Spring 1-1-2012

Statistical Methods for Water Demand Modeling to Help Mitigating Climate Induced Risk

Erik Clifford Haagenson
University of Colorado at Boulder, ehaagenson@gmail.com

Statistical Methods for Water Demand Modeling to Help Mitigating Climate Induced Risk

by

Erik Clifford Haagenson

B.S., Montana State University, 2008

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Masters of Science

Department Civil, Architectural and Environmental Engineering

2012

This thesis entitled:
Statistical Methods for Water Demand Modeling to Help Mitigate Climate Induced Risk
written by Erik Haagenson
has been approved for the Department of Civil, Environmental and Architectural Engineering

____
Balaji Rajagopalan

____
R. Scott Summers

____
Alan Roberson

Date_______________

The final copy of this thesis has been examined by the signatories, and we
Find that both the content and the form meet acceptable presentation standards
Of scholarly work in the above mentioned discipline.

Haagenson, Erik Clifford (MS, Civil Engineering)

Statistical Methods for Water Demand Modeling to Help Mitigating Climate Induced Risk

Thesis directed by Prof. Balaji Rajagopalan

Abstract. Water demand projections are crucial for efficient management of water utilities, especially given the stress from socio-economic growth and supply reductions and variations under climate variability. Currently, many water utilities rely largely on statistical models for average demand incorporating socio-economic and limited climate information. Water managers require projections of different demand attributes such as peak demand, demand exceeding desired thresholds, sustained period of exceedances, etc. There is increasing recognition that climate plays a substantial role in demand modulation and that it needs to be accounted better in water demand models. Two key gaps are identified with the state of knowledge that motivated this research – (i) the need to incorporate all attributes of weather and climate and (ii) the need to provide tools to model a variety of demand attributes that are simple and robust and can incorporate climate and socio-economic factors. This research makes the following contributions. (1) A suite of weather attributes are identified that are strongly related to attributes of water demand, these include hot/dry and wet/cold spells and average precipitation and temperature; (2) A Generalized Linear Modeling (GLM) framework to model monthly average, monthly peak and number days of demand exceeding a threshold, incorporating the weather attributes; (3) Introduction of Extreme Value Analysis (EVA) to model monthly peak demand and demand exceeding thresholds, which also uses climate attributes to model nonstationarity of the water demand variability (4) demonstration of the models from

EVA to make projections of water demand extremes under climate change. All of these are developed and demonstrated for water demand data from Aurora Water, the water utility of Aurora, CO. These methods are simple and robust and can easily be modified to incorporate other covariates such as social and economic. Furthermore, they can be applied with limited resources and thus, provides an effective set of tools to most water utilities.

Dedication

To my wife Brianne: I owe a great debt to her loving encouragement and steadfast patience.

Acknowledgements

I would like to firstly acknowledge Professor Balaji Rajagopalan who took an inexperienced graduate student under his wing and showed an amazing amount of patience and persistence in teaching me a great deal of skills, both related to hydrology and statistics and to life in general. Without your positive attitude, incredibly deep insights and constant encouragement this work would not have been possible. Thanks also to Professor Scott Summers who was there for me from beginning to end, interjecting real ideas and direction throughout my tenure at CU. And thanks to Alan Roberson, who kept the project on track and worked to make my research relevant to a broader audience.

I would also like to thank my friends and family who serve as my reality checkers and grand distracters from the long hours of graduate school and other obligations. Thank you Brianne, Ryan, Chris, Courtney, Kyle, Abby, Wes, Marissa, Bryce, Kate, Aaron, Matt, Steph, Paul, Rachel and Travis.

I would also like to thank my funding agencies including the NOAA Sector Applications Research Program (SARP), the American Water Works Association (AWWA) and the Bureau of Reclamation. I thank the National Center for Atmospheric Research and their Advanced Study Program 2011 and Eric Gilleland and Rick Katz for a solid Extreme Value Analysis foundation to build upon.

Contents

Tables

Table

Figures

Figure

1 Introduction

Successfully managing a water utility requires a careful balance of available water supply to meet water demand. Maintaining these skillful water demand forecasts is imperative. Water demand forecasts help water utility managers make decisions about additional supply acquisition, treatment infrastructure, water use restrictions, water conservation impacts, price elasticity and demand management (*Billing et al.* 2008). A large number of factors are known to impact water demand, including population, employment, technology, weather, climate, price, infrastructure efficiency, conservation programs, socioeconomics and water awareness. However, many of these variables are difficult to quantify and have high degree of uncertainty (*Billings et al.* 2008). It has been suggested that the threat of increasing water demand greatly outweighs other consequences of climate variability including variations in the hydrologic cycle in defining the status of global water systems (*Vorosmarty et al.,* 2000). Thusly, increasing the skill and usability of current water demand modeling efforts is imperative.

1.1 Water Demand Modeling

Water demand has been extensively studied since the 1960's with a primary focus on aggregate municipal demand or residential demand (e.g., *Gottlieb et al.* 1963; *Howe et al.* 1967; *Baumann et al.* 1997). Most research on water demand uses an economic modeling approach that attempts to utilize water price and socio-economic variables as covariates in statistical models for water demand (*Arbués et al.* 2003). Many studies have emphasized the use of price elasticities and income for water demand forecasting (*Katzman,* 1977). Such econometric models have been used to forecast water demand and have been expanded to incorporate price and incomes using traditional and generalized least squares technique (e.g., *Katzman,* 1977; *Babel et al.* 2007; *Malla et al., 1997*). Multiple regression models to analyze sustainable water consumption models (*Colyer et al.* 2000) have been developed and recently socio-economic indicators have been incorporated into these models (*Babel et al.* 2007). Discrete/continuous choice models have been created and found increased price elasticity when used (*Hewitt et al.* 1995). The aim of the work presented here is to expand upon the basic approaches to water demand modeling and incorporate modeling techniques not yet applied to water demand to create a nuanced picture of water demand risk due to climate change.

Seasonal average demand has been largely the focus of modeling efforts using linear statistical or econometric models (*McKee et al.* 1995). However, there are many important demand attributes in addition to average demand that the utility managers desire for efficient management such as; seasonal, monthly peaks, number of exceedances above a certain threshold etc. Seasonal maximum water demand models are used for planning purposes related to necessary maximum capacity or storage needs for decision making (*Arbués et al.* 2003). Generalized Linear Modeling of seasonal average, seasonal maximum and seasonal water demand exceedances is motivated in Chapter 2. This approach stands as a marked improvement over traditional, regression based analysis and can be applied by utilities of all sizes and sophistication.

1.2 Modeling Water Demand Extremes

A key attribute of water demand for planning and management is the extreme water demand events. Specifically, peak demand (either peak day and/or peak hour) is used for treatment plant capacity evaluation and for network design, which drives the raising of funds from internal capital and/or financial markets for capacity expansion, etc. (*Arbués et al.* 2003, *Billings et al.,* 2008); also of interest is the interval of time with elevated demand or demand above a critical threshold. Weather attributes also modulate water demand extremes. Current models are limited in their ability to model extremes and this is a major gap in demand management, which motivates the work in Chapter 3. The emerging field of Extreme Value Analysis (EVA) offers an attractive, simple and robust approach to modeling demand extremes that can be adapted by a wide swath of utilities regardless of economic resources. The desire to better model hydrologic risks was the early motivation for EVA (*Gumbel,* 1941). It has since matured and is now widely used in modeling hydrologic, meteorologic and sea level extremes (*Fiorentino*, 1984; *McNeil*, 1997; *Swift*, 1989; *Beguería*, 2006; *Coles*, 2003; *Katz*, 1995; *Eaton*, 1994; *Coles*, 2001; *Cooley*, 2009; *Smith*, 1989) and financial risk estimation (e.g., *Embrechts*, 1999; and for hydrological losses, *McNeil*, 1997).

At the core of the EVA is the recognition that extreme events by definition occur in the tail of the probability density function (PDF) and that the traditional PDFs have a thin tail, thereby underestimating the probability of extreme events. Therefore, heavy tailed PDFs are necessary, which are developed from theory (*Katz and Naveau*, 2002; *Coles*, 2001). Extreme events also vary with time and consequently so to do the probability of their occurrences – to capture this variability (i.e., nonstationarity) the PDFs should also vary. To this end, EVA proposes modeling of the parameters of the heavy tailed PDFs as a function of covariates. In this research, the covariates are the weather attributes identified in Chapter 2.

In Chapter 3, EVA for modeling water demand is introduced and the hope that this initiation to the field will lead to further application and expansion of this approach. EVA also supplies a nuanced method for understanding the risk associated with extreme water demand events and can also be used for making water demand projections under climate change (*Chapter 4*).

1.3 Projecting Water Demand Extremes Under Climate Change

Understanding variability in water demand and being able to make skillful projections of its future variability is crucial for efficient water supply management and for water demand forecasting. Water demand forecasts enable water utilities to plan for capacity expansion, pricing, conservation efforts etc. (*Billings et al.*, 2008). Water demand projections help water utility manages make decision about short-term operational decisions such as treatment plant production, balancing supply and demand, and potential water use restrictions in the event of a drought. Every year, water utility managers and financial staff make a prediction of how much water there are going to sell (water sales) and, subsequently, predict the resultant revenues as part of their annual budgeting process. Additionally, these demand projections help managers make long-term decisions about additional supply acquisition, additional treatment capacity, transmission mains, storage tanks, water conservation impacts, price elasticity, and demand management (*Billings et al.,* 2008). Longer-term financial sustainability for the water system is anchored by accurate long-term demand forecasts. Accurate water demand forecasts are needed for both short-term (operational and financial) and long-term (planning and financial). The costs of being wrong can be significant. Inaccuracies in longer-term forecasts can result in large costs to utilities and their customers, for example, in the form of stranded capital assets, insufficient supply reliability, or a reduced level of service due to treatment plant capacity limitations.

All utilities use water demand projections, however, many of them are from simplified models and fail to incorporate nonstationarity in variability from climate and other uncertainty. A large number of factors are known to impact water demand, including population, employment, technology, weather, climate, price, infrastructure efficiency, conservation programs, socioeconomics and water awareness (*Billings et al.,* 2008). Models have focused mainly on the social and economic variables—which are often difficult to project—resulting in complex models. Many of these models are outside the reach of utilities which are small in size with limited resources.

Utilities are specifically interested in the high impact, low probability water demand events which drive infrastructure planning decisions and the need for capital to fund such improvements. Accurate predictions of peak hour demand and peak day demand are necessary for planning capital improvements such as an alternate supply of source water, treatment plant capacity, transmission mains, storage tanks, and booster pumping stations. The EVA models used for projections are developed in Chapter 3.

1.4 Climate Change and Water Demand

Climate plays a fairly significant role in modulating water demand – for example, increased temperature and reduced precipitation leads to increased water demand and vice-versa. Recognizing this, some of the early research used precipitation during growing season to model agriculture water demand (e.g., *Foster et al.* 1979; *Foster et al.* 1981). Others used the difference between rainfall and evapotranspiration as a potential predictor variable for demand forecasting (e.g., *Billings,* 1982; *Agthe et al.* 1986; *Nieswiadomy et al.* 1988). Other amalgamations of climatic variables have also been used, such as a variable function of temperature, minutes of sunshine, and wind speed (*Al-Quanibet et al.* 1985). Micro-time series data—which aims to predict water demand on hourly or smaller time scales—has been used in econometric models incorporating covariates (*Danielson,* 1979). Typically, climatic variables are represented as temporal (i.e., seasonal, monthly etc.) averages but this approach may be limiting in that, the effect of precipitation may be more pronounced at first, and then decrease with time (*Maidment et al.* 1986). To address this issue, and others, seasonal spell statistics such as rainy days have been used in water demand modeling, (*Martínez-Espiñeira,* 2002). This idea is expanded upon and a full suite of weather attributes is created that are then used as covariates in water demand modeling.

Extreme water demand events exhibit significant variability over time driven by climate fluctuations – for example, an increasing trend in the frequency and intensity of precipitation events in the US has been known for more than 15 years (*Trenberth,* 1998, 1999). The IPCC (2010) projects the same for much of the world in the coming decades due to climate change. Researchers in water resources and hydrology have been working on translating the climate change projections from the global climate models to impacts on hydrology and water resources. This requires downscaling climate information which is often at coarse spatial and temporal scales to point scale of interest and, driving them through hydrologic models to generate ensembles of streamflows and consequently, impacts to water resources – a body of literature exists on understanding water resources variability under climate change in the Western US (e.g. *Christensen*, 2004; *Rajagopalan et al.*, 2009). Recently, joint efforts from water utilities have been developing tools for water supply and demand with a focus on planning decisions by water utilities (*JFRCCVS*, 2012, *Ray et al*., 2010). The traditional approach of downscaling climate information to process models can be computationally intensive and deter a majority of water utilities from being able to access them due to limited resources. Simple, effective and flexible tools are needed that can translate coarse climate information to specific attributes of decision variables without having to go through computationally intensive process models. *Towler et al.* (2010) developed an extreme value analysis models to translate monthly precipitation and temperature projections under climate change to projections of streamflow extremes and consequently water quality extremes in the Northwest United States and also for Aurora CO (*Towler et al.*, 2011). Such models are effective in providing projections of extreme events that are essential to water utility managers for planning purposes.

Water demand is somewhat different than water supply and water resources due to the potential human response factoring into demand. But the same weather factors such as temperature and precipitation play a significant role in predicting water demand. This research has expanded this EVA approach to apply to projecting water demand extremes, and the potential impacts from climate change to those extremes.

By introducing two novel modeling approaches—GLM modeling with weather attributes as covariates and Extreme Value Modeling—and then using the EVA models to project water demand extremes the goal is to expand general water demand modeling knowledge and advance the available tool set for water utilities for handling projections of water demand under climate change. Education is a critical component to climate change awareness and readiness for water utilities. The circumstances will be unique for each utility and a comprehensive analysis of the associated risks and potential mitigation strategies will need to be performed by each individual utility. The work presented here represents an expansion of the approaches available and serves as an important component to successful utility management.

