior approval or there is an imminent health hazard.[18] Also, according to the Texas Local Government Code[19] a municipality can issue a moratorium on development if there is a shortage of essential public facilities, such as a water utility, or a significant need for public facilities. Moreover, budgetary restrictions or time constraints may necessitate an earlier estimation of needed infrastructure improvements.

[18] Taken from the Texas Administrative Code, Part 1, Chapter 290 Subchapter D, Rule §290.40.

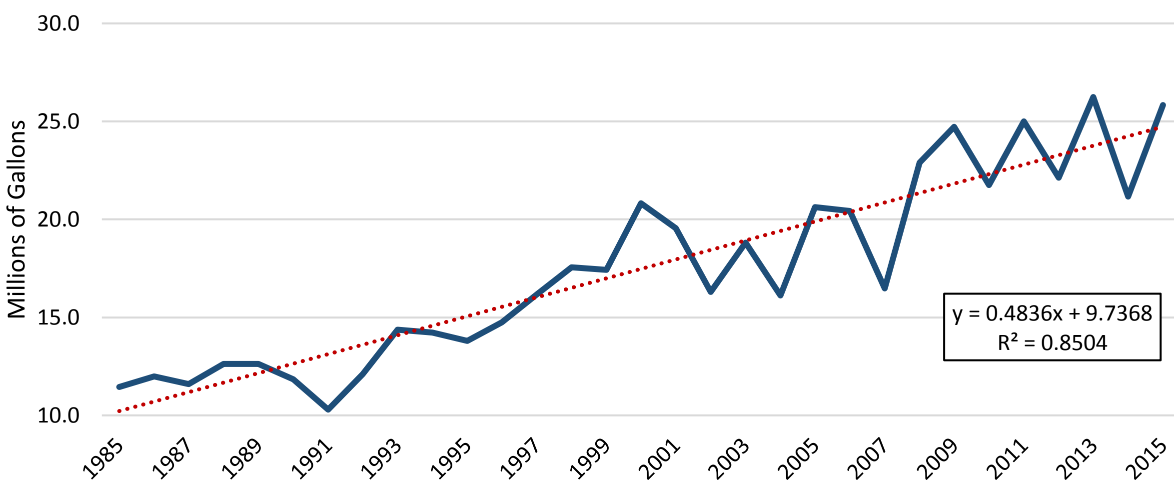
[19] Title 7, Subtitle A, Chapter 212, Subchapter E

In-House Long-term Forecasting Focuses on Peak Day Demand

When planning for system capacity, the most important forecasting measure is peak day demand. Specifically, a water utility must have the system capacity to serve all of its customers on the day each year that water demand is the highest. The City's Water Services' Department conducts this kind of forecasting in-house, which allows them to more effectively plan and adjust their capital improvements. For example, in combination with the Water Master Plan, this in-house forecasting has led to the scheduling of two capacity improvements for 2017 and 2022.

In-house time-series forecasting methods can be relatively accurate. Water Service's peak day forecasting methodology assumes that past growth in peak day water usage will be similar in the future. After reviewing peak day demand data from 1985 – 2015, there is an average yearly increase in peak day demand of about 0.5 million gallons of water. Figure 8 below illustrates that the linear trend of peak day demand represents relatively good fit to the actual data.[20] This indicates that a time-series methodology can be a relatively accurate way to forecast peak day demand if a linear trend is used. However, Water Service's does not use a linear trend to forecast, they use a percentage growth rate, which overstates peak day demand growth.

Figure 8: Peak Day Demand Growth (in millions of gallons)

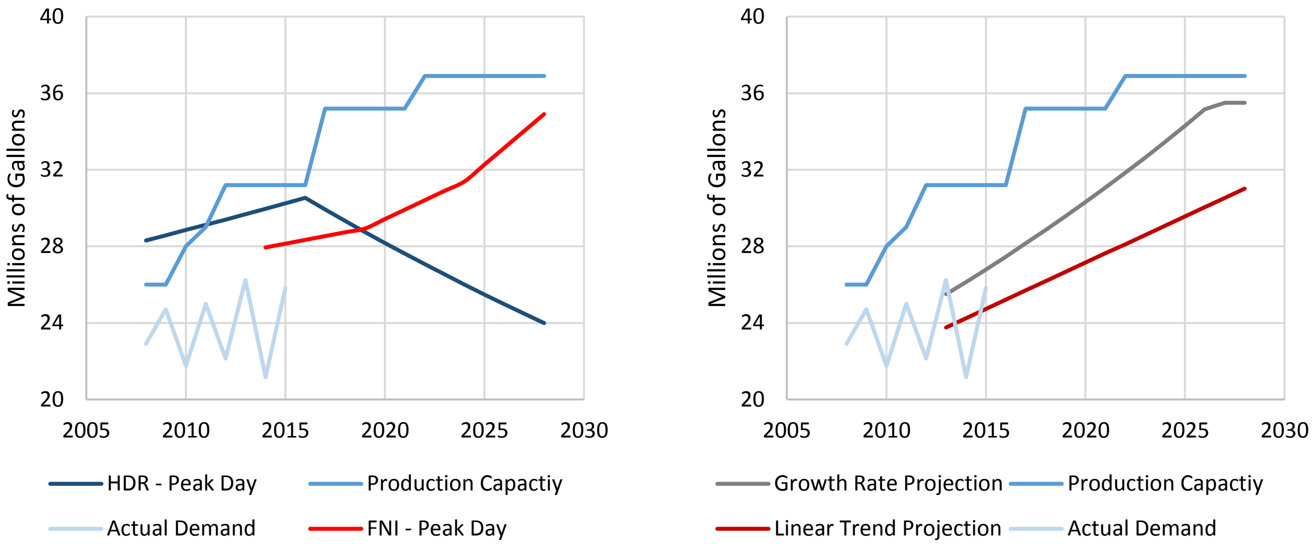


Forecasting method may be conventionalized to aid in understanding. When demand increases a similar amount each year, the growth rate is actually decreasing. This is illustrated further on the right in Figure 9 on the next page. As we can see, the range between the growth rate projection and the linear trend projection almost triples over a fifteen year forecasting horizon. This may have implications on planned system capacity improvements.

Though growth rate projections seem to over-estimate peak day demands, there may be reasons it is used. For instance, growth rates are commonly employed when communicating change over time. Using this type of metric may then make it easier for decision makers and other stakeholders to evaluate change in demand. Growth rates are also normalized, which may facilitate comparison and evaluation between other organizations of differing size and capacity.

[20] The R-squared value of 0.85 indicates that the model explains 85% the variability of the response data around its mean. For a time series model, r-squared values between 0.7 and 0.9 are considered to be quite good.

Figure 9: Peak Day Forecasting Comparisons



Consultant forecasting assumptions may be inconsistent with changes in the water demand environment. Figure 9 above shows four different peak day demand forecasts, each using a slightly different methodology. Both consultants use a sectoral, unit-use methodology to project peak day demands, however, our office and Water Services used two different time-series methodologies.[21] As we can see, the consultant forecasts (on the left side of the figure) are very distinctly inflated. As mentioned in previous sections, there may be reasonable justifications for this.

However, what is more striking is the divergence of the HDR and Freese & Nichols (FNI) projection lines. These two forecasts were made with two distinct sets of assumptions that were also separate from the forecasts made by our office and Water Services. This is perhaps most telling about consultant risks. Though this may have saved Water Service's time and staff, the HDR forecast assumes that after reaching 2016, peak day demands will decrease. Though this may have seemed reasonable at the time, it has essentially made this forecast ineffective after that point. This may be acceptable to Water Services, since a new forecast was developed by Freese and Nichols before this point was reached. However, it illustrates a need for constant adjustment and accuracy evaluation.

In-house forecasting has no connection to consultant forecasts. It is important to regularly evaluate the accuracy of forecasts so that they can be updated or adjusted if discrepancies are discovered. We found no evidence of any changes made to either consultant model after they completed their contracts. This may be due to resource or time limitations. Moreover, after examining an EXCEL spreadsheet provided by City staff, we found no evidence of any comparison between consultant projections and water demand actuals.

Reliance on Consultant Forecasts has Risks

There are several reasons that consultants should be hired. For example, an organization may occasionally lack the resources necessary to complete a certain project. In these cases, consultants can offer needed expertise and can take on projects that staff cannot complete due to time restraints or expertise. Currently in the City, the Water Services Department may not have the resources to complete

[21] Our office used the linear trend method (demonstrated in Figure 8). Water Services used a percentage growth rate method.

complicated in-house forecasting. However, it is necessary when hiring consultants to understand the risks associated with forecasts that either under-estimate or over-estimate water demand.

For instance, consultants may lack the institutional knowledge necessary to see all sides of a project. This may decrease the usefulness of their reports and could inadvertently harm City staff efforts. Also, when a consultant has completed their contract they move on to their next project, potentially leaving few people within City staff who completely understand the reports the consultants completed. This risk is greater when: 1) a consultant's work must be relied upon for a long period of time and 2) resource or time constraints may prevent an organization from recalibrating or updating consultant products in a timely manner.

General Conclusions

Using the book *Forecasting Urban Water Demand*, we identified four different forecasting types based on forecast horizon length. These are used for different purposes within a water utility and should be update and completed at different intervals. Over the course of our review, we identified all four forecasting activities, which were being covered through some combination of Water Services, Fiscal Services, and outside consultants.

City forecasts match their purpose. We found that all four forecasting types matched their purpose. For instance, the daily forecasting equation is used to help minimize the starting and stopping of pumps and wells and promote energy efficiency. Moreover, the medium- and long-term forecasts are used as part of rate studies and capital needs projections, respectively. The short-term forecast completed by Fiscal Services does not cover water demand, however, it is a revenue forecast.