2 Statistical Modeling of Seasonal Water Demand Attributes Using Climate Variables

2.1 Abstract

Water demand projections are crucial for efficient management of water utilities, especially given the stress from socio-economic growth and supply reductions and variations under climate variability. Traditional methods have relied on statistical models for average demand incorporating economic information and modest amount of climate features. Furthermore, they have varied significantly among utilities. We see two major gaps – (i) need for an approach to model a variety of demand attributes that are needed by utility managers such as – average demand, monthly peak and number of demand exceedances above a threshold and (ii) comprehensive use of climate information. We first identify a suite of climate variables that include high/low temperatures and wet/dry spells and the more traditional average temperature and precipitation. We find that these variables capture nuanced connections between climate and water demand, especially the lagged relationships. Second we propose a Generalized Linear Modeling (GLM) framework to model attributes of water demand by incorporating the climate variables. We demonstrate this approach to demand modeling for water utility in Aurora, CO. Demand attributes modeled include monthly average demand, monthly peak demand and monthly number of demand exceedances above a threshold.

2.2 Introduction

Successfully managing a water utility requires a careful balance of available water supply to meet water demand. Maintaining these skillful water demand forecasts is imperative. Water demand forecasts help water utility managers make decisions about additional supply acquisition, treatment infrastructure, water use restrictions, water conservation impacts, price elasticity and demand management (*Billing et al.* 2008). A large number of factors are known to impact water demand, including population, employment, technology, weather, climate, price, infrastructure efficiency, conservation programs, socioeconomics and water awareness. However, many of these variables are difficult to quantify and have high degree of uncertainty (*Billings et al.* 2008). It has been suggested that the threat of increasing water demand greatly outweighs other consequences of climate variability including variations in the hydrologic cycle in defining the status of global water systems (*Vorosmarty et al.,* 2000).

Water demand has been extensively studied since the 1960's with a primary focus on aggregate municipal demand or residential demand (e.g., *Gottlieb et al.* 1963; *Howe et al.* 1967; *Baumann et al.* 1997). Most research on water demand uses an economic modeling approach that attempts to utilize water price and socio-economic variables as covariates in statistical models for water demand (*Arbués et al.* 2003). Many studies have emphasized the use of price elasticities and income for water demand forecasting (*Katzman,* 1977). Such econometric models have been used to forecast water demand and have been expanded to incorporate price and incomes using traditional and generalized least squares technique (e.g., *Katzman,* 1977; *Babel et al.* 2007; *Malla et al., 1997*). Multiple regression models to analyze sustainable water consumption models (*Colyer et al.* 2000) have been developed and recently socio-economic indicators have been incorporated into these models (*Babel et al.* 2007). Discrete/continuous choice models have been created and found increased price elasticity when used (*Hewitt et al.* 1995).

Climate plays a fairly significant role in modulating water demand – for example, increased temperature and reduced precipitation leads to increased water demand and vice-versa. Recognizing this, some of the early research used precipitation during growing season to model agriculture water demand (e.g., *Foster et al.* 1979; *Foster et al.* 1981). Others used the difference between rainfall and evapotranspiration as a potential predictor variable for demand forecasting (e.g., *Billings,* 1982; *Agthe et al.* 1986; *Nieswiadomy et al.* 1988). Other amalgamations of climatic variables have also been used, such as a variable function of temperature, minutes of sunshine, and wind speed (*Al-Quanibet et al.* 1985). Micro-time series data—which aims to predict water demand on hourly or smaller time scales—has been used in econometric models incorporating covariates (*Danielson,* 1979). Typically, climatic variables are represented as temporal (i.e., seasonal, monthly etc.) averages but this approach may be limiting in that, the effect of precipitation may be more pronounced at first, and then decrease with time (*Maidment et al.* 1986). To address this issue, and others, seasonal spell statistics such as rainy days have been used in water demand modeling, (*Martínez-Espiñeira,* 2002).

Seasonal average demand has been largely the focus of modeling efforts using linear statistical or econometric models (*McKee et al.* 1995). However, there are many important demand attributes in addition to average demand that the utility managers desire for efficient management such as; seasonal, monthly peaks, number of exceedances above a certain threshold etc. - seasonal maximum water demand models are used for planning purposes related to necessary maximum capacity or storage needs for decision making (*Arbués et al.* 2003).

To summarize, water demand forecasting methods are largely simple linear models using socio-economic and modest climate variables – with some utilities, those with advanced capabilities having developed sophisticated statistical models (e.g., *Hanke et al., 1982; Tampa Bay Water,* 2010). Thus, there is a need for a simple and flexible approach that can model a variety of demand attributes and incorporate climate variables, which can be easily used regardless of the sophistication of the water utilities. Furthermore, the use of time varying climate variables can enable modeling nonstationarity in climate, which the traditional approaches do not consider at all. These important needs and gaps motivated our present research: We motivate the use of Generalized Linear Model (GLM) for modeling demand attributes – monthly average, monthly maximum (or peak) and number of days of exceeding desired threshold, incorporating suite of climate variables.

The paper is organized as follows. We present our methodology with application to data from water utility of Aurora, CO. A description of the case study water demand and climate data set is first presented. This is followed by a proposing a suite of climate attributes that relates to water demand attributes. Next we present the GLM approach to modeling three water demand attributes mentioned above. Results from the proposed method followed by summary and discussion of utility of GLM and further potential applications conclude the paper.

2.2.1 Case Study Data

For this study, daily water demand data from Aurora, CO was collected. We chose this location as it is water stressed with increasing growth and supply reductions. Figure 2.2 shows the study location along with some key water supply related statistics.

2.2.2 Water Demand Data

The data set used for this study is daily production data from Aurora, CO. Water is produced to meet the respective demand and therefore the water production data from this utility can be considered water demand. The data provided was in millions of gallons per day (MGD). The data from Aurora, CO was 21 years in length with daily data from January 1st, 1990 to December 31st, 2010. The daily demand data is shown as boxplots for each month (boxes represent the interquartile range, the whiskers 5th and 95th percentiles and, points beyond) in

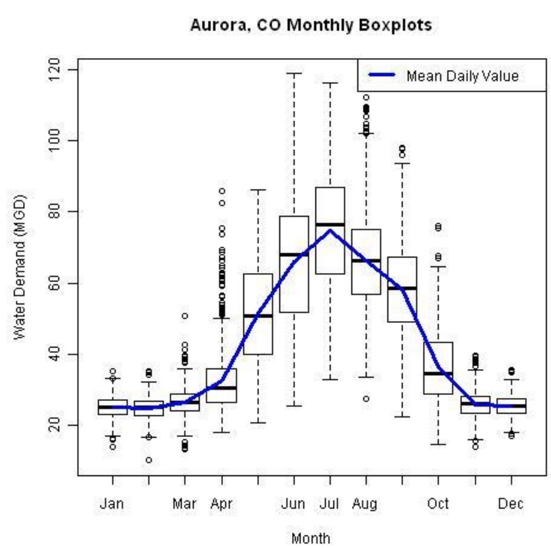


Figure 2.1 – Monthly boxplots of daily water demand values from 1990-2010 of a water utility at Aurora, CO – boxes represent the interquartile range, the whiskers are 5th and 95th percentiles and the points are beyond.

Figure 2.1 (the solid line represents monthly mean). It can be seen that the critical season of high demand is summer, specifically the three month period of June-Aug. From each month in this season, for each year, three water demand attributes were computed – they are, average monthly demand (averaging over all the daily demand values in a month); monthly maximum demand (maximum of all the daily demand values) and number of demand exceedances (number of days with demand exceeding 75^{th} percentile, which is 81.3 MGD, in each month). These are then modeled using GLM and climate attributes as described in the following sections.

2.2.3 Climate Data

Daily meteorological data – precipitation, maximum and minimum temperatures –for the study location is obtained from the National Climatic Data Center (NCDC) archive of National Oceanic and Atmospheric Administration (NOAA) (http://www.ncdc.noaa.gov). The NCDC data archive is extensive, quality checked and offers a variety of meteorological and

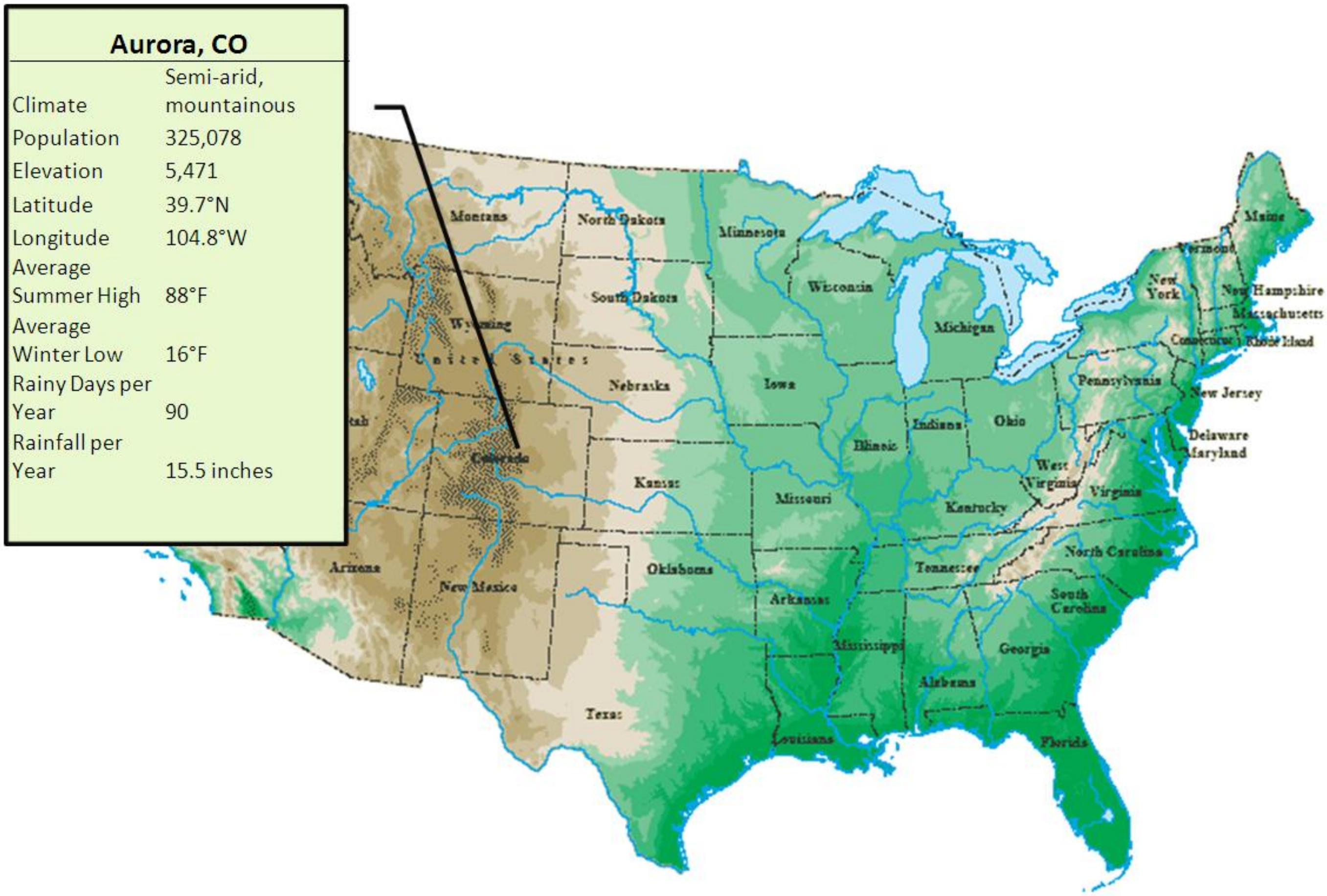


Figure 2.2 – Climate information for the Aurora, CO case study.

climate data for locations throughout the country. Location of the data closes to the study region can be selected from the data base and here we selected location CO-09, which is closest to Aurora, CO. Details of the data set can be found at the aforementioned web location.

2.3 Climate Attributes and Water Demand

Climate's impact on water supply and the associated phenomena are frequently studied. The correlation between water supply and climate is well understood (*Milly,* 1994; *Wu*

Weather Statistic Thresholds
Minimum Temperature

June	57°F
July	62°F
August	60°F
Maximum Temperature	
June	89°F
July	94°F
August	89°F
Precipitation	
June	0 in
July	0 in
August	0 in

Table 2.1 – 75th percentile values for weather in Aurora, CO. These values were used as threshold for developing spell statistics.

et al. 1993). The implications of climate on water demand are intuitive, especially for outdoor water uses, and have also been studied. For example, water demand elasticity has been modeled as a dependent on summer precipitation for all user categories including residential, commercial, industrial and agricultural (*Schneider et al.* 1991). Urban residential water demand has been modeled using rainfall and temperature data (*Nieswiadomy,* 1992). However, sequences of wet/dry and high/low temperature spells events have a much stronger role in modulating water demand, and to this end, spell statistics provide a more nuanced insight into this relationship. For instance, drought has long been considered a primary influence on water demand and can be modeled as a dry spell and incorporated into a water demand model (*Moncur,* 1987, *Arbues et al.* 2003). Similarly, consecutive days of high temperatures can drive evapotranspiration and outdoor water demand (*Smoyer-Tomic,* 2003).

Water Demand Covariates: Definitions and Correlations				
Covariate	**Definition**	**Correlation with Monthly Average Demand**	**Correlation with Monthly Maximum Demand**	**Correlation with Monthly Demand Exceedances**
Average Monthly Max Temperature	Daily maximum temperature values averaged over a month.	0.41	0.35	0.44
Average Seasonal Minimum Temperature	Daily minimum temperature values averaged over a month.	0.44	0.34	0.44
Average Seasonal Precipitation	Daily precipitation values averaged over one a.	0.06	0.04	0.05
TIME	A vector of values progressing chronologically.	0.20	-0.05	0.11
Preceding Month Average Water Demand	Average water demand from preceding month (lag 1).	0.50	0.55	0.46
Average Monthly Hot Spell Length (in days)	The average length in days of a spell with daily maximum temperatures over a given threshold for each month.	0.22	0.25	0.14
Total Monthly Days with Maximum Temperature Above Threshold	Total days the maximum daily temperature exceeds a given threshold in a month.	0.39	0.26	0.24
Maximum Hot Spell Length	The longest spell length of days with daily maximum temperatures over a given threshold in for each month.	0.26	0.25	0.15
Average Monthly Nightly Spell Length	The average length of consecutive days with daily minimum (i.e. nightly) temperatures above a threshold for each month.	0.40	0.39	0.32
Total Monthly Nights with Low Temperature Above Threshold	Total days with minimum temperature above a given threshold for each month.	0.55	0.48	0.44
Maximum Monthly Nightly Low Hot Spell	The longest number of consecutive days with minimum temperatures above a given threshold for each month.	0.53	0.50	0.47
Average Monthly Precipitation Spell in Days	The average length of consecutive days with precipitation above a given threshold for each month.	-0.19	-0.06	-0.15
Total Monthly Days with Precipitation Below Threshold	Total number of days with precipitation above a given threshold each month.	-0.44	-0.22	-0.35
Maximum Monthly Precipitation Spell	The longest number of consecutive days with precipitation above a given threshold each month.	-0.18	-0.05	-0.16

Table 2.2 – Climate Attributes and their correlation with water demand attributes shown in the last three columns.

Thus, we propose a suite of spell statistics for each month in addition to the average statistics computed from the daily meteorological data for potential modeling of water demand. These statistics include spells based on daily maximum temperatures, nightly low

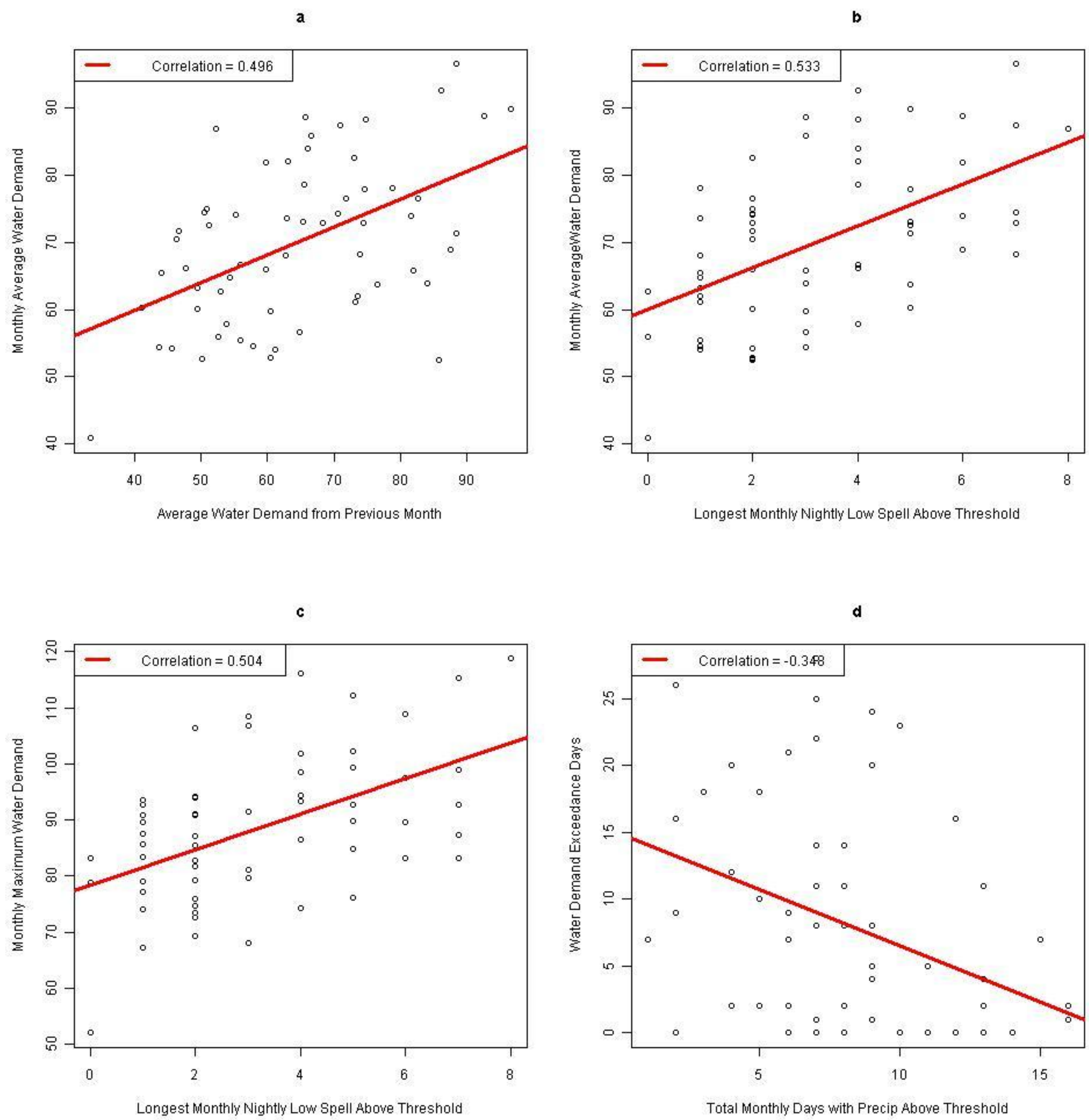


Figure 2.3 – Scatterplot of selected spell statistics and water demand. (a) average water demand in a month versus the same from previous month (antecedent conditions), (b) Longest monthly spell of minimum temperatures above each month's threshold versus monthly average water demand, (c) Longest monthly spell of minimum temperatures above each month's threshold versus monthly maximum water demand and (d) Total number of days with precipitation above 0 inches versus. number of days of water demand exceeding the 75th percentile water demand value for each month.

temperatures and precipitation thresholds. All the climate statistics are listed in Table 2.2 along with their correlation to the water demand attribute. Scatterplots of selected climate attributes and water demand attributes are shown in Figure 2.3. There is a strong positive correlation between average water demands from successive month (Fig 3a) indicative of persistence of antecedent climate conditions. Monthly average and maximum demands have a strong correlation with longest spell of minimum temperatures and, intuitively, water demand exceedance days is negatively correlated with the number of days in a month with precipitation greater than zero.

2.4 Methodology – Generalized Linear Model

Generalized Liner Modeling consists of a response or the dependent variable Y that can be assumed to be a realization from any distribution in the exponential family with a set of parameters (*McCullagh and Nelder*, 1989). A smooth and invertible link function transforms the conditional expectation of Y to a set of predictors (Equation 1).

$$\mathrm{G}(\mathrm{E}(\mathrm{Y})) = \eta = f(\mathrm{X}) + \varepsilon = \mathrm{X}\beta^{\mathrm{T}} + \varepsilon, \tag{2.1}$$

where G(..) is the link function, X is the set of predictors or independent variables, E(Y) is the expected value of the response variable and ε is the error, which is assumed to be Normally distributed with a given variance. If the dependent variable is assumed to be linear—as is the case with standard linear regression, the function G(..) is the identity and Y is assumed to be Normally distributed. However, often assumptions of linearity do not hold up and depending on the assumed distribution (nonlinear) of Y there exist appropriate link functions (*McCullagh and Nelder,* 1989). The model parameters (β) are estimated using an iterated weighted least squares method that maximizes the likelihood function, as opposed to an ordinary least squares method in linear modeling (*Myers et al.,* 2002).

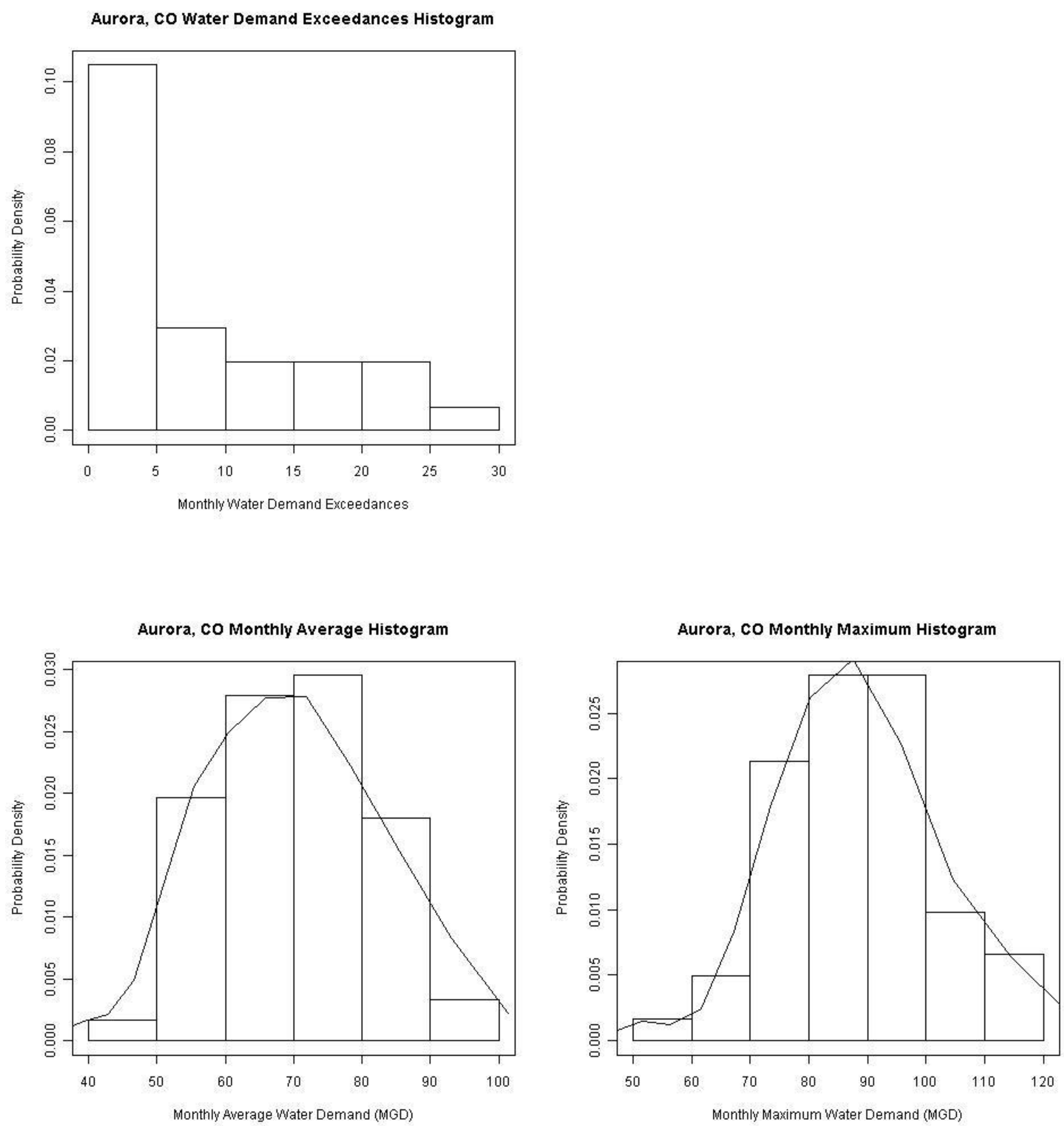


Figure 2.4 – Sample histogram and Probability Density Functions (shown as lines) of (a) number of days of demand exceeding 81.3 MGD (75th percentile value), (b) monthly average demand and (c) monthly maximum demand. The PDFs are based on kernel density estimators.

The GLM can be used to model a variety of response variables – for skewed variables with a lower bound of 0 (such as monthly maximum demand), the Gamma distribution assumption of Y and its associated link function is appropriate. For number of days of demand exceeding a threshold of 81.3 MGD (75^{th} percentile water demand value), which is a discrete variable, the Poisson distribution and its associated link functions can be used. For monthly average demand which has less skew Normal or Gamma distributions are appropriate. If modeling exceedance/non-exceedance of threshold a binomial distribution and its link function (i.e. logistic regression) is the approach. We refer the readers to McCullagh and Nelder (1989) for information about variety of distributions, link functions and parameter estimation.

To obtain the best set of predictors for the model there are a couple of choices for the objective criteria such as the Akaike Information Criteria (AIC) or Bayesian Information Criteria (BIC). In both case, the likelihood function for the number of parameters is penalized to optimize the model (*Venables and Ripley*, 2002). In this, multiple models are fit using all possible subsets of predictors and also link functions. For each model, the AIC and BIC are computed and the model with lowest AIC or BIC is selected as the 'best model'. Models can also be tested for significance against a null model or an appropriate subset model using a chi-squared test. We used BIC in this study as it tends to be more parsimonious compared to AIC.

As mentioned, the GLM gleans skill from its ability to address non-normality and non-continuous features of the dependent variables by using different link functions, such as: (i) Normal distribution for monthly average demand; (ii) Poisson for number of days of demand threshold exceedances, (iii) Gamma distribution for monthly maximum (or peak) demand. To underscore this histogram and probability density functions (PDFs) based on nonparametric smoothing of histograms (*Azzalini and Bowman*, 1989) of three demand attributes are shown in Figure 2.4. Skewed distribution can be seen for number of days of demand exceeding 81.3 MGD and monthly maximum demand, for which Poisson and Gamma distributions in the GLM are appropriate. The monthly average demand is close to Normal.

2.5 Results

We applied the GLM approach to model the three monthly demand attributes with the suite of predictors listed in Table 2.2. In addition to the climate predictors we also use time as a proxy for socio-trend growth and previous month's values for additional temporal dependency. Using BIC the best model, significance of the variables and R-squared for each demand attribute is shown in Table 2.3. Results from the models are shown in four plots; fit (scatterplots of observed versus predicted values), autocorrelation of the residuals, Quantile-Quantile plots of the residuals to test for Normality of the residuals and scatterplot of leave one out cross validation (LOOCV) estimates with the observed. In LOOCV each observation is dropped and the model fitted on the remaining data and subsequently, estimating the dropped point. While dropping a single point might not stress the model as much, we chose this due to limited data, dropping a fraction of observations and repeating a number of times to obtain distribution of skill measures would be much preferred when the data set is large (e.g., *Zachman et al.,* 2007, *Towler et al.,* 2007).