Generally, forecasts are completed and updated in a timely manner. Very-short-term forecasts are conducted daily, and short-term forecasts are conducted yearly, which matches the prescribed time period for updating. On the other hand, there have only been, effectively, two medium-term forecasts conducted in the last thirty years. This lack of rate analysis may have contributed to City risk in the past and unintentionally diminished City conservation goals. Furthermore, long-term forecasts should be compared to actuals on a yearly basis, and updated if drastic changes are identified in water usage or significant forecasting mistakes are discovered. A summary of these findings can be seen in Table 11 below:

Table 11: City Water Forecasting Summary

Forecast Type	Forecast Application	Forecast Update Period
Long-Term	Projecting Capital Improvement Needs	Every 5 Years
Medium-Term	Identifying Rate Structure Changes	About 15 Years
Short-Term	Forecasting Revenues	Yearly
Very-Short-Term	Optimizing System of Wells and Pumps	Daily

Recommendation: *More complex methods should be investigated in the future as the City grows and diversifies.* In the past, the forecasting methods utilized by the City have been sufficient. Though each forecast has associated risks, these have not had significant impact on Water Services operations in the past. However, as the City grows and diversifies these risks may become more apparent. As this occurs, the City could benefit from more complex in-house water demand forecasting approaches, as it allows for more thorough analysis and increases institutional knowledge.

Other Considerations

1. *Climate changes may increase water demand variability in future years.* According to the National Wildlife Federation, the most visible effect of climate change is an intensification of weather extremes. Simply, this means that hot days will be hotter, rain will be heavier, droughts will be more severe, etc. When reviewing effects of weather, we noted that average maximum summer temperatures had the most significant effect on annual water demands. As climate change and weather intensification continues, this will cause an increase in water demands that the City should prepare for or at least consider. Including weather variables into forecast models is most effective for medium-term and longer-term forecasting.

2. *Rate Structure may not currently incentivize all of Water Services Goals.* While the current revenue structure has effectively been covering costs (about 32% of rate revenues cover all O&M expenses), it may not have been encouraging conservation. This is evidenced by a strong correlation between average customer water usage and average summer maximum temperatures and may be due to the real price of water decreasing as inflation rates decrease. It is likely that inflation rates will increase in the future, causing the real price of water to increase. This may strengthen conservation efforts; however, it should be noted that revenues may decrease as real price increases.

3. *Some forecasts should be updated more regularly.* In-house (including very-short-, short-, and long-term) forecasts are performed and updated in a timely manner. This is most likely due to these forecasts being completed by City staff, making them relatively cheaper to perform. However, medium-term forecasts are not updated or completed in a timely manner. This is evidenced by the thirteen years that have passed since the last cost-of-service study. Furthermore, we found evidence that forecasting accuracy in the medium-term was not evaluated. This is most likely due to lack of resources and a perception that the current rate structure is adequate for covering costs. Though this may be true, it does not necessarily indicate that rate structures are effective in accomplishing all department goals.

4. *Dependence on consultant forecasts have some risks.* Long-term consultant forecasts are developed in a timely manner (every five years), however, forecasts should be updated when changes or mistakes are realized. We found no evidence that long-term forecasts were regularly monitored for accuracy — unlike very-short- and short-term forecasts where such evidence was apparent. Nonetheless, it may not be cost-effective to bring back a consultant every time forecasting assumptions or methodology need to be adjusted. Furthermore, consultants may use forecasting methodologies that cannot be easily updated, adjusted, or recreated by City staff.

Appendix A: Water Demand Weather Predictive Model Development

We began creating our weather model by combining daily pumped water data with daily weather data obtained from the National Oceanic and Atmospheric Administration. We had complete years of data for 2000 through 2015 for these two datasets, leaving us with 5,844 observations. We then computed several other measures based on these weather observations. Descriptions of each daily weather variable follow:

Daily Variable Descriptions. The variable PUMP is the total amount of water pumped by Water Services in millions of gallons (excludes water lost to cooling).

Precipitation. The PRECIP variable is the total amount of precipitation that fell each day expressed in inches. The $LAST_2$ variable is an indicator variable that had a value of 1 if there had been a precipitation event in the last two days, and a value of 0 if there had not been. The $DAYS_{SINCE}$ variable counts the number of days since a precipitation event occurred. The $WEEK_{FREQ}$ variable counts the number of days in the last week with a precipitation event.

Temperature. The T_{MAX} variable is the highest temperature recorded each day. The T_{MIN} variable is the lowest temperature recorded each day. The T_{AVG} variable is the average of each day's maximum and minimum temperature.

Weather Index. The WW_{INDEX} is a modified version of an equation taken from "Determinants of Demand for Water Used in Texas Communities" by David R. Bell and Ronald C. Griffin of the Department of Agricultural Economics at Texas A&M University (2005). The Bell and Griffin equation is designed to capture seasonal changes for monthly weather models. Therefore, the equation was modified to fit a seven day time frame. As a result, the modified equation takes into account the average temperature of the day and the frequency of precipitation in the last week. In our monthly models (see Table A-5), we used the Bell and Griffin equation as published. The modified equation is presented below:

Equation A-1: Weekly Weather Index

$$WW_{INDEX} = \left(\frac{T_{MAX}+T_{MIN}}{2}\right) * \left(1 - \frac{WEEK_{FREQ}}{7}\right)$$

Summary statistics for each variable can be seen in Table A-1. We then generated a correlation matrix which can be seen in Table A-2 on the next page.

Table A-1: Daily Variable Summary

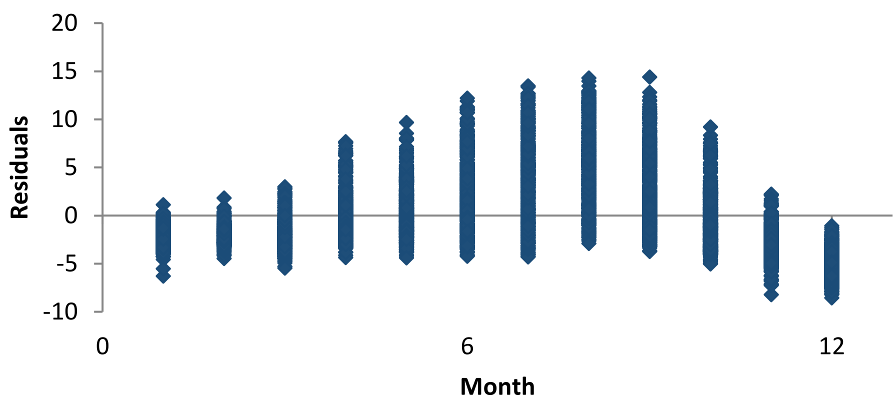
	PUMP	PRECIP	$LAST_2$	$DAYS_{SINCE}$	$WEEK_{FREQ}$	WW_{INDEX}	T_{MAX}	T_{MIN}	T_{AVG}
Mean	11.14	0.11	0.37	4.88	1.69	53.41	79.92	59.00	69.46
SE	0.05	0.01	0.01	0.09	0.02	0.26	0.19	0.19	0.19
SD	4.17	0.39	0.48	6.78	1.46	20.17	14.87	14.87	14.45
MIN	3.28	0.00	0.00	0.00	0.00	0.00	31.00	17.00	25.50
MAX	26.24	5.28	1.00	56.00	7.00	93.50	112.00	81.00	93.50

Table A-2: Daily Correlation Matrix

	PUMP	PRECIP	$LAST_2$	$DAYS_{SINCE}$	$WEEK_{FREQ}$	WW_{INDEX}	T_{MAX}	T_{MIN}	T_{AVG}
PUMP	1.00								
PRECIP	-0.10	1.00							
$LAST_2$	-0.29	0.13	1.00						
$DAYS_{SINCE}$	0.44	-0.20	-0.45	1.00					
$WEEK_{FREQ}$	-0.41	0.07	0.54	-0.50	1.00				
WW_{INDEX}	0.71	-0.07	-0.50	0.52	-0.84	1.00			
T_{MAX}	0.73	-0.09	-0.25	0.30	-0.27	0.71	1.00		
T_{MIN}	0.64	0.03	-0.16	0.19	-0.18	0.65	0.89	1.00	
T_{AVG}	0.70	-0.03	-0.21	0.25	-0.23	0.70	0.97	0.97	1.00

We developed three simple linear regressions using the PUMP variable as the response variable and the T_{MAX}, $DAYS_{SINCE}$, and $WEEK_{FREQ}$ weather variables as the explanatory variables. We also developed a simple linear regression using only the month as the explanatory variable. When we examined the residuals of this regression (Figure A-1), a clear pattern can be seen, implying that a separate weather model should be developed for each month.

Figure A-1: Monthly Residual Plot



Noting this, we developed a model on a daily time-step for each month. For each of these models, we added the variable most strongly correlated to the response variable (due to residual plot shapes, the PUMP variable was transformed by logarithm base 10 and used as the response variable in the models). Each remaining variable was then correlated with the residuals of the previous regression. The next most strongly correlated variable was then added to the model. This process continued until the regression with each added variable did not meet the criteria in Table A-3. At this point, the previous regression model (the one excluding the insignificant variable) was then considered to be the best-fit predictive model. The results from these analyses are presented in Table A-4 on the next page.