2.5.1 Monthly Average Demand

Monthly average demand is modeled using a Gamma distribution with its canonical link function; the inverse link function (*McCullagh and Nelder,* 1989). For monthly average longest

Aurora, CO					
Monthly Average Demand					
Independent Variable	Coefficient Estimate	Standard Error	t Value	Pr(>|t|)	R^2
Y Axis Intercept	5.19E+01	5.36E+00	9.69	1.17E-13	5.48E-01
Longest Spell with Min Temp Above Threshold	1.72E+00	5.99E-01	2.87	5.69E-03	
Total Days with Precip Above Threshold	-1.26E+00	3.05E-01	-4.13	1.18E-04	
Average Water Demand from Preceding Month	3.58E-01	8.08E-02	4.43	4.36E-05	
Monthly Maximum Demand					
Independent Variable	Coefficient Estimate	Standard Error	t Value	Pr(>|t|)	R^2
Y Axis Intercept	2.30E+01	2.75E+01	0.84	4.07E-01	3.36E-01
Longest Spell with Min Temp Above Threshold	2.91E+00	7.44E-01	3.92	2.44E-04	
Monthly Average Min Temp	1.09E+00	5.17E-01	2.10	3.99E-02	
TIME	-4.50E-02	2.11E-02	-2.14	3.69E-02	
Number of Days per Month with Demand Exceeding a Given Threshold					
Independent Variable	Coefficient Estimate	Standard Error	z Value	Pr(>|t|)	R^2
Y Axis Intercept	-4.20E+00	9.86E-01	-4.26	2.05E-05	4.09E-01
Average Spell with Max Temp Above Threshold	2.34E-01	7.90E-02	2.96	3.03E-03	
Longest Spell with Max Temp Above Threshold	-2.20E-01	4.40E-02	-5.00	5.70E-07	
Total Days with Min Temp Above Threshold	6.97E-02	1.72E-02	4.06	4.82E-05	
Longest Spell with Min Temp Above Threshold	8.52E-02	3.18E-02	2.68	7.44E-03	
Average Spell with Precip Above Threshold	-5.43E-01	1.59E-01	-3.41	6.43E-04	
Total Days with Precip Above Threshold	-1.50E-01	2.34E-02	-6.42	1.33E-10	
Longest Spell with Precip Above Threshold	3.00E-01	6.78E-02	4.43	9.36E-06	
Monthly Average Min Temp	1.28E-01	1.84E-02	6.96	3.54E-12	
TIME	-3.48E-03	8.87E-04	-3.93	8.62E-05	

Table 2.3 – Best GLM models selected using BIC for the three demand attributes. The Pr(>|t|) value notifies significance for the coefficient estimate.

spell with minimum temperature below the month's 75th percentile value and number of days with precipitation above zero inches are the main climate variables, in addition to average demand from the previous month. The precipitation provides the moisture input while temperature spell provides information on losses from evaporation. The model does very well

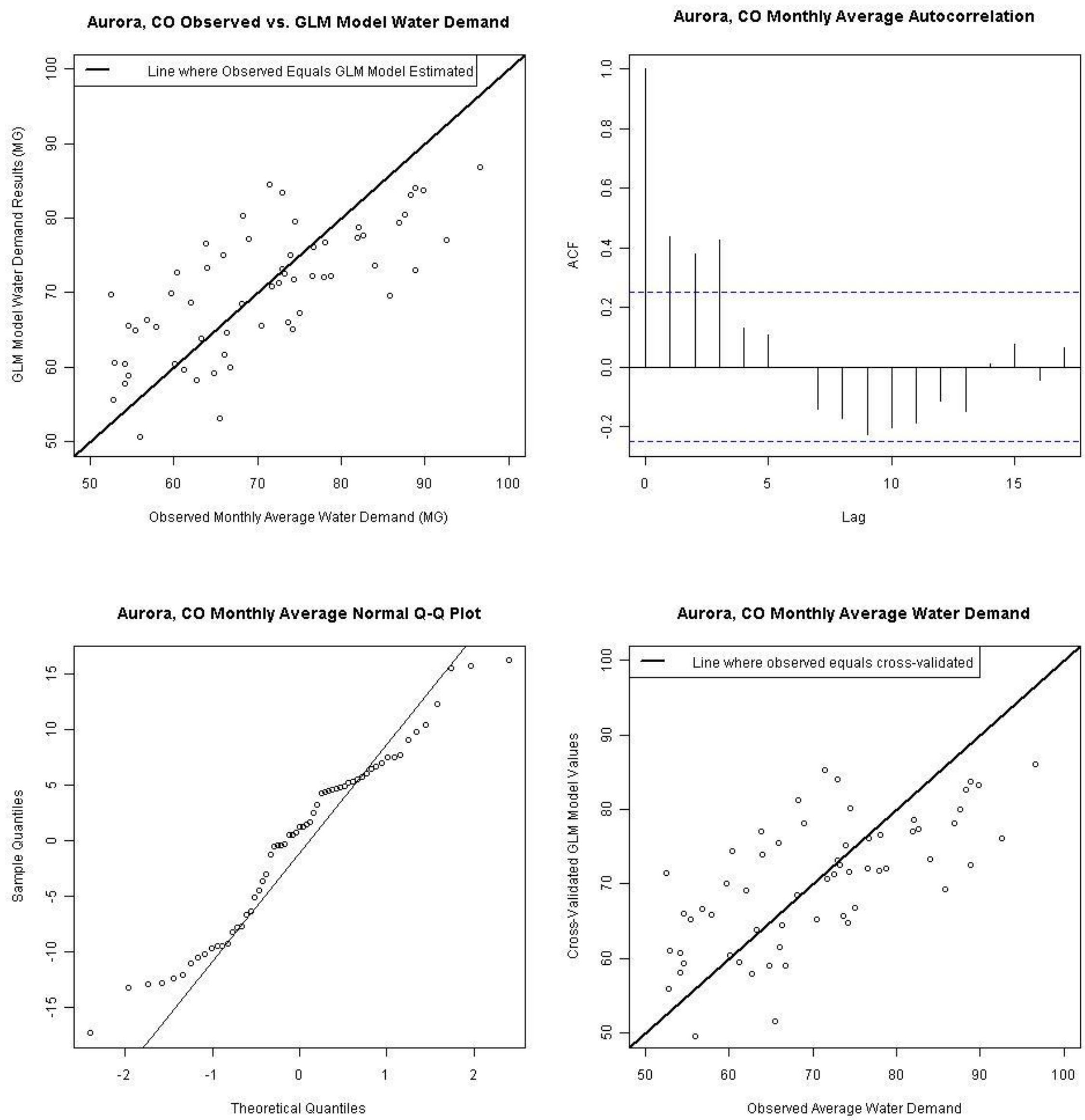
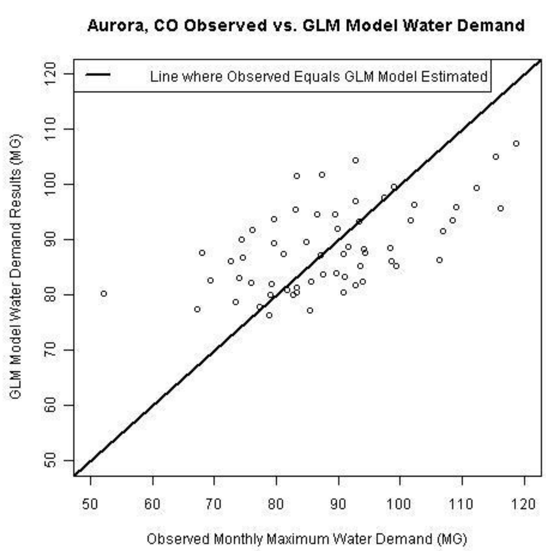
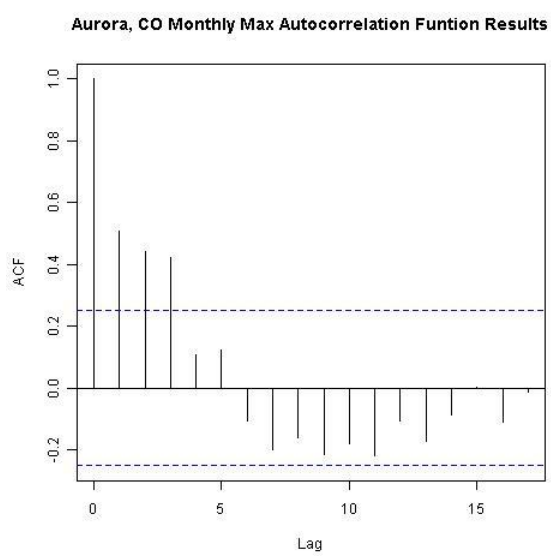


Figure 2.5 – Diagnostics plots for monthly average demand – (a) Scatterplot of observed vs. model estimates with the 1:1 line, (b) autocorrelation of model residuals, (c) Q-Q plot of model residuals and (d) leave one out cross validated estimates versus observed values with the 1:1 line.

in estimating the monthly average demand with the correlation between observed and model estimates at 0.739 (Figure 2.5). There is a slight underestimation of higher average demand values. Autocorrelation of residuals at lags 1-3 are slightly significant, the time predictor could help but BIC optimization selected to neglect it. The Q-Q plot indicates that the residuals are quite Normal. The cross validated estimations also perform well, suggesting that the model has good predictive capability with the correlation of the cross-validated model versus observed of 0.694.

2.5.2 Monthly Maximum Demand





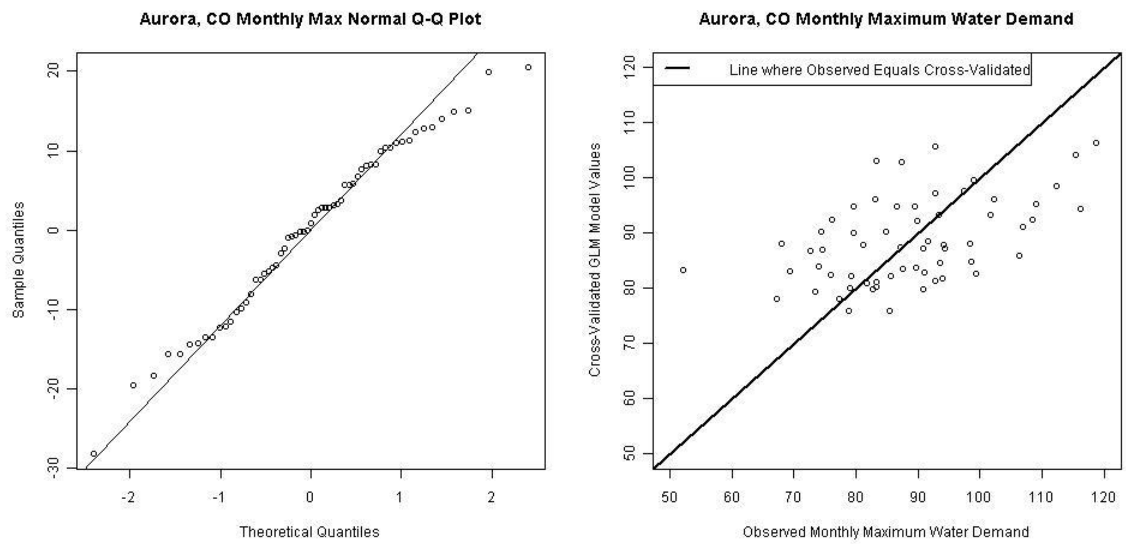


Figure 2.6 – Same as Figure 4 but for monthly maximum demand.

Here too the Gamma distribution with an inverse link function is used in the GLM. For the monthly maximum demand average minimum temperature, time (a proxy for socio-economic trend) and longest spell of minimum temperature below each month's 75th percentile are the predictors. It is interesting to note that in addition to climate, maximum demand increases with growth and hence, time is selected as one of the predictors. The model performance (Figure 2.6) is quite similar to that of the monthly average described above. Apart from the over estimation of few points at the low end the model does very well. This is largely due to the Gamma distribution as it has a lower bound of zero and a positive tail.

2.5.3 Number of days of exceeding threshold

Using the Poisson distribution with the log link function in the GLM is appropriate for discrete variables—which this facet of water demand is. The results of this GLM are seen in figure 2.7. The number of days of demand exceeding a threshold of the 75th percentile value (81.3MGD) is somewhat of a noisy variable, hence, a number of variables are necessary to model it as can be seen in Table 2.3. Note that using a Poisson distribution in the GLM the estimates are discrete and obviates rounding to integers as would be necessary if using traditional linear regression models. The model performance is quite good with low autocorrelation and normality of residuals.

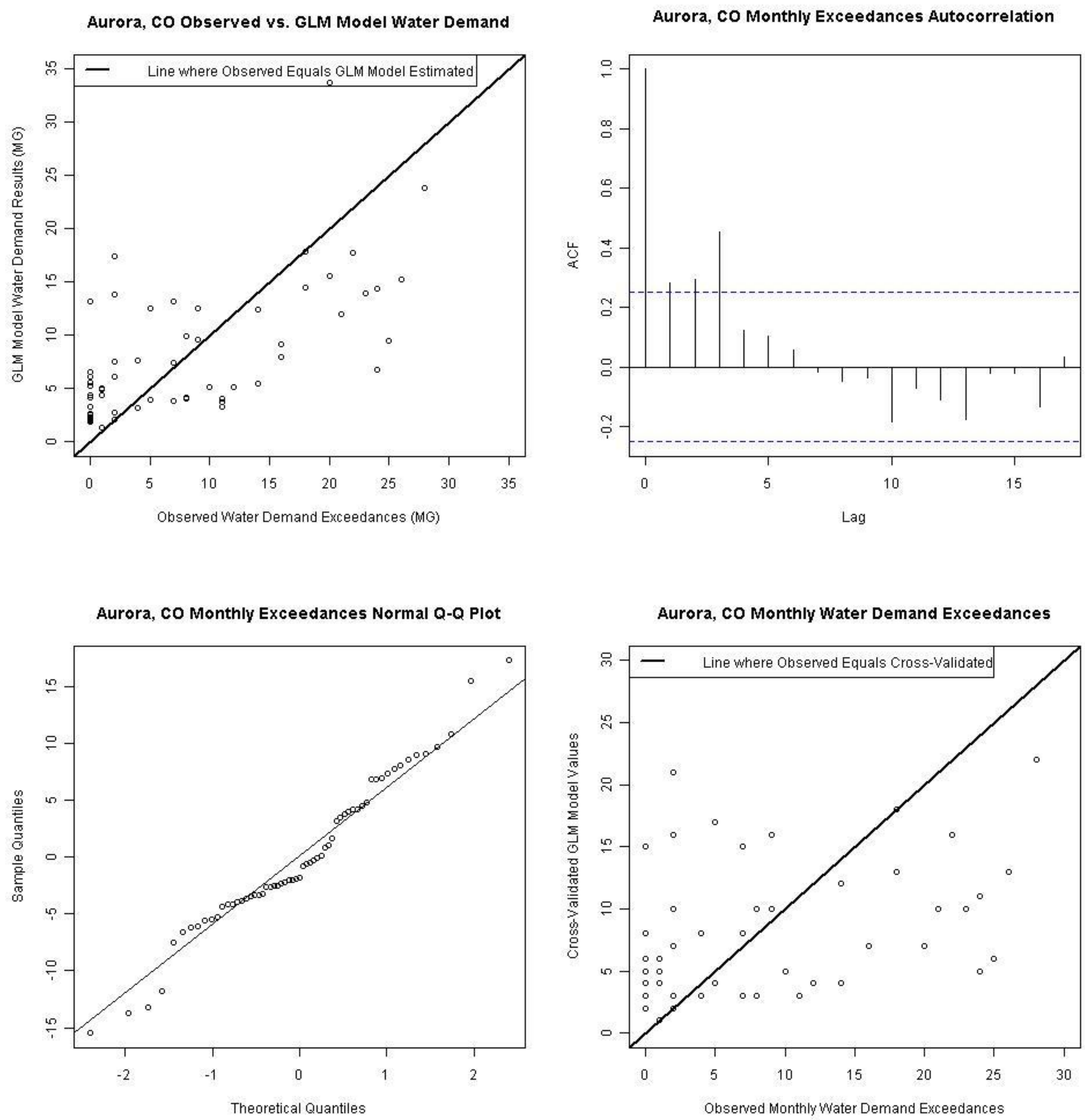


Figure 2.7 – Same as Figure 4 but for number of days of demand exceeding a threshold of 81.3 MGD (75th percentile value).

2.6 Summary and Discussion

Using the GLM and climate attributes to model water demand provides multiple advantages to traditional methods. First, climate data is easily ascertained and from the data climate attributes can be simply calculated as opposed to socioeconomic variables, which can be difficult to define and the data hard to collect. Secondly, the inherent flexibility of Generalized Linear Modeling lends itself well to water demand modeling as the water demand variable of interest may be nonnormal or discrete. The GLM modeling results presented suggest that using climate covariates to describe monthly average and monthly maximum water demand is more parsimonious than using them to model monthly exceedance days. This shortcoming is fundamental to modeling monthly exceedance days, which is a particularly noisy water demand variable. Approaches to circumventing this issue are subsequently discussed.

The developed models can be applied for short term (seasonal) management by incorporating seasonal climate forecasts – such applications are common in water resources management (e.g., *Regonda et al.,* 2011) and recently for water quality management (*Towler et al.,* 2010).

While the demand attributes considered in this research are relevant for utility managers, a more important variable is the variability of 'extreme demand events'. In this, the managers would like to know the fluctuations in peak demand, threshold exceedance and length of exceedances. While individual components can be modeled using GLM as described above, extreme value analysis offers an attractive alternative to model these in an integrated manner incorporating predictor variables (e.g., *Towler et al.,* 2010; *Katz*, 2010; *Katz and Naveau,* 2002). Future research will broach the topic of water demand extremes modeling.

Water demand projections are important for efficient management and planning of water supply. Traditionally these projections are made using statistical and econometric models that account for socio-economic growth and limited climatic variability. Given the sophistication and uncertainty in economic projections these models are accessible only to relatively well-off utilities. We see two major gaps – (i) need for an approach to model a variety of demand attributes that are needed by utility managers such as – average demand, monthly peak and number of demand exceedances above a threshold and (ii) comprehensive use of climate information. We first identify a suite of climate variables that include high/low temperature and wet/dry spells and the more traditional average temperature and precipitation. We find that these variables capture nuanced connections between climate and water demand, especially the lagged relationships. Second we propose a GLM framework to model attributes of water demand by incorporating the climate variables. We demonstrate the utility of this approach to demand modeling for a water utility in Aurora, CO. The various dependent variables (monthly average demand, monthly peak demand and monthly number of demand exceedances above a threshold) are modeled quite well and the models have the ability to capture a variety of water demand attributes and distributions. Additional socio-economic predictors can be easily incorporated, although here we use 'time' as a proxy for this. Furthermore, the GLM based modeling approach with climate variables provides the capability to make future water demand projections under climate change. In this, the climate attributes can be derived under climate change scenarios which can then be used with the developed GLM models to make demand projections. Above all, this method is simple and can run with limited computational resources; thus, accessible to all utilities.

3 Modeling Water Demand Extremes Using Extreme Value Analysis

3.1 Abstract

Understanding and modeling variations in water demand extremes are important for efficient utility management and reliability of water supply, as inaccurate projections can lead to suboptimal infrastructure and management practices. Traditional methods using linear statistical models capture only the average demand variability, however water utilities make decision based on peak (or extreme) demands. We demonstrate the use of the emerging field of Extreme Value Analysis to model water demand extremes incorporating climate covariates. Two approaches in this analysis are used – (i) Generalized Extreme Value analysis to model monthly peak water demand and (ii) Points Over a Threshold analysis to model frequency of demand exceedances above a threshold and the magnitude of exceedances. Climate covariates can be incorporated in these models to capture nonstationarity due to climate change. We demonstrate the utility of this approach by applying it to model water demand extremes for Aurora, CO.

3.2 Introduction

Understanding variability in water demand and being able to make skillful projections of its future variability is crucial for efficient water supply management and for water demand forecasting. Water demand forecasts enable water utilities to plan for capacity expansion, pricing, conservation efforts, etc. (*Billings et al.*, 2008). Demand is influenced by a large number of factors from the socio-economic realm – employment, economic activity, population, infrastructure efficiency, geographic constraints, and also the weather and climate realm –

meteorological and hydrologic variability. Most of the demand models are based on socio economic factors (e.g., *Arbués et al.* 2003; *Katzman*, 1977; *Babel et al.*, 2007). Socio-economic projections such as population and economic growth are highly uncertain which translates into uncertain water demand projections (*Arbués et al*. 2003). Climate variability plays a strong role in modulating water demand and this varies with location: for example, increased temperatures lead to increased evapotranspiration and therefore to higher agricultural and municipal demand, likewise increased precipitation reduces the plant water demand and consequently the overall demand of the water utility (*Billings*, 2008). Average measures of precipitation, temperature, variations of these and rudimentary measures of spells such as number of rainy days have all been used in the past to model water demand (e.g., *Smoyer-Tomic*, 2003; *Foster* et al., 1981; *Billings*, 1982). Furthermore, seasonal or monthly average demand has largely been the focus of current modeling efforts (*McKee et al.*, 1995; *Hanke et al.*, 1982). In Chapter 2 a suite of weather attributes that include hot/cold and wet/dry spells was developed to model aspects of monthly water demand with substantial skill. The existing water demand modeling efforts are mainly for average demands incorporating basic weather attributes.

A key attribute of water demand for planning and management is the extreme water demand events. Specifically, peak demand (either peak day and/or seasonal peak) is used for treatment plant capacity evaluation, network design and for securing capital for new infrastructure, capacity expansion, etc. (*Arbués et al*., 2003, *Billings et al.*, 2008). Also of interest is the interval of time with elevated demand or demand above a critical threshold. Weather attributes do a lot to modulate the water demand extremes. Current models are limited in their ability to model extremes and this is a major gap in demand management, which motivates this research. The emerging field of Extreme Value Analysis (EVA) offers an attractive, simple and robust approach to modeling demand extremes that can be adapted with a wide range of utilities of economic resources. The desire to better model hydrologic risks was the early motivation for EVA (*Gumbel*, 1941). It has since matured and is now widely used in modeling hydrologic, meteorologic and sea level extremes (*Fiorentino*, 1984; *McNeil*, 1997; *Swift*, 1989; *Beguería*, 2006; *Coles*, 2003; *Katz*, 1995; *Eaton*, 1994; *Coles*, 2001; *Smith*, 1989); financial risk estimation (e.g., *Embrechts*, 1999); and for hydrological losses (*McNeil*, 1997).

Weather Statistic Thresholds	
Minimum Temperature	
June	57°F
July	62°F
August	60°F
Maximum Temperature	
June	89°F
July	94°F
August	89°F
Precipitation	
June	0 in
July	0 in
August	0 in

Table 3.1 - 75th percentile values for weather in Aurora, CO. These values were used as threshold for developing climate attributes.

At the core of the EVA is the recognition that extreme events by definition occur in the tail of the probability density function (PDF); that the traditional PDFs have a thin tail thereby, underestimating the probability of extreme events; thus heavy tailed PDFs are necessary, which are developed from theory (*Katz and Naveau*, 2002; *Coles*, 2001). Extreme events also vary in time and space and consequently the probability of their occurrences – to capture this variability (i.e., nonstationarity) the PDFs should also vary. To this end, EVA proposes modeling of the parameters of the heavy tailed PDFs as a function of covariates. In this research, the covariates are the weather attributes identified in recent research (see *Chapter 2*).

We introduce EVA for modeling water demand and hope that this initiation to the field will lead to further application and expansion of this approach. The paper is structured as follows: We introduce the readers briefly to the two extreme value analysis methods – block maxima and Points Over a Thresholds (POT) – in this we also describe the use of these methods with covariates to model nonstationary features. We then demonstrate the utility of these

Weather Attributes (Covariates)	**Definition**	**Correlation with Monthly Maximum Demand**
Average Monthly Max Temperature	**Daily maximum temperature values averaged over one month.**	0.35
Average Seasonal Minimum Temperature	Daily minimum temperature values averaged over one month.	0.34
Average Seasonal Precipitation	Daily precipitation values averaged over one month.	0.04
TIME	A vector of values progressing chronologically.	-0.05
Preceding Month Average Water Demand	**Average water demand from preceding month (lag 1).**	0.55
Average Monthly Daily High Hot Spell in Days	The average length in days of a spell with daily high temperatures over a given threshold for each month.	0.25
Total Monthly Days with High Temperature Above Threshold	Total days the maximum daily temperature exceeds a given threshold in a month.	0.26
Maximum Monthly Daily High Hot Spell	The longest spell with daily high temperatures over a given threshold in a given month.	0.25
Average Monthly Nightly Low Hot Spell in Day	The average length in days of a spell with nightly low temperatures above a threshold for each month.	0.39
Total Monthly Nights with Low Temperature Above Threshold	**Total days with the nightly low temperature above a given threshold for each month.**	0.48
Maximum Monthly Nightly Low Hot Spell	The longest spell in days with nightly low temperatures above a given threshold for each month.	0.50
Average Monthly Precipitation Spell in Days	The average length in days of a spell with precipitation above a given threshold for each	-0.06

	month.	
Total Monthly Days with Precipitation Below Threshold	**Total days with the precipitation above a given threshold each month.**	-0.22
Maximum Monthly Precipitation Spell	The longest spell in days with precipitation above a given threshold each month.	-0.05

Table 3.2 – Weather attributes and their correlation with monthly maximum water demand. Bolded values are statistically significant at 95% confidence level and, bolded rows are variables used in the nonstationary GEV as covariates.

methods to modeling water demand extremes from water utility of Aurora, CO and conclude with a summary and discussion of the results and the methods.

3.3 Methods for Modeling Extremes - Extreme Value Analysis

The approach methodologies used in this study are detailed in this section. Information regarding climate data collection, weather spell statistics extrapolation, correlation methods and an introduction to the Monthly Maxima approach and the POT approach can be found in this section.

As described above, understanding the distribution of extreme events is crucial for efficient water resources management. There are two types of extremes that would be of interest to utilities – (i) the extreme demand over a certain period or block, such as a month (block maxima approach) and (ii) the frequency of exceeding a fixed threshold and the magnitude of exceedances (POT approach). Below we describe methods to model these two types.

3.3.1 Modeling Block Extremes - Generalized Extreme Value Approach

Extreme values can be thought of as occurrences from an underlying distribution. However, when the maximum of a sample (or block) is considered results from extreme value theory suggests that the block extremes follow a Generalized Extreme Value (GEV) distribution in the limit (that is in the presence of large number of observations, a sufficiently large block and assuming the extremes to be independent and identically distributed) (*Leadbetter*, 1983; *Coles*, 2001). The GEV distribution is given with the following cumulative distribution function;

$$G(z;\theta) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right]\right\} \tag{3.1}$$

where $\theta = [\mu, \sigma, \xi)$ are the location, scale and shape parameters, respectively, and z is a value of the seasonal monthly water demand maxima. This notation is most common in statistical literature; however, in hydrology it is common for the notation to include $-\xi$ instead of ξ (*Katz*, 2002). The location parameter, μ, indicates the center of the distribution, the scale parameter, σ, represents the spread of the distribution, and the shape parameter, ξ, represents the behavior of the distribution in the tail. Positive value of shape parameter indicates heavy tail, a value of zero suggests thin tail and a negative value points to a bounded tail. Also, for the heavy-tailed case, it must be true that,

$$\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right] \geq 0.$$

The parameters are estimated using Maximum Log Likelihood (MLE) procedure, wherein parameters are obtained as those that maximizes the log likelihood function (*Coles*, 2001; *Katz and Naveau*, 2002). The above form of GEV is fitted to the entire data resulting in a single model to describe its distribution. However, the extreme values show significant variability and often the variability is driven by external factors including trends. This is true with many natural phenomena such as floods, droughts and certainly is the case with water demand. Modeling such variability means obtaining an appropriate GEV for each value, also known as a nonstationary GEV. To achieve this in a parsimonious manner, the parameters of the GEV distribution are allowed to vary as a function of covariates that orchestrate the variability of the extremes (*Katz and Naveau*, 2002). The model is given as,

$$\mu(t) = \beta_{0,\mu} + \beta_{1,\mu}x_1 + \cdots + \beta_{n,\mu}x_n \tag{3.2}$$

$$\log[\sigma(t)] = \beta_{0,\sigma} + \beta_{1,\sigma}x_1 + \cdots + \beta_{n,\sigma}x_n \tag{3.3}$$

$$\xi(t) = \xi \tag{3.4}$$

where the variables $x_1, x_2, \dots, x_n$ are the covariates which can include 'Time' to consider a temporal trend, and the betas are the coefficients of the regression. Shape parameter is generally not varied as it tends to be noisy, but it too can be varied as the other two. Notice that the above form is reminiscent of a Generalized Linear Model (GLM, *McCullagh and Nelder*, 1989). The coefficients are estimated using the maximum likelihood approach that is commonly used in GLM (*Katz and Naveau*, 2002). The best combination of covariates, if there are several of them, is selected using a likelihood ratio test. In this, models are fit with various combinations of the covariates and are then compared in pairs on their values of likelihood function, which selects a parsimonious model (*Katz and Naveau*, 2002; *Coles*, 2001).

Variations of the nonstationary GEV idea have been developed for flood frequency modeling (*Sankarasubramanian and Loll*, 2003; *Rajagopalan et al*., 2010) and recently Towler et al., 2010 applied this for modeling maximum streamflow and consequently water quality extremes – for both seasonal time scale and multidecadal time scales under climate change.

The two forms of GEV modeling provide a strong capability to characterize extremes at a variety of time scales that will be of use in water supply planning.