Table A-3: Parameter Selection Criteria

Statistic	Selection Basis
Residual Plots	Seemingly Random
R^2 (whole model)	Increase from previous model
Mean of Squared Error	Decrease from previous model
F-Statistic (whole model)	Significance F < 0.01
T-Statistic (all parameters)	P-value < 0.05
Variation of Inflation (all parameters)	VIF < 10

Table A-4: Daily weather models for each month (n>450 observations for each analysis)

Month	Coeff.	Variable	Partial R^2	P-value	Model R^2	F-Value
January	0.8289				**0.0615**	**11.8070**
	0.0018	$DAYS_{SINCE}$	0.0359	0.0009		
	-0.0009	T_{MIN}	0.0158	0.0052		
	0.0011	T_{MAX}	0.0098	0.0005		
February	0.8394				**0.1060**	**17.6999**
	0.0010	T_{MAX}	0.0756	0.0000		
	-0.0045	$WEEK_{FREQ}$	0.0215	0.0034		
	-0.0109	*PRECIP*	0.0089	0.0458		
March	0.9152				**0.2562**	**84.9161**
	0.0064	$DAYS_{SINCE}$	0.1771	0.0000		
	-0.0169	$WEEK_{FREQ}$	0.0791	0.0000		
April	0.7435				**0.4684**	**104.6514**
	0.0071	$DAYS_{SINCE}$	0.3545	0.0000		
	0.0042	T_{MAX}	0.0697	0.0000		
	-0.0016	T_{MIN}	0.0250	0.0005		
	-0.0100	$WEEK_{FREQ}$	0.0192	0.0005		
May	0.8518				**0.5847**	**137.9476**
	0.0027	WW_{INDEX}	0.3543	0.0000		
	0.0049	$DAYS_{SINCE}$	0.1150	0.0000		
	-0.0039	T_{MIN}	0.0575	0.0000		
	0.0032	T_{MAX}	0.0415	0.0000		
	-0.0221	*LAST2*	0.0164	0.0045		
June	-0.0179				**0.7149**	**237.7402**
	-0.0256	$WEEK_{FREQ}$	0.3202	0.0000		
	0.0032	$DAYS_{SINCE}$	0.3048	0.0000		
	0.0124	T_{MAX}	0.0686	0.0000		
	-0.0185	*LAST2*	0.0114	0.0199		
	0.0186	*PRECIP*	0.0099	0.0304		
July	-0.1560				**0.6576**	**473.4102**
	0.0144	T_{MAX}	0.3692	0.0000		
	-0.0350	$WEEK_{FREQ}$	0.2884	0.0000		
August	0.1781				**0.4713**	**146.1675**
	0.0110	T_{MAX}	0.3572	0.0000		
	-0.0199	$WEEK_{FREQ}$	0.1030	0.0000		
	-0.0219	*LAST2*	0.0111	0.0190		
September	0.4961				**0.4900**	**152.4266**
	0.0029	WW_{INDEX}	0.3411	0.0000		
	0.0065	T_{MAX}	0.1381	0.0000		
	-0.0017	T_{MIN}	0.0108	0.0230		
October	0.7492				**0.5250**	**135.6569**
	0.0030	WW_{INDEX}	0.3842	0.0000		
	0.0021	$DAYS_{SINCE}$	0.0722	0.0000		
	-0.0025	T_{MIN}	0.0506	0.0000		
	0.0034	T_{MAX}	0.0180	0.0000		
November	0.8043				**0.2699**	**58.6435**
	0.0021	WW_{INDEX}	0.2416	0.0000		
	-0.0012	T_{MIN}	0.0152	0.0069		
	0.0017	T_{MAX}	0.0131	0.0004		
December	0.7613				**0.0967**	**26.3919**
	0.0018	T_{AVG}	0.0548	0.0000		
	0.0041	$DAYS_{SINCE}$	0.0419	0.0000		

As we can see from Table A-4, weather appears to be a more significant factor in water demand as we move towards the summer months, peaking in June with a model R^2 value of 0.7149, which means that this model explains 71.5% of the variation in water pumped. This is also reflected in the weather models we developed on a monthly time-step, where June had the most explanatory model with a model R^2 value of 0.9482, the results of which can be seen in Table A-5. January is absent from Table A-5 because no significant weather variables could be identified.

Table A-5: Weather Model Results – Monthly Time-Step

Model	Coeff.	Variable	Partial R^2	P-value	Model R^2	F-value
General	-12.856				**0.9031**	**346.88**
	0.0112	MW_{INDEX}	0.7851	0.0001		
	0.0194	$MONTH_{FREQ}$	0.0565	0.0001		
	0.0073	*YEAR*	0.0518	0.0001		
	0.0018	$DAYS_{MSINCE}$	0.0070	0.0004		
	-0.0035	$PRECIP_M$	0.0027	0.0247		
February	2.0893				**0.2591**	**4.900**
	0.0039	T_{MMAX}	0.2591	0.0440		
March	2.0948				**0.6640**	**27.67**
	0.0067	MW_{INDEX}	0.6640	0.0001		
April	2.4007				**0.4824**	**13.05**
	0.0067	$DAYS_{MSINCE}$	0.4824	0.0028		
May	2.2151				**0.7996**	**25.93**
	0.0062	$DAYS_{MSINCE}$	0.5752	0.0019		
	0.0041	MW_{INDEX}	0.2244	0.0021		
June	0.0100				**0.9482**	**118.96**
	0.0281	T_{MMAX}	0.9249	0.0000		
	-0.0071	$PRECIP_M$	0.0233	0.0311		
July	-0.0672				**0.6883**	**30.91**
	0.0287	T_{MMAX}	0.6883	0.0001		
August	-0.1946				**0.7047**	**15.51**
	0.0306	T_{MMAX}	0.5784	0.0006		
	-0.0041	$DAYS_{MSINCE}$	0.1263	0.0347		
September	-0.0104				**0.6019**	**21.16**
	0.0328	T_{MAVG}	0.6019	0.0004		
October	2.1817				**0.6826**	**30.10**
	0.0068	MW_{INDEX}	0.6826	0.0001		
November	2.5072				**0.5066**	**14.37**
	-0.0157	$PRECIP_M$	0.5066	0.0020		
December	1.9889				**0.2540**	**4.78**
	0.0072	T_{MAVG}	0.2540	0.0465		

*In these models, $PUMP_M$ variable was LOG_{10} transformed, except for the General model.

Each model presented in Table A-5 was developed using the same statistical methodology as the daily model. It is important to note that each daily model per month had between 452 and 496 observations for each month (days/month x 16 years); meanwhile, each monthly model per month had 16 observations (16 years, 2000-2015) while the General model had 192 observations (16 years x 12 months). This would partially explain the discrepancy between the month based predictor variables and the day based predictor variables, as a large number of observations can sometimes cause a model to be "over fit" to a particular set of data. Noting this, we felt that developing the models in both ways would allow us to better identify the most significant variables. Below is a description of each monthly variable used to develop the models in Table A-5:

Monthly Variable Descriptions. The $PUMP_M$ variable is the total amount of water pumped by Water Services each month in millions of gallons (excludes water lost to cooling).

Precipitation. The $PRECIP_M$ is the total amount of precipitation that fell each month expressed in inches. The $LAST_{M2}$ variable is the average of the $LAST_2$ variable for each month; the $LAST_2$ variable is an indicator variable that had a value of 1 if there had been a precipitation event in the last two days, and a value of 0 if there had not been. The $DAYS_{MSINCE}$ variable was calculated for each day by identifying the highest $DAYS_{SINCE}$ value in the last thirty days for each day and then averaging over each month. The $MONTH_{FREQ}$ variable counted the number of days a precipitation event occurred each month.

Temperature. The T_{MMAX} variable is the maximum temperature of each day averaged for each month. The T_{MMIN} variable is the minimum temperature of each day averaged for each month. The T_{MAVG} variable is the average of the T_{MMAX} and T_{MMIN} variables for each month.

Weather Index. The MW_{INDEX} is a monthly version of the WW_{INDEX} and is taken from "Determinants of Demand for Water Used in Texas Communities" by David R. Bell and Ronald C. Griffin of the Department of Agricultural Economics at Texas A&M University (2005). The equation below was used to calculate this variable:

Equation A-2: Monthly Weather Index

$$MW\ Index = \left(\frac{T_{MMAX}+T_{MMIN}}{2}\right) * \left(1 - \frac{MONTH_{FREQ}}{DAYS}\right)$$

Table A-6 is a summary of the monthly variables described above and a correlation matrix of each variable is presented in Table A-7.

Table A-6: Monthly Variable Summary

	$PUMP_M$	$PRECIP_M$	$LAST_{M2}$	$DAYS_{MSINCE}$	$MONTH_{FREQ}$	MW_{INDEX}	T_{MMAX}	T_{MMIN}	T_{MAVG}
Mean	339.03	3.31	0.38	13.56	7.35	53.05	79.85	58.93	69.39
SE	8.56	0.19	0.01	0.54	0.25	1.04	0.93	0.92	0.92
SD	118.65	2.63	0.16	7.45	3.50	14.47	12.88	12.79	12.76
MIN	193.93	0.00	0.00	4.90	0.00	22.41	54.74	35.58	45.16
MAX	708.45	12.89	0.77	53.55	18.00	85.98	103.84	78.03	90.94

Table A-7: Monthly Correlation Matrix

	$PUMP_M$	$PRECIP_M$	$LAST_{M2}$	$DAYS_{MSINCE}$	$MONTH_{FREQ}$	MW_{INDEX}	T_{MMAX}	T_{MMIN}	T_{MAVG}
$PUMP_M$	1.00								
$PRECIP_M$	-0.27	1.00							
$LAST_{M2}$	-0.57	0.52	1.00						
$DAYS_{MSINCE}$	0.50	-0.23	-0.51	1.00					
$MONTH_{FREQ}$	-0.51	0.60	0.95	-0.43	1.00				
MW_{INDEX}	0.86	-0.38	-0.77	0.46	-0.76	1.00			
T_{MMAX}	0.84	-0.11	-0.41	0.35	-0.36	0.87	1.00		
T_{MMIN}	0.78	0.00	-0.29	0.25	-0.23	0.80	0.98	1.00	
T_{MAVG}	0.82	-0.05	-0.35	0.30	-0.29	0.84	0.99	0.99	1.00

Appendix B: Water Demand Environment Causal Model Development

We compiled daily pumped data into monthly and then yearly time-steps (Table B-1) based on data from 2004 through 2015 (this time period was chosen due to the lack of complete water rate data before 2004). Summary statics for this data set can be seen below. A correlation matrix of the data set was also generated and can be seen at the end of Appendix B (Table B-4 and Table B-5).