3.3.2 Model for Threshold Exceedance of Extremes - Points Over a Threshold (POT) Approach

Another attribute of extremes, in addition to the extreme value in a block described in the previous section, is exceedances with respect to a threshold. Typically, if values exceed a threshold they have direct consequences – in the case of floods it is the risk to infrastructure and human life, in the case of insurance damage it can be risk to payouts by the insurance company and in the case of water demand it can be the risk to adequate water supply or additional expense to meet high demand. In modeling threshold exceedances there are two aspects – (i) the probability of exceedance or non-exceedance of the threshold described as a Poisson Process (PP) and (ii) the magnitude of exceedances when there is an exceedance described by a Generalized Pareto Distribution (GPD) – similar to the GEV (see *Coles*, 2001; *Katz and Naveau*, 2002).

The GPD cumulative distribution function for values exceeding a threshold is as follows,

$$F(x;\sigma,\xi) = 1 - \left[1 + \xi(\frac{x}{\sigma})\right]^{-\frac{1}{\xi}} \tag{3.5}$$

where, similarly to the GEV distribution, σ is the scale parameter and ξ is the shape parameter with the following necessarily being true,

$$\sigma > 0,$$

$$1 + \xi\left(\frac{x}{\sigma}\right) > 0.$$

The scale parameter can be allowed to vary with respect to a chosen set of covariates to create a non-stationary model useful for understanding evolving risks.

The threshold can be user specified—something that might have significance in a decision making context—or it can be selected in a quasi-objective manner. Wherein, different thresholds are selected and for each the scale parameter is estimated and the region where the parameter remains stable is used to select the threshold. A higher threshold results in fewer exceedances and a lower one makes every event an extreme. Thus we suggest selecting a threshold based on the system requirements and the parameter sensitivity.

To model variability in the exceedances the parameters of the GPD need to vary with covariates. The scale parameter can be modeled in a GLM framework, similar to the nonstationary GEV described in the previous section, as:

$$\log[\sigma(t)] = \beta_{0,\sigma} + \beta_{1,\sigma}x_1 + \cdots + \beta_{n,\sigma}x_n \tag{3.6}$$

$$\xi(t) = \xi \tag{3.7}$$

where the variables $x_1, x_2, \ldots, x_n$ are the covariates which can include 'Time' as one of them to consider a temporal trend, and betas are the coefficients. The shape parameter is generally not varied as it tends to be noisy, but it too can be varied as the other two if desired.

As mentioned above a companion procedure for the occurrence of exceedance is necessary to complete the model. The occurrence is typically modeled as a Poisson Process (*Katz and Naveau*, 2002; *Coles*, 2001). The rate parameter, λ, is generally obtained as the average number of occurrences. Nonstationarity in this modeled similarly using a GLM (also known as Poisson regression);

$$\log(\lambda) = \beta_{0,\lambda} + \beta_{1,\lambda}x_1 + \cdots + \beta_{n,\lambda} \tag{3.8}$$

where the variables $x_1, x_2, \ldots, x_n$ are the covariates.

A point process approach has also been proposed (*Katz and Naveau*, 2002; *Coles*, 2001) to model both, the occurrence of exceedances and the magnitudes with a GPD, simultaneously. This has flexibility in terms of relating to GEV and in estimating uncertainties.

Here too Maximum Likelihood Estimation (MLE) method is used to estimate the parameters of the GPD and the Poisson Process. As is the case of GEV, models are fitted with different combinations of covariates and the pairwise comparison using likelihood ratio test is performed to consequently select the best model. All of these extreme value analysis methods can be easily implemented in the software and programming language R (*Stephenson and Gilleland*, 2006; *Gilleland and Katz*, 2005).

3.4 Application and Data

The data set used for this study is daily production data from the city of Aurora, CO in millions of gallons per day (MGD). Water is produced to meet the respective demand and therefore the water production data from this utility can be considered water demand. The data covers the period from January 1st, 1990 to December 31st, 2010. The critical season of high demand is summer, specifically the three month period of June through August. For the block maxima analysis monthly maximum demand is computed from the daily data. For POT a threshold of 90 MGD is chosen – this threshold could be related to a specific utility's capacity.

This threshold selection is reasonable for many reasons; it is well within the safe sensitivity range for GPD modeling and it also is sufficiently high to be of concern for the utility and will be described in the following section. As has been mentioned, threshold selection can be varied based on the particular needs of an individual utility.

To incorporate climate attributes as covariates to model nonstationarity, daily meteorological data—precipitation, maximum and minimum temperatures—for the study location is obtained from the National Climatic Data Center (NCDC) archive of National Oceanic and Atmospheric Administration (NOAA) (http://www.ncdc.noaa.gov). The NCDC data archive is extensive, quality checked and offers a variety of meteorological and climate data for locations throughout the country. Location of the data closest to the study region can be selected from the data base. Details of the data set can be found at the aforementioned web location. We computed hot/cold and dry/wet spells with the thresholds shown in Table 3.1. In Chapter 2, the use of weather spells for modeling demand attributes is motivated. A suite of weather attributes are developed and their correlations with monthly maximum water demand are shown in Table 3.2. For computational ease variables with significant correlations are selected for use in modeling water demand extremes with nonstationarity.

3.5 Results

The results are presented for both the block maxima and the POT methods for the stationary and nonstationary cases.

3.5.1 Monthly Maximum Demand

A GEV model was fitted to the block maxima with the block being a month in this case. Monthly maximum water demand, in MGD, was computed from the daily demand data for the three months (June-August) over the 21 year period, thus providing 63 values (Figure 3.1). Given the paucity of data we chose a smaller block of a month – this was used in Towler, 2010 , for modeling monthly maximum of turbidity in drinking water supply. In this study they computed the maximum flow for each of the winter months (Nov-Mar) and fitted a GEV to them. Using months as a block can circumvent issues created by smaller data sets such as the case here.

The stationary GEV model was fit using the methodology described in the previous section – see Table 3.3 for parameter values. The diagnostics of the fit (see *Chapter 3*) show that the observed quantiles are well captured by the model and corroborated by the Probability Density Function (PDF) fits.

Nonstationarity in the monthly water demand extremes is quite apparent from Figure 3.1. To model this, nonstationary GEV models have to be fitted. Following the method described in

		Non-Stationary Models		
Variable	Stationary Model	Varying Scale Parameter: $\sigma = \beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)$	**Varying Location Parameter:** $\boldsymbol{\mu = \beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)}$	$\sigma = \beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)$ & $\mu = \beta_5(a) + \beta_6(b) + \beta_7(c) + \beta_8(d)$
β_1	-	-0.02	**0.36**	0.57
β_2	-	0.20	**-1.65**	-0.93
β_3	-	0.07	**0.28**	-0.44
β_4	-	-0.10	**0.40**	0.37

β_5	-	0.02	**-0.02**	0.17
β_6	-	-	**-**	-0.02
β_7	-	-	**-**	0.51
β_8	-	-	**-**	0.05
Location Parameter (μ)	65.27	65.53	**31.41**	83.10
Scale Parameter (σ)	12.17	7.86	**8.40**	-38.25
Shape Parameter (ξ)	-0.29	0.05	**-0.42**	-0.47
Negative Log Likelihood	246.70	249.91	**219.62**	224.96
Significance (p-value)	-	Fail (0.159)	**Yes (0.005)**	Fail (0.215)

**Where (a) is the total monthly days with minimum temperatures above a threshold, (b) is the total monthly days with precipitation above the threshold, (c) is the monthly average maximum temperature and (d) is the preceding month average water demand.

Table 3.3 – Generalized Extreme Value Distribution model results. The column in bold is the best fit model.

the previous section the location and scale parameters of the GEV were allowed to each vary individually and then both vary simultaneously, resulting in three different non-stationary models with statistically significant weather attributes from Table 3.2 as covariates (Equations 2-3). The best model was chosen based on its significance in a likelihood ratio test and the diagnostic plots. Here the resulting best model was one in which the location parameter varied with covariates, shown in bold in Table 3.3. This nonstationary model is used to compute the GEV parameters for each month and consequently, the various quantiles. This is shown in Figure 3.2 where the dotted lines correspond to the 0.05 and 0.95 quantiles, the solid line is the observed monthly peak demand; the dashed lines are the quantiles from the stationary GEV. Notice that the quantiles from the stationary model remains fixed while the nonstationary model captures

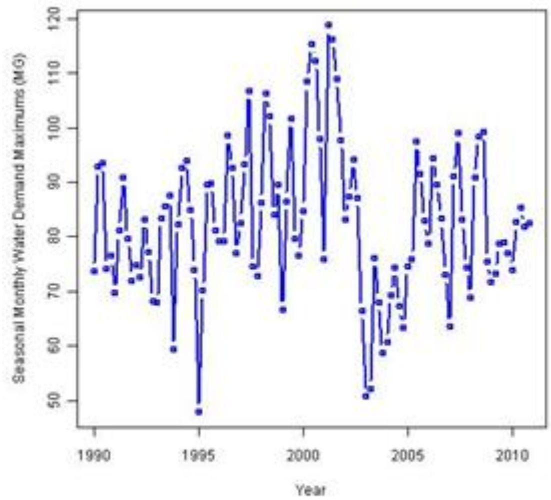


Figure 3.1 – Monthly maximum water demand (MGD) (Jun-Aug).

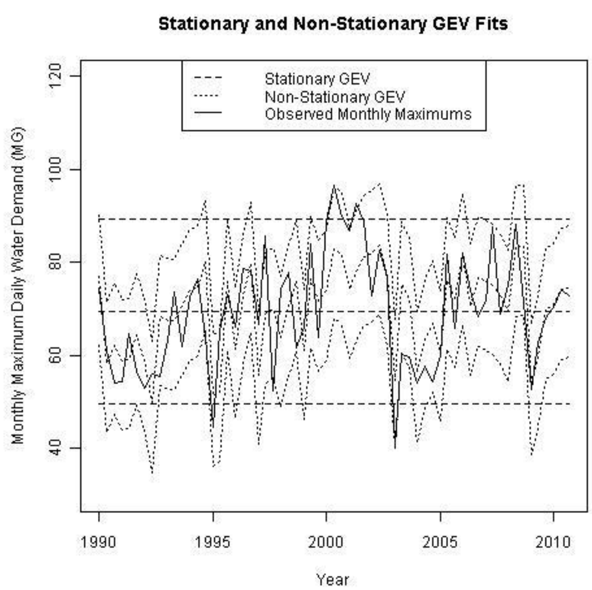


Figure 3.2 – Quantiles from stationary and nonstationary GEV models. Solid line shows the monthly maximum water demand, dotted lines are the 0.05, median and 0.95 quantiles from the nonstationary GEV, and dashed lines are the quantiles from stationary GEV model.

the variability of the observed peak demand very well. This ability to capture nonstationarity has significant implications for realistic assessment of demand risks that can be of immense help in planning decisions; because water demand in not a stationary phenomenon, it is essential to have models that can incorporate covariates and become nonstationary. With seasonal climate forecasts, this modeling approach can be used to make projections of distribution of peak demand and therefore help in devising effective decision strategies. This approach and its variations have been developed for nonstationary flood frequency forecast (*Rajagopalan et al.*, 2010; *Arumugam and Loll*, 2006). In this research we have one of the earliest arguments for application in water demand modeling.

3.5.2 Threshold Exceedances of Water Demand

Often water managers would like to know the risk of demand exceeding a certain threshold that has operation implications. As described earlier the POT approach enables modeling this in a comprehensive manner. Figure 3.3 shows the time series of daily water demand during the months of June-August for all the years with a threshold of 90 MGD shown as a dotted line. The threshold can be user specified, something that might have significance in a decision making context, or it can be selected in a quasi-objective manner. Different thresholds are selected and for each selected threshold the scale parameter is estimated as shown in Figure 3.4, the stable region extends from 70MGD through 100MGD after which the scale parameter changes sharply. A higher threshold results in fewer exceedances and a lower one makes every event an extreme. Thus we suggest selecting a threshold based on the system requirements and the sensitivity plots such as Figure 3.4 – this led to the threshold of 90MGD in the present study.

Variable	Stationary Model	**Non-Stationary Model** $\boldsymbol{\sigma = \beta_1(T) + \beta_2(MaxT) + \beta_3(P) + \beta_4(MinT)}$
β1	-	**0.59**
β2	-	**-0.08**
β3	-	**1.73**
β4	-	**-0.38**
Scale Parameter (σ)	9.44	**86.80**
Shape Parameter (ξ)	-0.26	**-0.10**
Negative Log Likelihood	653.16	**640.50**
Significance (p-value)	-	**Yes (0.000)**

*** (T) = Time, (MaxT) = Daily Maximum Temperature, (P) = Precipitation Days Below Threshold, (MinT) = Daily Minimum Temperature

Table 3.4 – Generalized Parateo Distribution model results. The nonstationary model statistically significant in comparison to the stationary model and is shown in bold.

We fit a GPD model to all the exceedances of 90MGD – i.e., stationary model, the parameter estimates of which are shown in Table 3.4. It can be seen that the model performs quite well in terms of capturing the distribution of the observed daily demand data. The daily demand data and the 0.05 and 0.95 quantiles are shown in Figure 3.5 as solid and dashed lines, respectively. We see that the variability of the demand over time is not well captured, which necessitates the nonstationary GPD model. As with the GEV, models for varying the scale parameter are developed with combinations of the predictor variables and using a likelihood ratio test the best model is selected and is shown in Table 3.3. The shape parameter is generally not varied, especially with short data length, because it can be quite noisy and thus change the GPD distribution erratically. In the nonstationary model the aim is to model each daily

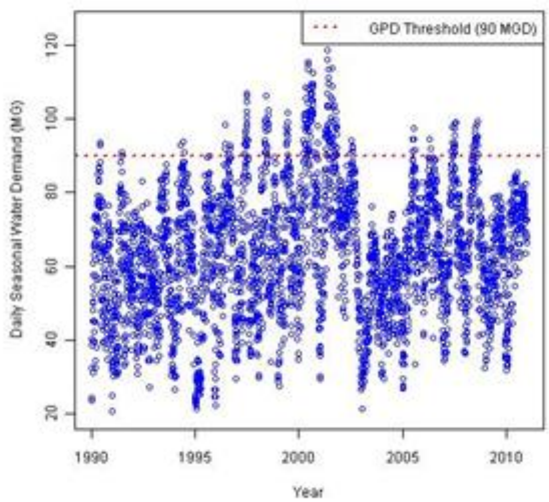


Figure 3.3 – Daily water demand during Jun-Aug of each year. The dotted line is the threshold corresponding to 90 MGD, considered to be high enough that exceeding it would stress the water supply of the utility.

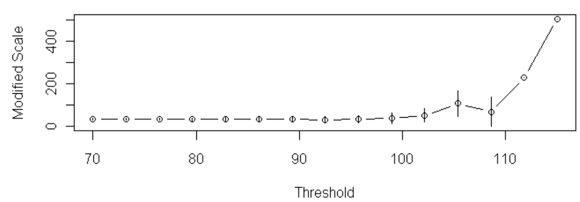


Figure 3.4 – Sensitivity of the scale parameter to choice of threshold. A threshold selection of 90 MGD is in the stable region and an appropriate choice for GPD modeling.

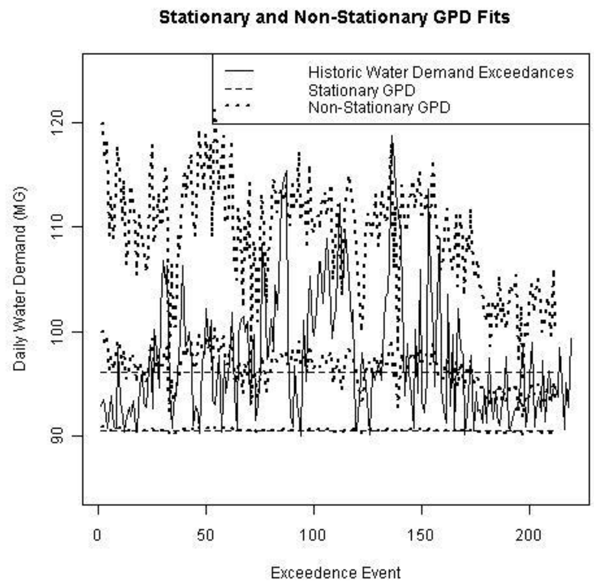


Figure 3.5 – Quantiles from stationary and nonstationary GPD models. Solid line shows the daily water demand, dotted lines are the 0.05, median and 0.95 quantiles from the nonstationary GPD, and dashed lines are the quantiles from stationary GPD model.

exceedance with the appropriate GPD so the diagnostics is performed on the residuals and they are seen to be well modeled (see 10, *Chapter 3*). Furthermore, the 0.05 and 0.95 quantiles from this are shown as dotted lines in Figure 3.5, which shows that the temporal variability of the demand is very well captured relative to the stationary model.

The occurrences of exceedances are modeled as a Poisson Process. The rate parameter, λ, is the average number of occurrences per year over the record of the historical data and here it is estimated to be 10.4 per year or 3.37 per month. The monthly occurrences are shown in Figure 3.6. The average occurrence is shown as the dashed line and the nonstationary occurrences, whose derivation is shown below, are shown at the median and quantile values as the dotted line. To model the month-to-month variability seen in the occurrences we fitted a GLM (Poisson regression) to the number of occurrences each month using Equation (8) with the set of covariates in Table 3.2. The 0.05 and 0.95 quantiles from this model are shown as dotted lines in Figure 3.6 – and it can be seen that the month to month variability of occurrences of exceedances is quite well captured by the nonstationary model.

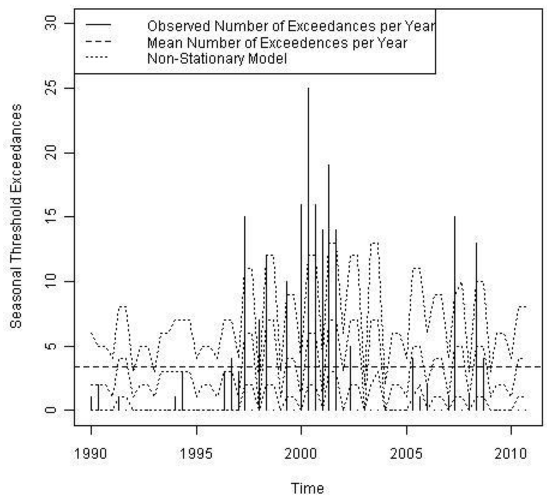


Figure 3.6 – Quantiles from stationary and nonstationary Poisson models. Solid line shows the monthly exceedances, dotted lines are the 0.05, median and 0.95 quantiles from the nonstationary Poisson model, dashed lines are the quantiles from stationary model.

3.6 Discussion

We proposed two extreme value analysis techniques for modeling water demand extremes – one to model maximum demand values over a block of a month (i.e., monthly maximum) using GEV and the other to model exceedances above a threshold using POT. They are complementary and together provide comprehensive insights into the variability of extremes. We found that the variability in extremes is modulated by climate attributes and, to model this nonstationary versions of these methods are proposed. These powerful and emerging methods offer attractive alternative approaches to modeling extremes in the face of variability from climate or other sources. We demonstrated the utility of these methods by applying them to modeling water demand extremes from Aurora Water, a water utility in Aurora, CO.

Traditional statistical methods based on linear models using Normal Distributions capture only the average variability and are incapable of modeling extreme events or events in the tail of the distributions. However, extreme water demand estimates critically are important for a variety of decisions made by utility managers ranging from plant operations, regulatory compliance, plant capacity expansion, operational viability, infrastructure updates, etc. The Extreme Value Analysis methods proposed here can model the high impact extreme events in a more robust manner than traditional, Normal methods. These methods can provide scenarios of water demand extremes at seasonal time scales incorporating seasonal climate forecasts for near term operational management and planning (e.g. *Towler et al.*, 2010) and also at longer time scales incorporating climate change projections (*Towler et al.*, 2011, *Chapter 5*).

We used climate attributes as covariates but other socio-economic variables and their projections, when available, can also be included in the model. Moreover, any variables of known importance to the utility can be tested using the methods described and statistically significant variables should be included in the specific utility's model.

The EVA models can be applied generally across locations under different climate and socio-economic conditions. To test this we applied these approaches to water demand from Tampa Bay, FL. The Tampa Bay, FL utility provides a stark contrast of climate, economy, population growth, and overall size from Aurora, CO. Tampa Bay, FL is dominated by maritime climate with substantially more rain as well as a less pronounced seasonality but with much higher population and a larger service area. The Extreme Value Analysis approach proved capable of modeling the water demand extremes as robustly and skillfully as they did in Aurora, CO and the results are described in (*Chatper 3*, see Appendix A for full results).

4 Projecting Demand Extremes under Climate Change Using Extreme Value Analysis

4.1 Abstract

Accurate projections of water demand extremes are crucial for all aspects of the planning process; defining target growth, securing financing, etc. The role of climate on water demand variability is pronounced and will likely be more so in the future due to climate change. Thus, a flexible approach is needed to project water demand extremes that can incorporate climate change scenarios. An integrated approach was developed with two components: (i) extreme value models for water demand maxima and threshold exceedances that are based on climate attributes are fitted to the historic water demand and climate data and, (ii) ensembles of future weather sequences conditioned on climate change projections generated using stochastic weather generator. Together they provide scenarios of water demand extremes under various climate projections. Aurora, CO was used as a case study to demonstrate the utility of this approach.

4.2 Introduction

Water demand projections help water utility managers make short-term operational decisions such as treatment plant production, balancing supply and demand, and potential water use restrictions in the event of a drought. Every year, water utility managers and financial staff make a projection of how much water there are going to sell (water sales) and, subsequently, project the resultant revenues as part of their annual budgeting process. Additionally, these demand projections help managers make long-term decisions about additional supply acquisition, additional treatment capacity, transmission mains, storage tanks, water conservation impacts, price elasticity and demand management (*Billings et al.,* 2008). Longer-term financial sustainability for the water system is anchored by accurate long-term demand forecasts. Accurate water demand forecasts are needed for both short-term (operational and financial) and long-term (planning and financial). The costs of being wrong can be significant. Inaccuracies in longer-term forecasts can result in large costs to utilities and their customers. For example, costs can manifest in the form of stranded capital assets, insufficient supply reliability, or a reduced level of service due to treatment plant capacity limitations.

All utilities use water demand projections, however, many of them are from simplified models and fail to incorporate nonstationarity and variability from climate change and other uncertainty. A large number of factors are known to impact water demand; including population, employment, technology, weather, climate, price, infrastructure efficiency, conservation programs, socioeconomics and water awareness (*Billings et al.,* 2008). Models have focused mainly on the social and economic variables—which are often difficult to project—resulting in complex, highly uncertain models, many of which are outside the reach of utilities that are small in size or have limited fiscal resources.

Both climate and weather play a fairly significant role in modulating water demand and some research initiatives have used precipitation and temperatures to help model water demand (e.g., *Foster et al.* 1981; *Billings,* 1982; *Nieswiadomy et al.* 1988). Typically, climatic variables are represented as temporal (i.e., seasonal, monthly, annual etc.) averages but this approach may be limiting in that, the effect of precipitation and temperatures may be more pronounced at first, and then decrease with time (*Maidment et al.* 1986). To address this issue, and others, spell statistics (i.e. droughts, heat waves, etc.) have been considered (e.g., *Smoyer-Tomic*, 2003; *Foster et al.,* 1981; *Billings,* 1982) and more recently in Chapter 2 for use in statistical demand models.

Utilities are specifically interested in the high impact, low probability water demand events which drive infrastructure planning decisions and the need for capital to fund such improvements. Accurate projections of peak hour demand and peak day demand are necessary for planning capital improvements such as an alternate supply of source water, treatment plant capacity, transmission line upgrades, storage tanks, and booster pumping stations.

These events, by definition, are 'extreme events,' and one approach that this research found to better model these events is the use of Extreme Value Analysis (EVA). Climate change and more extreme weather events will likely change peak demands and will need to be considered by water utilities in future planning, design, and operations. Extreme Value Analysis has been used in a wide variety of fields including the financial industry (e.g., *Embrechts,* 1999), civil engineering (e.g., *Holmes,* 1999), ecology (e.g., *Eaton,* 1994), water quality (*Towler*, 2010) and especially climatology (e.g., *Beguería,* 2006; *Coles*, 2003; *Katz,* 1995; *Furer et al.*, 2011). EVA has been used in hydrology to estimate and forecast flood frequency, model financial loss related to flooding events and to model extreme hydrological events in water sheds of various sizes (*Fiorentino,* 1984; *McNeil*, 1997; *Swift,* 1989; *Beguería,* 2006; *Coles*, 2003; *Katz,* 1995, *Eaton, 1995; Coles,* 2001; *Smith,* 1989). For a full discussion on EVA and hydrology, see *Katz,* 2002. However, EVA has not seen employment in the water demand sector.

Extreme events exhibit significant variability over time driven by climate fluctuations. For example, an increasing trend in the frequency and intensity of precipitation events in the US has been known for more than 15 years (*Trenberth,* 1998 & 1999). The International Panel on Climate Change (IPCC 2010) projects the same for much of the world in the coming decades due to climate change. Researchers in water resources and hydrology have been working on translating the climate change projections from the global climate models to impacts on hydrology and water resources. This requires downscaling climate information which is often at coarse spatial and temporal scales to scales of interest and driving them through hydrologic models to generate ensembles of streamflows and consequently, impacts to water resources. A body of literature exists on understanding water resources variability under climate change in the Western US (e.g. *Christensen*, 2004; Rajagopalan et al., 2009). Recently, joint efforts from water utilities have been developing tools for water supply and demand with a focus on planning decisions by water utilities (*JFRCCVS*, 2012, *Ray et al*., 2010). The traditional approach of downscaling climate information to process models can be computationally intensive and deter a majority of water utilities from being able to access them. Simple, effective and flexible tools are needed that can translate coarse climate information to specific attributes of decision variables without having to go through computationally intensive process models. *Towler et al.* (2010) developed an extreme value analysis models to translate monthly precipitation and temperature projections under climate change to projections of streamflow extremes and consequently water quality extremes in the Northwest United States and also for Aurora, CO

(*Towler et al.*, 2011). Such models are effective in providing projections of extreme events that are essential to water utility managers for planning purposes.

Demand is somewhat different than water supply and water resources due to the potential human response factoring into demand. But the same weather factors such as temperature and precipitation play a significant role in projecting water demand. This research has expanded this EVA approach to apply to projecting water demand extremes, and the potential impacts from climate change to those extremes.

The paper is organized as follows. A brief description of the study application, data and climate change projections and impacts are first presented. The approach to generating ensembles of water demand extremes under climate change projections is next presented followed by the results and a summary.

4.3 Study Area – Data and Climate Change Impacts

The water utility of Aurora, CO was chosen as the study location, which is a rapidly growing suburb of Denver with limited water supply in a semi-arid region. Recent dry spells (*Pielke, et al.*, 2005) have exacerbated the situation and climate projections portend further stress to the water resources for the system. They have instituted short and long term demand management (*Kenney et al.*, 2008) in addition to embarking on developing a waste water reuse facility to augment their limited mountain water supply (see details at http://www.prairiewaters.org/). Climate change projections from multi-model global

circulation models in conjunction with hydrologic models (*Yates et al*., 2005) show 10-25% reduction in the runoff in the Colorado basin and 5-10% reduction in the Arkansas basin over a 50-year period, both sources provide water to Aurora (*Ray et al*., 2008). Clearly, Aurora, CO (and potentially many other utilities) needs to minimize errors in future water demand projections, and also understand the potential impacts of climate change on these projections. Realistic projections of water demand extremes should be of immense use for all water managers as costly infrastructure decisions are necessarily being made, as the cost of being wrong could be significant.

Weather Statistic Thresholds	
Minimum Temperature	
June	57°F
July	62°F
August	60°F
Maximum Temperature	
June	89°F
July	94°F
August	89°F
Precipitation	
June	0 in
July	0 in
August	0 in

Table 4.1 – Threshold of daily minimum and maximum temperatures and, precipitation, to compute weather attributes for use as covariates in the extreme value analysis models.