Table B-1: Yearly Weather Model (Obs = 12)

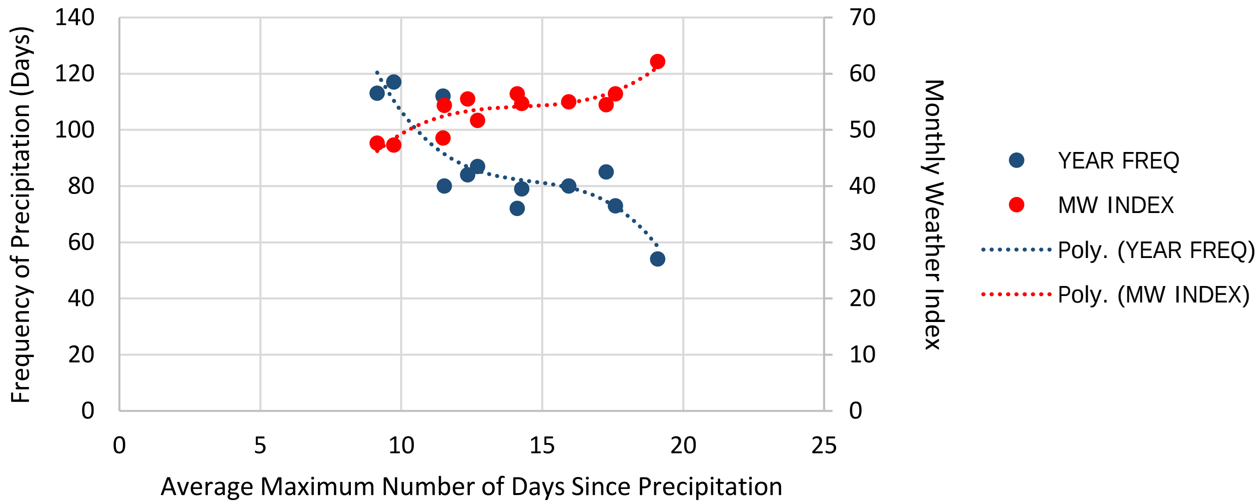
Variable	Average	Stnd Err	Stnd Dev	Min	Max
$PUMP_Y$	4235.97	150.03	519.71	3268.14	5320.09
$PRECIP_Y$	39.04	3.36	11.63	19.90	58.31
$DAYS_{YSINCE}$	13.77	0.92	3.18	9.15	19.10
$YEAR_{FREQ}$	86.33	5.42	18.78	54.00	117.00
T_{YMAX}	80.12	0.46	1.59	77.73	83.42
T_{YMIN}	59.33	0.26	0.91	57.77	61.17
T_{YAVG}	69.73	0.32	1.12	67.75	71.72
$T_{SUMMERMAX}$	94.31	0.70	2.41	90.98	99.71
YW_{INDEX}	53.68	1.24	4.28	47.30	62.15
POP	92963.50	2397.62	8305.58	80214.00	106465.00
POP_{Δ}	2.60	0.29	1.02	1.34	4.50
STUDENT	31889.85	669.71	2319.94	29302.91	36945.79
$STUDENT_{\Delta}$	2.78	0.89	3.07	-0.68	10.40
UNEMP	4.64	0.26	0.89	3.30	6.30
CONS	0.42	0.15	0.51	0.00	1.00
BARREL	11.08	5.25	18.20	0.00	63.00
TOILET	71.00	41.24	142.85	0.00	485.00
SAVINGS	445.49	283.00	980.36	0.00	3493.35
PRICE	2.20	0.02	0.08	2.03	2.26
INFLATE	2.44	0.03	0.10	2.29	2.59

The variables were then divided into two sets. One only included weather variables (Weather Set) and the other included demographic, economic, and conservation variables (Environment Set). We began with the weather set by graphing each explanatory variable with the response variable. Most notable was the polynomial shape to the $T_{SUMMERMAX}$. Due to this, a $T_{SUMMERMAX}^2$ term was computed and graphed.

Each pair of weather variables were then graphed. For the $PRECIP_Y$ variable, there were obvious trends in each graph (scatter plot) except for the T_{YMAX} and T_{YMIN} variables. However, not all of these trends were highly linear. Correlation coefficients indicate a strong relationship between the $PRECIP_Y$ variable and the $YEAR_{FREQ}$ and the YW_{INDEX} variables. Scatter plots reinforce these coefficients and indicate that a model should not include all three of the variables. This is expected, as the total precipitation should be affected by the number of times it rains, and the monthly weather index includes the number of times it rains (as frequency) into the computation.

The $DAYS_{YSINCE}$ variable appears to be linear and most strongly correlated with the $T_{SUMMERMAX}$ variable. However, notably, there appears to be a cubic shape to the scatter plot of the $DAYS_{YSINCE}$ and $YEAR_{FREQ}$ variables as well as the YW_{INDEX} variable. These plots are shown on the next page:

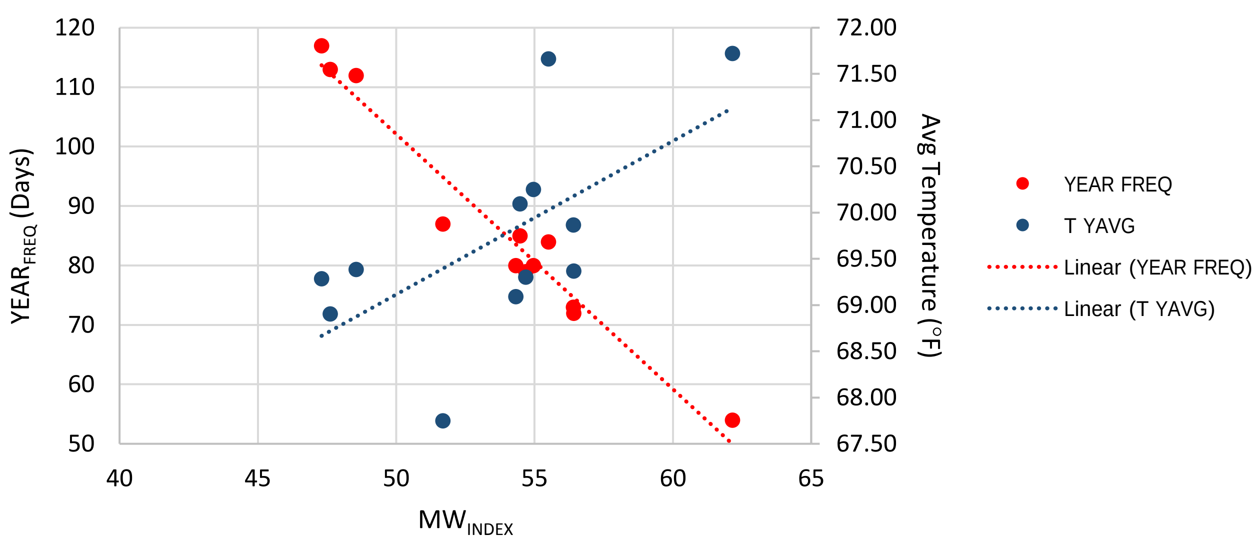
Figure B-1: $DAYS_{YSINCE}$ Scatter Plot



As we can see, there are very strong third degree polynomial relationships between these variables (Figure B-1). Although none of these are our variable of interest, it is still important to note relationships like these when selecting variables for inclusion in the regression model, especially since it is not a predictive model.

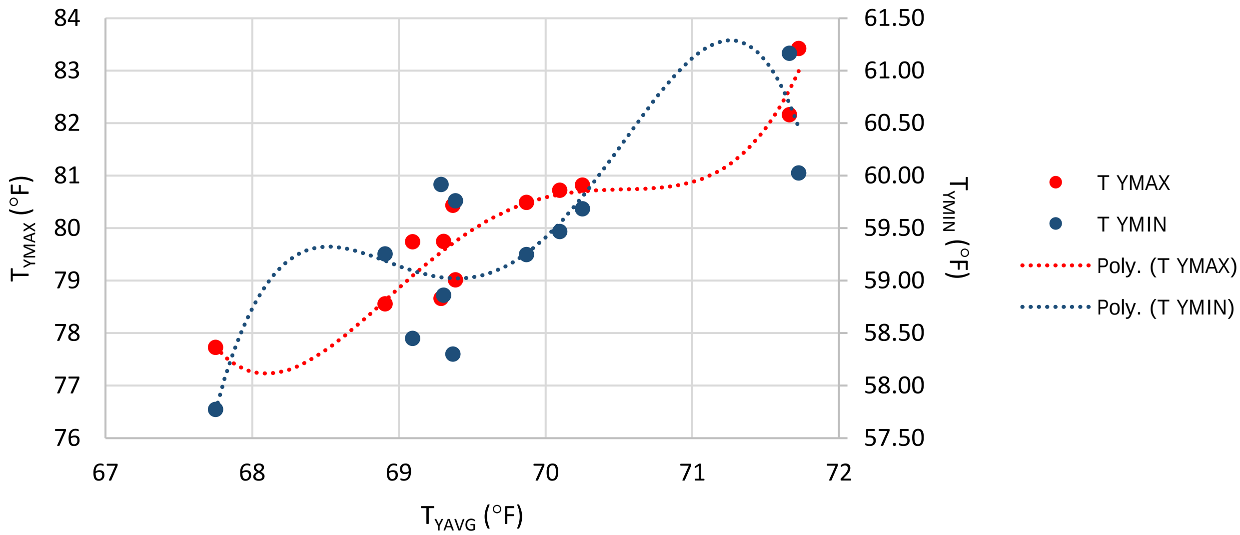
When graphed with other variables, the $YEAR_{FREQ}$ variable is not visually correlated with any temperature variables. However, the YW_{INDEX} scatter plot is almost a straight line (Figure B-2), which is indicative of how the YW_{INDEX} is calculated. It is also evidence of the impact of type of weather event on the YW_{INDEX}. As we can see from Figure B-2, the frequency of precipitation events seems to have a much stronger impact (i.e., tighter fit to the linear line) in the value of the YW_{INDEX} than the average temperature.

Figure B-2: YW_{INDEX} Variable Make Up



When the T_{YMAX} variable is graphed with other weather variables, they support the strong correlations between T_{YMAX} and $T_{SUMMERMAX}$, T_{YAVG}, and YW_{INDEX} identified in the correlation matrix. Interestingly, a fourth degree polynomial trend line fits each of these scatter plots best. Though this may be a slight over fit to the data, it could also be indicative of the effect average maximum daily temperature has on the temperature during the rest of the day. For example, the duration of high temperatures most likely matters to water demand. The T_{YAVG} and T_{YMAX} graph may be indicative of some sort of duration term, since the T_{YAVG} variable is not truly the average daily temperature, but it is instead an interaction term between the maximum daily temperature and minimum daily temperature variables. This is an important factor to note as we move forward with variable selection. This interaction is easier to see in the figure below (Figure B-3), which shows the almost inverse relationship between the maximum and minimum temperatures even as they both trend upwards.

Figure B-3: Temperature Interaction Term (T_{YAVG}) Make Up



Finally, there is also a fairly strong correlation between the YW_{INDEX} variable and the $T_{SUMMERMAX}$ variable when the scatter plot is graphed. This relationship is expected due to temperatures being a factor included in the YW_{INDEX}. Also, both variables are strongly correlated with the $PUMP_Y$ variable, so this may be inflating the correlation between these two.

The Environment Set of variables include demographic, economic, and conservation variables. We plotted each variable with the response variable.

The demographic variables, POP and STUDENT have a positive trend, this is what we would intuitively expect from population growth variables. For economic variables, there is a positive trend for all except the INFLATE variable. This is most likely due to declining inflation rates since 2004 with few increases in nominal price. Intuitively, we would not expect the UNEMP and PRICE variables to have positive trends, however, this is most likely due to the rapid growth rates in College Station during this time period. Also, all conservation variables appear to have a positive trend, which may be due to population increases over time as well.

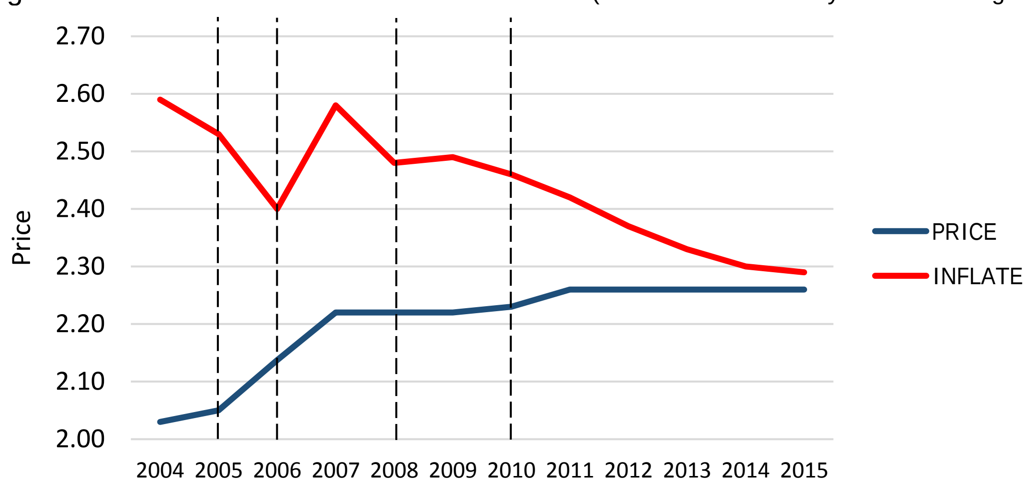
Next we look at the scatter plots with the POP variable. As we would expect, there is a strong positive, linear correlation between the POP and STUDENT variables. Other than that, the scatter plots with other variables tend to look like a time-series plot, indicating that population has a strong time factor driving it. This may disrupt, or cause false collinearity with other variables with strong time factors. The STUDENT variable is similar though less extreme. This may indicate a need to include a different type of variable, such as the growth rates of both the overall population and the student population when developing predictive model. By using growth rates, there will be no undue correlation between population growth and other variables with large time factors.

Most other graphs result in no visually striking correlation. There is a positive correlation between the TOILET and SAVINGS variables, this is most likely due to the number of toilets (and rain barrels) having a direct influence on the number of additional estimated saved gallons. This may indicate that the TOILET and BARREL variables should not be included in a model with the SAVINGS variable. Also, the CONS and INFLATE variables are correlated, but this is most likely due to time factors.

As one might image, there are relationships within the conservation variable set. It is important to note that the CONS variable is an indicator variable (which acts like a switch), signaling which years the conservation programs were in effect. It may not be necessary to include this variable if the other conservation variables are included in our model. Also as mentioned previously the BARREL and TOILET variable directly influence the SAVINGS variable.

Finally, when graphed as a scatter plot, there is no real visual relationship between the PRICE and INFLATE variables. However, it is interesting to note how they have changed over time. As we can see, real price (INFLATE) has been decreasing over the past twelve years, even as rates have increased. Though nominal price (PRICE) is currently as close as ever to real price, this is due to low inflation.

Figure B-4: Lowest Volumetric Price of Water (dashed lines indicate year rates changed)



Now that we have a better understanding of how each variable interacts with the others within each data set, we began selecting variables. To do this, we examined each set of variables (weather, demographic, economic, and conservation) and determined which variables have the strongest f-stat performance and p-value based on a simple linear regression. From each set of variables the following were chosen: $T_{SUMMERMAX}$, POP, CONS, and PRICE. However, we also ran a regression replacing the PRICE variable with the INFLATE variable. The CONS variable was chosen over the SAVINGS variable due to SAVINGS being an inadequate measure. The results of the regressions are shown in Table B-2 on the next page.

Table B-2: Yearly Water Demand Regression Results

	Potential Models (1 - 5)				
Variable	**1**	**2**	**3**	**4**	**5**
$T_{SUMMERMAX}$	189.73***	161.50***	158.38***	155.50***	162.98***
POP		0.0269***	0.0202**	0.0130	0.0107
CONS			137.67	165.78	22.23
PRICE				701.42	
INFLATE					-1434.06*
Constant	-13656.97	-13492.12	-12636.04	-13247.17	-8642.72
Model R^2	0.7760	0.9431	0.9493	0.9520	0.9681
Model F-Stat	34.64	74.63	49.89	34.68	53.03
Model Sig	0.0002	<0.0000	<0.0000	0.0001	<0.0000

Note: * indicates significance above the 90% level, ** indicates significance above the 95% level, and *** indicates significance above the 99% level.

If the coefficients presented above were to be used for predictive purposes, then only Model 2 should be considered because it provides the best-fit model explaining 94.3 % (R^2 = 0.9431, $p < 0.0001$) of the variation in water pumped per year ($PUMP_Y$) with two significant variables $T_{SUMMERMAX}$ and POP, where $PUMP_Y = -13492.12 + 161.50\ (T_{SUMMERMAX}) + 0.0269\ (POP)$. In the other models, although the R^2 may be slightly higher, not all variables were significant, and therefore should not be used as a predictive model.

Variable Descriptions. The $PUMP_Y$ variable is the total amount of water pumped by Water Services each year expressed in millions of gallons (excludes water lost to cooling).

Precipitation. The $PRECIP_Y$ variable is the total amount of precipitation that fell during each year expressed in inches. The $DAYS_{YSINCE}$ variable is maximum number of days without rain in the last thirty days averaged over each year. The $YEAR_{FREQ}$ variable is the total number of days each year with a precipitation event above 0.01 inches.

Temperature. The T_{YMAX} variable is the daily maximum temperature averaged over each year. The T_{YMIN} variable is the daily minimum temperature averaged over each year. The T_{YAVG} variable is the daily average of the minimum and maximum temperature averaged over each year. The $T_{SUMMERMAX}$ variable is the daily maximum temperatures in June through September averaged for each year.

Weather Index. The YW_{INDEX} was an equation take from “Determinants of demand for Water Used in Texas Communities” by David R. Bell and Ronald C. Griffin of the Department of Agricultural Economics at Texas A&M University (2005). Equation B-1 is the equation presented in their paper, where CI is calculated for each month and then averaged over each year to equal the YW_{INDEX}.

Equation B-1: YW_{INDEX} Equation

$$CI = \left(\frac{\max Temp + \min Temp}{2} \right) * (1 - Frequency\_of\_Rain)$$

Demographics. The POP variable is the estimated College Station population for each year based on certificates of occupancy. The POP_{Δ} variable is the growth rate in estimated College Station population expressed in percentage. The STUDENT variable is the weighted average of Texas A&M enrollment over each calendar year. The $STUDENT_{\Delta}$ variable is the growth rate in the weighted average of Texas A&M enrollment expressed in percentage.

Conservation. The CONS variable is an indicator variable, expressed as a 1 in years with an active water rebate program (2010 – 2015) and a 0 in years without an active water rebate program (2004 – 2009). The BARREL variable is the number of additional rain barrels rebated over the course of a year. The TOILET variable is the number of additional high efficiency toilets rebated over the course of a year. The SAVINGS variable is the estimated amount of additional water saved due to additional rebated rain barrels and toilets expressed in millions of gallons. Assumptions for estimated savings calculations are summarized below:

Table B-3: Estimated Rebate Savings Assumptions

High Efficiency Toilets		
Replaced Toilet Age	Uses per Day	Saved Flush Volume
1950 - 1980	5.1	3.72
1980 - 1994	5.1	2.22
Rain Barrels		
Gallons of Rain per Year		41,040
Rainfall collected per Barrel		5.00%
Annual Savings per Barrel		2,052

Economy. The UNEMP variable is the average unemployment rate taken from the US Bureau of Labor Statistics for College Station. The PRICE variable is the nominal price of water at the lowest volumetric residential water rate. When a rate change occurred (usually starting in October), a weighted average of water price was calculated for that year. The INFLATE variable is the PRICE variable, expressed in real (inflation adjusted) 2016 dollars.

Table B-4: Environment Set Correlations

	$PUMP_Y$	POP	POP_Δ	STUDENT	$STUDENT_\Delta$	CONS	BARREL	TOILET	SAVINGS	UNEMP	PRICE	INFLATE
POP	0.66	1.00										
POP_Δ	-0.21	0.02	1.00									
STUDENT	0.50	0.95	0.15	1.00								
$STUDENT_\Delta$	-0.43	0.01	0.29	0.22	1.00							
CONS	0.65	0.81	-0.21	0.78	0.06	1.00						
BARREL	0.35	0.48	-0.51	0.34	-0.19	0.66	1.00					
TOILET	0.76	0.31	-0.36	0.18	-0.21	0.52	0.15	1.00				
SAVINGS	0.74	0.30	-0.29	0.20	-0.18	0.53	0.17	0.94	1.00			
UNEMP	0.44	0.04	-0.71	-0.17	-0.37	0.00	0.27	0.54	0.45	1.00		
PRICE	0.65	0.87	0.04	0.71	-0.24	0.63	0.46	0.37	0.35	0.17	1.00	
INFLATE	-0.62	-0.84	0.07	-0.84	0.03	-0.81	-0.50	-0.23	-0.22	0.10	-0.65	1.00

Table B-5: Weather Set Correlations

	$PUMP_Y$	$PRECIP_Y$	$DAYS_{YSINCE}$	$YEAR_{FREQ}$	T_{YMAX}	T_{YMIN}	T_{YAVG}	$T_{SUMMERMAX}$	MW_{INDEX}
$PUMP_Y$	1.00								
$PRECIP_Y$	-0.52	1.00							
$DAYS_{YSINCE}$	0.57	-0.66	1.00						
$YEAR_{FREQ}$	-0.73	0.84	-0.82	1.00					
T_{YMAX}	0.66	-0.52	0.69	-0.72	1.00				
T_{YMIN}	0.05	0.22	0.07	0.12	0.57	1.00			
T_{YAVG}	0.49	-0.28	0.52	-0.47	0.94	0.81	1.00		
$T_{SUMMERMAX}$	0.88	-0.66	0.75	-0.79	0.81	0.18	0.65	1.00	
MW_{INDEX}	0.76	-0.80	0.84	-0.98	0.85	0.08	0.64	0.86	1.00

Highlighted cells indicate evidence of multicollinearity; across set correlations were calculated but not shown because no evidence of multicollinearity was found.

Appendix C: Very-Short-Term Replica Model Development

After discussing very-short-term (day-to-day) water demand forecasting with Water Services staff, we received three equations (combined into Figure 4 in the report on page 17) used to forecast daily water demand. We also received a spreadsheet containing their final daily demand predictions, which we used to compare their accuracy to other predictive models. It is important to note that we did not receive how these final predications were calculated. These equations use the following variables to forecast the next day's water demands: (1) Previous Days Flow, (2) Weather Constant (cloud cover), (3) City Activity (normal, Home Football, Parent's Weekend, Christmas Weekend, Thanksgiving Weekend), (4) Day of Week, and (5) Rain Day Interval.

Each of these variables is included in the three simple equations that Water Services uses. The equations are broken out by day (one for Tuesdays, one for Thursdays, and one for the other days). According to City staff, Tuesday and Thursday have their own equation because they are traditionally low irrigation days.

Variable Definition. We then attempted to recreate, refine, and combine Water Service's three equations using regression modeling. Using daily pumped water and weather data from 2000 through 2015, we recreated some of the variables utilized in the daily water demand equations utilized by Water Services.

The Previous Days Flow variable was recreated using the previous day's pumped water. The Weather Constant could not easily be recreated since we could not obtain cloud cover data for any extended period of time. However, according to the International Satellite Cloud Climatology Project, the main effects that cloud cover has on climate is the temperature of the earth. Due to this, we believed it best to use a temperature variable (T_{MAX}, T_{MIN}, or T_{AVG}) in place of this variable. The T_{MAX} variable was ultimately chosen, because it had the strongest correlation with the amount of water pumped each day.

The next variable is City Activity. In Water Services' equations there are four events that they have specified as significantly changing water demand: 1) Home Football, 2) Parents Weekend, 3) Christmas Weekend, and 4) Thanksgiving Weekend. In our regression model, these variables were represented by binary indicator variables. Essentially, they act as switches. When the variable has a value of 1, it's "on" and contributes to the model, but when it has a value of 0, it's "off" and contributes nothing to the model. These variables are explained further below.

The next variable is then Day of Week. We examined the effects of this variable in two different ways: (1) as a discrete variable with values between 1 and 7 and (2) by including Day of Week Fixed Effects (FE_{DOW}) in the model. As a discrete variable, each number is equal to a different day. For example, 1 would indicate the observation occurred on a Sunday, 2 would indicate the observation occurred on a Monday and so on. On the other hand, Day of Week Fixed effects act similarly to the "Event" variables and estimates the average effects of, for example, a Monday on water demand.

When we added the discrete day of week variable (DOW), we examined the residual plot. As we can see, the plot had a slight pattern to it (see Figure C-1). We also calculated if there was a significant difference between the mean water pumped on each individual day versus the mean of all days. This analysis can be seen in Table C-1.

Based on the results that are described in Figure C-1 as well as the analysis summarized in Table C-1, were led to include the day of week fixed effects (FE_{DOW}) in our final model versus the discrete DOW variable.

Figure C-1: Day of Week (DOW) Residual Plot

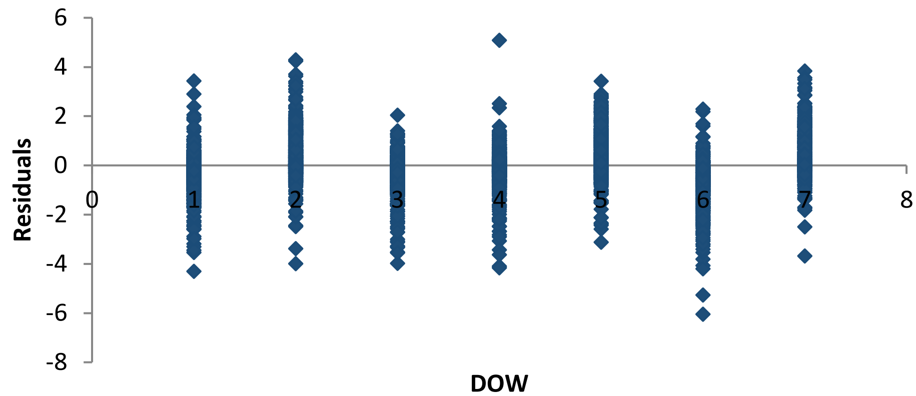


Table C-1: Day of Week Difference between Means

	SUN	MON	TUE	WED	THR	FRI	SAT	ALL
$AVERAGE_{PUMP}$	10.98	11.62	10.71	11.27	10.90	11.45	11.03	11.14
$AVERAGE_{PREV\ PUMP}$	11.03	10.98	11.62	10.71	11.27	10.91	11.44	11.14
Proportion	1.00	1.06	0.92	1.05	0.97	1.05	0.96	1.00
Stnd Dev	4.10	4.42	3.84	4.26	4.01	4.31	4.26	4.17
Obs	835	835	835	835	835	834	835	5844
t-value	1.04	-3.09	2.81	-0.84	1.56	-2.00	0.71	
Significant	**NO**	**YES**	**YES**	**NO**	**NO**	**YES**	**NO**	

This set of day-of-week (FE_{DOW}) variables allows us to account for the differences in activity level in the City as it affects water demand between, for example, a Monday and a Wednesday. This otherwise would be unaccounted for, since the previous pumped water variable (PREV PUMP) can only account for one day prior to current day use. In other words, Day of Week Fixed Effects accounts for the standard activity level of the City on any given day of the week; for example, more water is pumped on Mondays and Fridays, on average, than other days of the week (Table C-2).

Table C-2: Variable Summary

Variable	Mean	SE	Min	Max
PUMP	11.14	0.05	3.28	26.24
PREV PUMP	11.14	0.05	3.28	26.24
T_{MAX}	79.92	0.19	31.00	112.00
CHRIST	0.01	0.00	0.00	1.00
THANKS	0.01	0.00	0.00	1.00
GAMEDAY	0.02	0.00	0.00	1.00
RINGDAY	0.01	0.00	0.00	1.00
DOW	4.00	0.03	1.00	7.00
$DAYS_{SINCE}$	4.88	0.09	0.00	56.00

Finally, the last variable was Rain Day Interval. In our replica model, we included this as a $DAYS_{SINCE}$ variable. This variable is a running discrete variable that counts the number of days since the last precipitation event (greater than 0.01 inches). When developing the model, we used forward selection, meaning we added one variable at a time, checking for individual and model significance (p-values and F-statistic), as well as individual variable contribution to the model (partial R^2).

Replica Model Development. To begin the replica model development, we first added the PREV PUMP variable into our regression variable. As expected (due to high correlation between water pumped in the current day (PUMP) and water pumped in the previous day (PREV PUMP)), this simple regression model has a very high R^2 (0.93), a significant F-statistic, and a high t-statistic (both significances are above the 99% level). Also, when we plotted the residuals with the predicted values, they appeared to be randomly scattered, indicating no need to transform the response variable (PUMP) in this model.

We then added several of the Day of Week Fixed Effects variables (MON, WED, and FRI). Together, these three variables explain about 1.41% of variation in pumped water. Next, the T_{MAX} (maximum temperature) variable was added due to a lack of cloud cover data. The partial R^2 of this variable was 0.0030, indicating that only 0.3% of the variation can be explained by changes in the maximum daily temperature.

We then added another Day of Week Fixed Effect variable (SUN). This variable explained only about 0.14% of variation in pumped water. Next, the $DAYS_{SINCE}$ variable was added. When this variable was added, the R^2 value increased slightly. Also, the partial R^2 value of this variable was 0.0013. This indicates that the $DAYS_{SINCE}$ variable explains 0.13% of the variation in water demand (PUMP variable) when the PREV PUMP variable was already included. When viewing the residual plot of this model, there was no indication of the need to transform the response variable. Also, all t-statistics and the F-statistic are significant above the 99% level. We continued to add the other variables in sequence including only those that were significant and followed our criteria for model inclusion (Table A-3). When this was done, we found that all variables were significant except the variables indicating Parents Weekend, Christmas, and Thanksgiving. This can be seen in Table C-3 below:

Table C-3: Daily Water Demand Model – Replica Output Summary

Variable	Coefficients	Partial R^2	P-value	Variation Inflation Factor	Model R^2	F-value
Intercept	-1.52432				0.9468	
$PUMP_{PREV}$	0.89974	0.9256	<0.0001	2.28		72649.9
MON	1.47476	0.0038	<0.0001	1.72		316.1
WED	1.37265	0.0042	<0.0001	1.73		370.6
FRI	1.34778	0.0061	<0.0001	1.77		586.0
T_{MAX}	0.02086	0.0030	<0.0001	2.03		300.6
SUN	0.77143	0.0014	<0.0001	1.76		148.9
$DAYS_{SINCE}$	0.02446	0.0013	<0.0001	1.22		143.5
THR	0.49706	0.0004	<0.0001	1.72		40.6
SAT	0.45131	0.0009	<0.0001	1.76		100.9
GAMEDAY	0.17770	0.0001	0.0002	1.08		9.6

Model R^2 = 0.9468; Mean of Squared Error = 0.9256; 5844 Observations (days/month x 12 months x 16 years)

As we can see from Table C-3 on the last page, all of these variables are significant above the 99% level (except GAMEDAY). However, many of the variables explain very little of the variation in pumped water. This is mostly because they are only intended to signify certain events (such as home football games) that over the course of a year are not very impactful.

It is good that Water Services has recognized that on these days, water consumption is significantly different than on regular days. However, it may not be necessary to include these less explanatory variables in a predictive model. Due to this, we have excluded any variables that explain less than 0.1% of variability in pumped water. Table C-4 below is the final predictive model used to evaluate accuracy in Table 7 of the report:

Table C-4: Daily Water Demand Model – Replica Output Summary

Variable	Coefficients	Partial R^2	P-value	Variation Inflation Factor	Model R^2	F-value
Intercept	-1.21229				0.9454	
$PUMP_{PREV}$	0.89909	0.9256	<0.0001	2.25		72649.9
MON	1.15107	0.0038	<0.0001	1.15		316.1
WED	1.04898	0.0042	<0.0001	1.15		370.6
FRI	1.04554	0.0061	<0.0001	1.15		486.0
T_{MAX}	0.02108	0.0030	<0.0001	2.03		300.6
SUN	0.46811	0.0014	<0.0001	1.15		148.9
$DAYS_{SINCE}$	0.02486	0.0013	<0.0001	1.22		143.5

Model R^2 = 0.9454; Mean of Squared Error = 0.9492; 5844 Observations (days/month x 12 months x 16 years)

Appendix D: Developed Forecasting Methodologies – Per Capita

This forecasting method is based solely on historic pumped water data and population projections. It is the easiest method to use and requires the least amount of time and effort. Like the other methods that will be discussed here, it forecasts daily average water demand. Below is an example simple forecast:

Table D-1: Simple Water Demand Forecast*

2015 population	106,465
2015 average water production, mgd	12.2504
2015 per capita water use, gpcd	115.0651
2015 year peak-to-average day ratio	2.1088

Year	Population Forecast	Water Demand Forecasts	
		Average Day (mgd)	Peak Day (mgd)
2020	113,665	13.08	27.58
2025	124,219	14.29	30.14
2030	134,772	15.51	32.70

*Where mgd stands for millions of gallons per day and gpcd stands for gallons per capita per day

In the above example, 2015 daily pumped water data (mgd, millions of gallons per day) was averaged to produce the "2015 average water production, mgd." The December 2015 population estimate purported by the City of College Station's Planning and Development Services Department was used as the "2015 population." The average water production was then divided by the 2015 population to obtain the "2015 per capita water use, gpcd." The "2015 year peak-to-average day ratio" was calculated by dividing the highest water production day in 2015 (25.8336 mgd) by the average daily water production in 2015.

The Population Forecast numbers were taken from the City of College Station Comprehensive Plan as adopted in May 2009 (page 1-13). These population forecasts were then multiplied by the calculated per capita water use to obtain the Average Day (mgd) Water Demand Forecast. This was then multiplied by the year peak-to-average day ratio to obtain the Peak Day Water Demand Forecast.

General Conclusions:

The simple forecasting method is the easiest to use. This forecasting method requires only two sources of data, which are readily available to Water Services. This method also takes into account daily water loss and forecasts peak day demands, which is important when planning for infrastructure improvements.

Simple forecasts have limited uses. Though the simple forecast is useful for long term infrastructure planning, it is not useful as a revenue projection tool. The simple forecast only provides a general daily usage for the entire customer population, which cannot be easily translated into a revenue stream. Also, it relies on population projections, which are prone to inaccuracy. Due to the presence of Texas A&M University, College Station's population is more transient and variable than many other cities. This variability may require a more detailed approach to forecasting.

Appendix E: Developed Forecasting Methodologies - Sectoral

This forecasting methodology is a disaggregated, variant unit-use water demand forecast. We had complete billed water data from 2008 through 2015 for individual water location IDs. We then connected each location ID to its 2016 property type, assigned using the Brazos Central Appraisal Districts database. A full list of property type codes can be seen in Table E-4 at the end of this appendix.

Once we had identified the property type for each metered location, we created an average consumption pattern and an aggregate consumption pattern for each property type group. We then correlated these patterns together to generate an average correlation for each pair of property types.

For example, the Commercial (F1) and Industrial (F2) properties' consumption patterns for one year and resulting correlations are shown in Table E-1. Only one year of each consumption pattern is shown in this table, however, each pattern was generated for the period between 01/2008 and 11/2015.

Table E-1: Example Correlation Calculations

	Commercial (F1)		Industrial (F2)	
Month	Average Pattern	Aggregate Pattern	Average Pattern	Aggregate Pattern
01/2008	28.45	42607.65	37.36	410.96
02/2008	27.56	41410.69	35.95	415.66
03/2008	29.35	45308.90	47.96	541.46
04/2008	36.83	58574.00	80.43	884.77
05/2008	46.09	72599.11	107.46	1182.05
06/2008	61.57	94987.35	98.15	1079.68
07/2008	64.33	100143.32	138.20	1580.72
08/2008	52.27	83620.41	93.24	1118.92
09/2008	45.34	73807.43	126.80	1462.39
10/2008	37.25	59938.63	109.68	1156.53
11/2008	36.86	54435.22	90.67	885.02
12/2008	34.02	53605.65	99.27	893.47
Average Consumption Pattern Correlation				0.88
Aggregate Consumption Pattern Correlation				0.90
Average Correlation				0.89

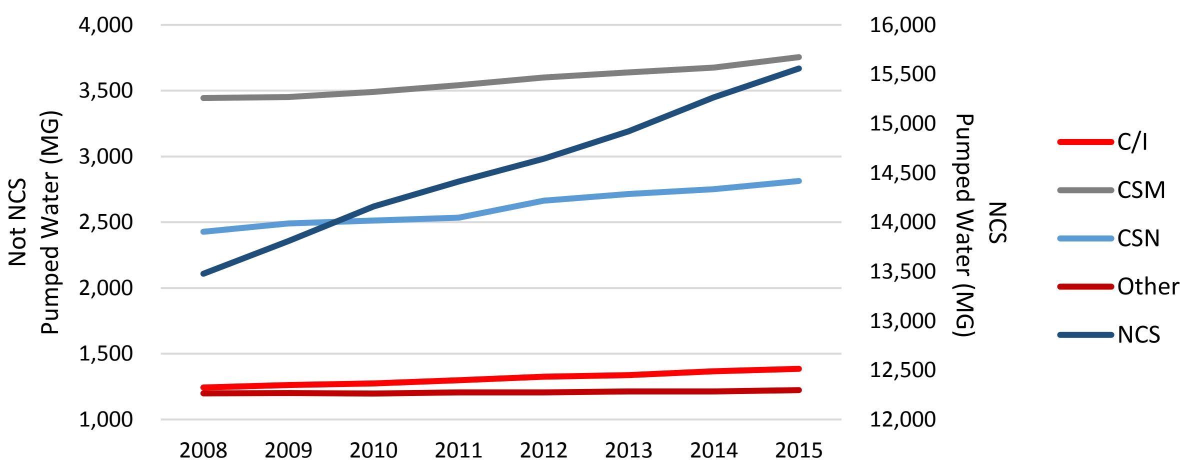
Based on these correlations, the water use categories were identified as shown in Table E-2. A brief description of our recommended sectoral forecasting methodology follows; the full calculations can be seen at the end of this appendix.

Table E-2: Water Use Categories

Property Type Category	2015 Metered Location	2013	2014	2015	Avg.
No Common Space (NCS)					
Single Family	15,099	60.91%	61.12%	61.05%	61.03%
Patio Home	449	1.68%	1.73%	1.82%	1.74%
Manufactured Home	10	0.04%	0.04%	0.44%	0.04%
NCS Totals:	**15,558**	**62.63**	**62.89%**	**62.91%**	**62.81%**
Commercial & Industrial (C/I)					
Commercial	1,376	5.57%	5.59%	5.56%	5.57%
Industrial	9	0.04%	0.04%	0.04%	0.04%
C/I Totals:	**1,385**	**5.61%**	**5.63%**	**5.60%**	**5.61%**
Shared Common Space – Not Resident Maintained (CSN)					
Apartments	2,379	9.73%	9.69%	9.63%	9.68%
Condominium	434	1.68%	1.67%	1.75%	1.70%
CSN Totals:	**2,813**	**11.41%**	**11.36%**	**11.38%**	**11.38%**
Shared Common Space – Resident Maintained (CSM)					
Duplex	2,496	10.44%	10.25%	10.09%	10.26%
Townhome	787	2.92%	2.98%	3.18%	3.03%
Homeplex	426	1.73%	1.72%	1.72%	1.72%
Triplex	42	0.18%	0.17%	0.17%	0.17%
Fraternity/Sorority	4	0.02%	0.02%	0.02%	0.02%
CSM Totals:	**3,755**	**15.29%**	**15.14%**	**15.18%**	**15.2%**
All Other Categories	**1,223**	**5.09%**	**5.00%**	**4.94%**	**5.01%**

To begin water usage rates were calculated for each water use category by dividing the annual water consumption by the number of locations for each category. These were then averaged over an eight year period and forecasted forward. Below is a graph of categorical location growth from 2008 through 2015.

Figure E-1: Categorical Location Growth



The metered location growth of the NCS, C/I, CSN, and CSM categories are all fairly linear, so linear trend lines were generated to forecast this growth (Figure E-1). A second degree polynomial trend line was used for the "Other" category, since it provided the best fit to the data, but did not predict unreasonable growth.

The forecasted metered location amounts were then multiplied by the calculated average water usage rates to obtain the average daily water demand for each category. These are then added together to equal the “metered usage.”

Water loss was forecasted using Water Services’ 2014 Water Conservation Plan. In this plan, Water Services states their water loss goal to be 8 gallons per capita per day. Noting this, the water loss is then calculated by multiplying the forecasted population (taken from the City’s 2009 Comprehensive Plan) amount by 8 and dividing by one million. The average daily metered usage plus the average daily water loss then equals the total system demand.

Table E-3 Sectoral Water Demand Forecast

Year	Population Forecast	Water Demand Forecasts (mgd)							
		NCS	C/I	CSN	CSM	Other	Metered Usage	Water Loss	Total System Demand
2020	113,665	6.05	2.78	0.70	2.82	0.83	13.17	0.91	14.08
2025	124,219	6.57	2.97	0.74	3.07	0.87	14.23	0.99	15.22
2030	134,772	7.08	3.16	0.78	3.33	0.94	15.29	1.08	16.37

Table E-4: State Property Types – Full

State Code (BCAD)	Description
A1	Residential-Single Family
A2	Residential-Manufactured Home
A3	Residential-Duplex
A4	Residential-Triplex
A5	Residential-Fourplex
A6	Residential-Condominium
A7	Residential-Townhome
A8	Residential-Patio Home
A9	Residential-Homeplex
B1	Rental-Apartments
B2	Rental-Duplex
B3	Rental-Triplex
B4	Rental-Fourplex
B10	Rental-Fraternity/Sorority House
C1	Vacant-Residential Lot
C2	Vacant-Commercial Lot
C3	Vacant-Rural or Recreational Lot

State Code (BCAD)	Description
D1	OS-Land Qualified For Open Space
D2	OS-Farm & Ranch Improvements
E1	Rural-Single Family
E4	Rural-Land
EA2	Rural-Manufactured Home
EB1	Rural-Apartments (5+)
EB2	Rural-Duplex
EB3	Rural-Triplex
EB4	Rural-Fourplex
F1	Commercial
F2	Industrial
F3	Commercial-Improvement Only
F4	Industrial-Improvement Only
M1	Personal Property Manufactured Home

Table E-5: Property Type Correlations – Expanded

	A1	A2	A3	A5	A6	A7	A8	A9	B1	B2	B3	B4	B10	F1	F2
A1	1.00														
A2	0.94	1.00													
A3	0.58	0.53	1.00												
A5	0.08	0.07	0.12	1.00											
A6	0.76	0.69	0.53	0.11	1.00										
A7	0.86	0.82	0.52	0.06	0.75	1.00									
A8	0.95	0.92	0.53	0.07	0.78	0.84	1.00								
A9	0.85	0.77	0.52	0.15	0.76	0.77	0.86	1.00							
B1	0.53	0.47	0.23	0.04	0.78	0.54	0.57	0.57	1.00						
B2	0.75	0.68	0.54	0.14	0.67	0.66	0.73	0.80	0.49	1.00					
B3	0.77	0.74	0.52	0.06	0.68	0.69	0.81	0.70	0.47	0.77	1.00				
B4	0.33	0.29	0.39	0.01	0.43	0.40	0.32	0.35	0.26	0.48	0.49	1.00			
B10	0.77	0.71	0.54	0.10	0.76	0.66	0.82	0.77	0.55	0.80	0.74	0.41	1.00		
F1	0.93	0.91	0.57	0.07	0.81	0.87	0.96	0.82	0.59	0.68	0.75	0.31	0.78	1.00	
F2	0.89	0.86	0.52	0.07	0.74	0.80	0.87	0.76	0.56	0.58	0.68	0.26	0.69	0.89	1.00

Color Key

Color	Correlation
Very Strong	0.90-0.99
Strong	0.80-0.89
Semi Moderate	0.70-0.79
Moderate	0.60-0.69

Table E-6: Historic Annual Water Usage by Category

DATE	**ANNUAL USAGE (millions of gallons)**				
Year	NCS	C/I	CSM	CSN	Other
2008	1,828.4	797.2	237.2	786.8	263.9
2009	1,799.5	802.3	236.5	827.3	283.5
2010	1,799.8	816.3	225.0	844.3	277.1
2011	2,337.2	1,044.8	255.5	939.6	386.8
2012	1,894.6	893.8	227.6	875.3	220.7
2013	1,900.8	913.6	227.7	884.0	273.5
2014	1,730.4	899.0	218.0	893.3	286.0
2015	1,783.3	998.0	220.9	898.3	308.1

Table E-7: Sectoral Forecast Calculations – Expanded (millions of gallons)

Year	NCS	C/I	CSM	CSN	Other	NCS	C/I	CSM	CSN	Other	NCS	C/I	CSM	CSN	Other	Metered Usage	Water Loss	Total System Demand
2008	370.7	1,752.3	188.2	885.8	601.3	13,477	1,243	3,444	2,427	1,199	5.0	2.2	0.6	2.1	0.7	10.69	1.00	11.69
2009	357.0	1,743.1	187.7	910.3	647.2	13,808	1,261	3,451	2,490	1,200	4.9	2.2	0.6	2.3	0.8	10.82	0.97	11.79
2010	348.3	1,754.1	176.6	920.5	634.3	14,159	1,275	3,491	2,513	1,197	4.9	2.2	0.6	2.3	0.8	10.86	0.89	11.75
2011	444.3	2,203.5	197.7	1,015.9	879.4	14,411	1,299	3,541	2,534	1,205	6.4	2.9	0.7	2.6	1.1	13.60	0.98	14.58
2012	353.5	1,843.0	172.7	898.0	500.0	14,642	1,325	3,601	2,663	1,206	5.2	2.4	0.6	2.4	0.6	11.23	0.92	12.15
2013	349.0	1,873.6	171.4	892.4	617.7	14,922	1,336	3,640	2,714	1,213	5.2	2.5	0.6	2.4	0.7	11.51	1.23	12.73
2014	310.6	1,803.0	162.5	889.4	645.4	15,266	1,366	3,676	2,752	1,214	4.7	2.5	0.6	2.4	0.8	11.03	0.53	11.56
2015	314.0	1,974.1	161.1	874.9	690.2	15,558	1,385	3,755	2,813	1,223	4.9	2.7	0.6	2.5	0.8	11.53	0.76	12.29
2016	355.9	1,868.4	177.2	910.9	651.9	15,836	1,404	3,779	2,867	1,230	5.6	2.6	0.7	2.6	0.8	12.34	0.84	13.18
2017	355.9	1,868.4	177.2	910.9	651.9	16,126	1,424	3,824	2,924	1,238	5.7	2.7	0.7	2.7	0.8	12.55	0.86	13.41
2018	355.9	1,868.4	177.2	910.9	651.9	16,416	1,445	3,870	2,980	1,248	5.8	2.7	0.7	2.7	0.8	12.76	0.88	13.63
2019	355.9	1,868.4	177.2	910.9	651.9	16,706	1,466	3,915	3,037	1,258	5.9	2.7	0.7	2.8	0.8	12.96	0.89	13.86
2020	355.9	1,868.4	177.2	910.9	651.9	16,997	1,486	3,960	3,093	1,269	6.1	2.8	0.7	2.9	0.8	13.17	0.91	14.08
2021	355.9	1,868.4	177.2	910.9	651.9	17,287	1,507	4,006	3,150	1,282	6.2	2.8	0.7	2.9	0.8	13.38	0.93	14.31
2022	355.9	1,868.4	177.2	910.9	651.9	17,577	1,527	4,051	3,206	1,295	6.3	2.9	0.7	2.9	0.8	13.59	0.94	14.54
2023	355.9	1,868.4	177.2	910.9	651.9	17,867	1,548	4,096	3,263	1,309	6.4	2.9	0.7	3.0	0.9	13.80	0.96	14.76
2024	355.9	1,868.4	177.2	910.9	651.9	18,157	1,568	4,142	3,319	1,324	6.5	2.9	0.7	3.0	0.9	14.01	0.98	14.99
2025	355.9	1,868.4	177.2	910.9	651.9	18,448	1,589	4,187	3,376	1,341	6.6	3.0	0.7	3.1	0.9	14.23	0.99	15.22
2026	355.9	1,868.4	177.2	910.9	651.9	18,738	1,610	4,232	3,432	1,358	6.7	3.0	0.8	3.1	0.9	14.44	1.01	15.45
2027	355.9	1,868.4	177.2	910.9	651.9	19,028	1,630	4,278	3,489	1,376	6.8	3.0	0.8	3.2	0.9	14.65	1.03	15.68
2028	355.9	1,868.4	177.2	910.9	651.9	19,318	1,651	4,323	3,545	1,395	6.9	3.1	0.8	3.2	0.9	14.87	1.04	15.91
2029	355.9	1,868.4	177.2	910.9	651.9	19,608	1,671	4,368	3,602	1,416	7.0	3.1	0.8	3.3	0.9	15.08	1.06	16.14
2030	355.9	1,868.4	177.2	910.9	651.9	19,899	1,692	4,414	3,658	1,437	7.1	3.2	0.8	3.3	0.9	15.29	1.08	16.37

Appendix F: Management Response

The following is the Water Services Department's response to the recommendations made in the City Internal Auditor's Office Water Demand Forecasting Audit. The audit recommendation is followed by a response describing how the recommendation will be addressed by the Water Services Department.

1. Audit Recommendation: More complex methods should be investigated in the future as the City grows and diversifies. In the past, the forecasting methods utilized by the City have been sufficient. Though each forecast has associated risks, these have not had significant impact on Water Services operations in the past. However, as the City grows and diversifies these risks may become more apparent. As this occurs, the City could benefit from more complex in-house water demand forecasting approaches, as it allows for more thorough analysis and increases institutional knowledge.

Management Response

Management concurs with this recommendation and will begin the transition to in-house water demand forecasting. We are in the process of hiring a new position, an Engineer-in-Training to work under our Utility Engineer. With this added manpower, we can begin to dedicate the time to train and implement more sophisticated and accurate water demand forecasting methods.