Daily water production data in million gallons per day (MGD) was available for the period 1990-2010. In general, water is produced to meet the respective demand and therefore the production data can be considered water demand. The critical season of high demand for

Covariate	Definition	Correlation with Monthly Maximum Demand
Average Monthly Max Temperature	**Daily maximum temperature values averaged over one month.**	0.35
Average Seasonal Minimum Temperature	Daily minimum temperature values averaged over one month.	0.34
Average Seasonal Precipitation	Daily precipitation values averaged over one month.	0.04
TIME	**A vector of values progressing chronologically.**	**-0.05**
Average Monthly Daily High Hot Spell in Days	The average length in days of a spell with daily high temperatures over a given threshold for each month.	0.25
Total Monthly Days with High Temperature Above Threshold	Total days the maximum daily temperature exceeds a given threshold in a month.	0.26
Maximum Monthly Daily High Hot Spell	The longest spell with daily high temperatures over a given threshold in a given month.	0.25
Average Monthly Nightly Low Hot Spell in Day	The average length in days of a spell with nightly low temperatures above a threshold for each month.	0.39
Total Monthly Nights with Low Temperature Above Threshold	**Total days with the nightly low temperature above a given threshold for each month.**	0.48
Maximum Monthly Nightly Low Hot Spell	The longest spell in days with nightly low temperatures above a given threshold for each month.	0.50
Average Monthly Precipitation Spell in Days	The average length in days of a spell with precipitation above a given threshold for each month.	-0.06
Total Monthly Days with Precipitation Below Threshold	**Total days with the precipitation above a given threshold each month.**	-0.22
Maximum Monthly Precipitation Spell	The longest spell in days with precipitation above a given threshold each month.	-0.05

Table 4.2 – Weather attributes and their correlation with monthly maximum water demand. Bolded values are statistically significant at 95% confidence level and, bolded rows are variables used in the nonstationary GEV as covariates.

Aurora is summer, specifically the three month period of June-Aug. Monthly maximum demand was computed from the daily data for use in projections. Daily weather data consisting of precipitation, maximum and minimum temperature were obtained from the National Climatic Data Center (NCDC, http://ncdc.noaa.gov/). A suite of weather attributes (hot/dry, wet/cold spells along with average weather variables) that have been identified to be related to average water demand and water demand extremes (*Chapters 2, 3*) were computed from the daily weather data (thresholds used for spells are shown in Table 4.1) for the summer months in the data period (the attributes used are shown in Table 1 and 2).

Water demand variability is often more dependent on weather phenomena like droughts and extreme temperatures or weather spells than traditionally employed socioeconomic variables (*Arbués,* 2003). For this study, weather spell statistics (Table 4.2) were used as covariates in this analysis, in addition to the daily weather information ascertained from the NCDC. We computed hot/cold and dry/wet spells with the thresholds shown in Table 4.1. In Chapter 2 we motivated the use of weather spells for modeling demand attributes. A suite of weather attributes were developed and their correlations with monthly maximum water demand are shown in Table 4.2. For computational ease, variables with significant correlations were selected for use in modeling water demand extremes with nonstationarity. Moreover, the spell statistics can also be calculated for projected water demand data. For more background on creating weather spell statistics, see Chapter 2.

Projections of precipitation and temperature for the region under climate change scenarios were solicited from the Joint Front Range Climate Change Vulnerability Study completed in 2012 (*JFRCCVS*, 2012). We used two projections from the A2 middle emissions scenario; increasing temperatures and decreasing precipitation (warm/dry scenario) and increasing temperatures and increasing precipitation (warm/wet scenario). The details of these scenarios can be seen in Table 4.3. Change in the annual average temperature over the 30-year

Scenario	GCM/Ensemble	30 Year Precipitation Change (%)	30 Year Average Temperature Change (°F)
Warm/Wet	ncar_ccsm3_0.2	3.77	3.40
Warm/Dry	Micro3_2_medres.1	-8.51	3.40

Table 4.3 – Two climate change scenarios used in stochastic weather generation.

planning horizon (2010-2040) was used along with two different forecasted precipitation scenarios; wet and dry. The temperature/precipitation scenarios will henceforth be referred to as warm/dry and warm/wet. These projected changes are used in a stochastic weather generator, which is described later, to generate ensembles of daily weather sequences.

4.4 Approach for Projecting Water Demand Extremes

As described above, two attributes of water demand extremes are of interest – monthly maximum and the number of water demand exceedance days above a threshold. We propose an approach to projecting these extremes attributes based on future climate projections. The proposed approach has two main components – (i) extreme value analysis models for the two water demand extreme dependent variables based on weather attributes and (ii) a model for generating daily weather sequences and consequently the weather attributes, which are then combined with the extreme value analysis models to generate projections of water demand extremes. The components of this approach are described below.

4.4.1 Extreme Value Analysis Models for Water Demand Extremes

In Chapter 3 we describe in detail the motivation for extreme value analysis models along with their development for water demand extremes data for the same study location as in this research. However, for the benefit of the readers we provide a brief description of the two models and refer to the above paper and other references for details.

Monthly maximum demand, which is the maximum of a block of a seasonal (June-August) month, is modeled using Generalized Extreme Value (GEV) distribution. This distribution results from theory as the limiting distribution of the maximum of a series of independent and identically distributed observations (*Leadbetter,* 1983; *Coles*, 2001; *Katz and Naveau*, 2002). The cumulative distribution function of the GEV is given as:

$$G(z;\theta) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right]\right\} \tag{4.1}$$

where $\theta = [\mu, \sigma, \xi]$ are the location, scale and shape parameters, respectively, and z is the dependent variable, monthly maximum demand in this case. The location parameter, μ, indicates the center of the distribution or the mean, the scale parameter, σ, represents the spread of the distribution or the variance, and the shape parameter, ξ, represents the behavior of the distribution in the tail. The GEV better models heavy-tailed distributions compared to traditional probability distributions and therefore preferable for modeling extremes (*Katz and Naveau*, 2002).

Extremes vary with time and also are modulated by external variables - i.e. nonstationarity. For example, water demand extremes have been shown to be related to weather attributes (Table 4.2) (see *Chapter 2*) and streamflow extremes have been shown to be modulated by large scale climate and meteorological variables (*Sankarasubramanian and Loll*, 2003; *Towler* et al., 2010). Thus to model the nonstationarity the parameters of the GEV have to be related to the external variables or covariates. To this end, the GEV distribution parameters can be modeled as a function of covariates as follows (*Katz and Naveau*, 2002):

$$\mu(x) = \beta_{0,\mu} + \beta_{1,\mu}x_1 + \cdots + \beta_{n,\mu}x_n \tag{4.2}$$

$$\log[\sigma(x)] = \beta_{0,\sigma} + \beta_{1,\sigma}x_1 + \cdots + \beta_{n,\sigma}x_n \tag{4.3}$$

$$\xi(t) = \xi \tag{4.4}$$

where the variables $x_1, x_2, \dots, x_n$ are the covariates which can include 'Time' to consider a temporal trend, and the betas are the coefficients of the regression. Shape parameter is generally not varied as it tends to be noisy, but it too can be varied as the other two. Notice that the above form is reminiscent of a Generalized Linear Model (*McCullagh et al.*, 1989). The coefficients are estimated using the maximum likelihood approach that is commonly used in GLM (*Katz and Naveau*, 2002). Best combination of covariates, if there are several of them, is selected using a likelihood ratio test. In this, models are fit with various combinations of the covariates and are then compared in pairs based on their values of the likelihood function (*Katz and Naveau*, 2002; *Coles*, 2001), which selects a skillful yet parsimonious model.

Variations of the nonstationary GEV idea have been developed for flood frequency modeling (*Sankarasubramanian and Loll*, 2003; *Rajagopalan et al.*, ,2010) and recently Towler et al., 2010 applied this for modeling maximum streamflow and consequently water quality extremes – for both seasonal time scale and multidecadal time scales under climate change.

The two forms of GEV modeling provide a strong capability to characterize extremes at a variety of time scales that will be of use in water supply planning.

4.4.2 Threshold Exceedance Model of Water Demand

Block extremes, as described in the previous section, are based on selecting a single value within a block, thus discarding large amounts of data, which can be undesirable especially when the data set is limited. A complimentary approach is to model threshold exceedances using the Points Over a Threshold (POT) method. In this, a threshold is selected, often based on the system application, and the probability of exceedance is modeled as a Poisson Process (PP) with parameter λ while the magnitude of exceedance is modeled using Generalized Parateo Distribution (GPD) (see *Coles*, 2001; *Katz and Naveau*, 2002). The GPD is given as:

$$F(x;\sigma,\xi) = 1 - \left[1 + \xi\left(\frac{x}{\sigma}\right)\right]^{-\frac{1}{\xi}},$$

where, σ is the scale parameter and ξ is the shape parameter.

To model nonstationarity, similar to the GEV described above, the dependence of the parameters to covariates is introduced as:

$$\log[\sigma(x)] = \beta_{0,\sigma} + \beta_{1,\sigma}x_1 + \cdots + \beta_{n,\sigma}x_n \tag{4.5}$$

$$\xi(tx) = \xi \tag{4.6}$$

$$\log(\lambda) = \beta_{0,\lambda} + \beta_{1,\lambda}x_1 + \cdots + \beta_{n,\lambda} \tag{4.7}$$

where the variables $x_1, x_2, \ldots, x_n$ are the covariates which can include 'Time' as one of them to consider a temporal trend, and betas are the coefficients. Shape parameter is generally not varied as it tends to be noisy, but it too can be varied as the other two if desired. As with the case of GEV the parameters are estimated via maximum likelihood estimation and the best combination of covariates are selected using likelihood ratio test (*Katz and Naveau*, 2002; *Coles*, 2001). The Point Process approach has also been suggested as a way to model the threshold exceedance process in a single model (*Smith,* 1989, *Furrer et al.,* 2010).

4.4.3 Generating Daily Weather Sequences under Climate Change

The covariates in the extreme value models are weather attributes that are based on daily weather information as mentioned above (see Tables 4.1, 4.2) thus, daily weather sequences are necessary. There is a rich literature on stochastic weather generation – traditional parametric (e.g., *Richardson*, 1984; *Parlange and Katz*, 2000; *Furrer and Katz*, 2007) and recent nonparametric techniques (e.g. *Rajagopalan and Loll*, 1996; *Yates et al.*, 2003; *Apipattanavis*, 2007). Nonparametric weather generator based on time series bootstrapping (*Rajagopalan and Loll*, 1996; *Yates et al*., 2003) are simple and robust to implement. Furthermore, they can easily be modified to generate weather sequences conditioned on climate forecasts or projections (*Apipattanavis, et al*., 2009). Here we propose to us a block bootstrapping approach (*Efron and Tibshirani,* 1993); in this approach blocks of historical years are selected at random and with it the entire daily weather sequences. Sampling blocks of historical data and combining them to generate a rich variety of weather sequences is akin to generating from the multivariate probability density function (*Efron and Tibshirani*, 1993). This captures the variability at low frequency, besides being easy to implement. In this application, we chose 30-year blocks. Other sampling approaches based on K-nearest neighbor bootstrap on a daily time scale can be employed (*Apipattanavis et al.*, 2009), but we chose the block bootstrap for sake of parsimony. Precipitation and temperature trends for the simulation horizon are obtained from climate change projections (Table 4.3) and these trends are imposed on to the daily weather sequences from the block bootstrap to obtain sequences conditioned on climate change projections. This approach is widely used in hydrologic applications, recently Rajagopalan et al., 2009 used this to simulate streamflow in the Colorado River Basin to investigate water supply risk under climate change. By executing bootstrap resampling with an imposed trend, water demand projections can be used to inform decisions related to water resource needs.

4.5 Results

A planning horizon of 30-year period was considered and we chose two plausible scenarios suggested from the climate change report for the study region (*Ray et al.*, 2010). The scenarios are warm/wet and warm/dry, and Table 4.3 shows the associated precipitation and temperature trends. Daily weather sequences were generated based on the methods described in the previous section and from it the weather attributes. Nonstationary models for maximum monthly demand and threshold exceedances were fitted to the historical data – they are same as those developed in Chapter 3 and are shown in Table 4.4.

4.5.1 Monthly Maximum Water Demand Projections

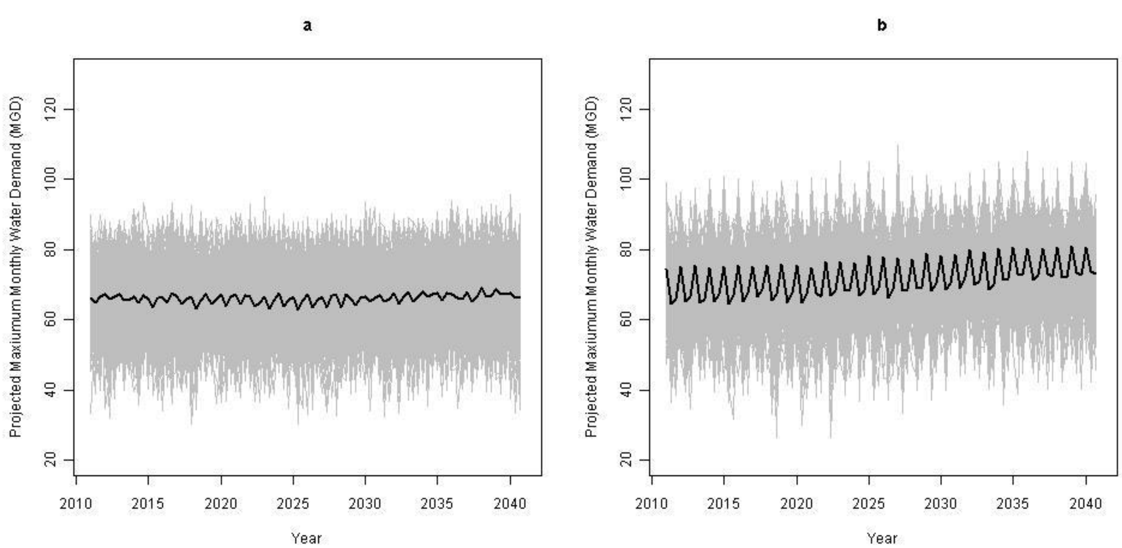
We generated 100 weather ensembles each 30-years long for the two climate change projection scenarios and also without the climate change trends – i.e. 'natural variability.' Table 4.2 shows the weather attributes from these ensembles that were used in the stationary GEV model to generate projections of monthly maximum water demand. Figure 4.1 (a,b,c) shows

Variable	Non-Stationary Model: $\mu = \beta_0+\beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)$
β_0	91.08
β_1	0.60
β_2	-0.97
β_3	-0.29
β_4	0.33
Location Parameter (μ)	(nonstationary)
Scale Parameter (σ)	9.98
Shape Parameter (ξ)	-0.15
Negative Log Likelihood	238.82

**Where (a) is the longest monthly spell with minimum temperatures above a threshold, (b) is the longest monthly spell with precipitation above the threshold, (c) is the monthly average maximum temperature and (d) is time.

Table 4.4 – Generalized Extreme Value distribution model.

the water demand projections for the natural variability warm/wet and warm/dry scenarios, respectively. Clearly, water demand trends upward for Aurora under climate change projections. Even under a warm/wet projection where rainfall is projected to increase, the



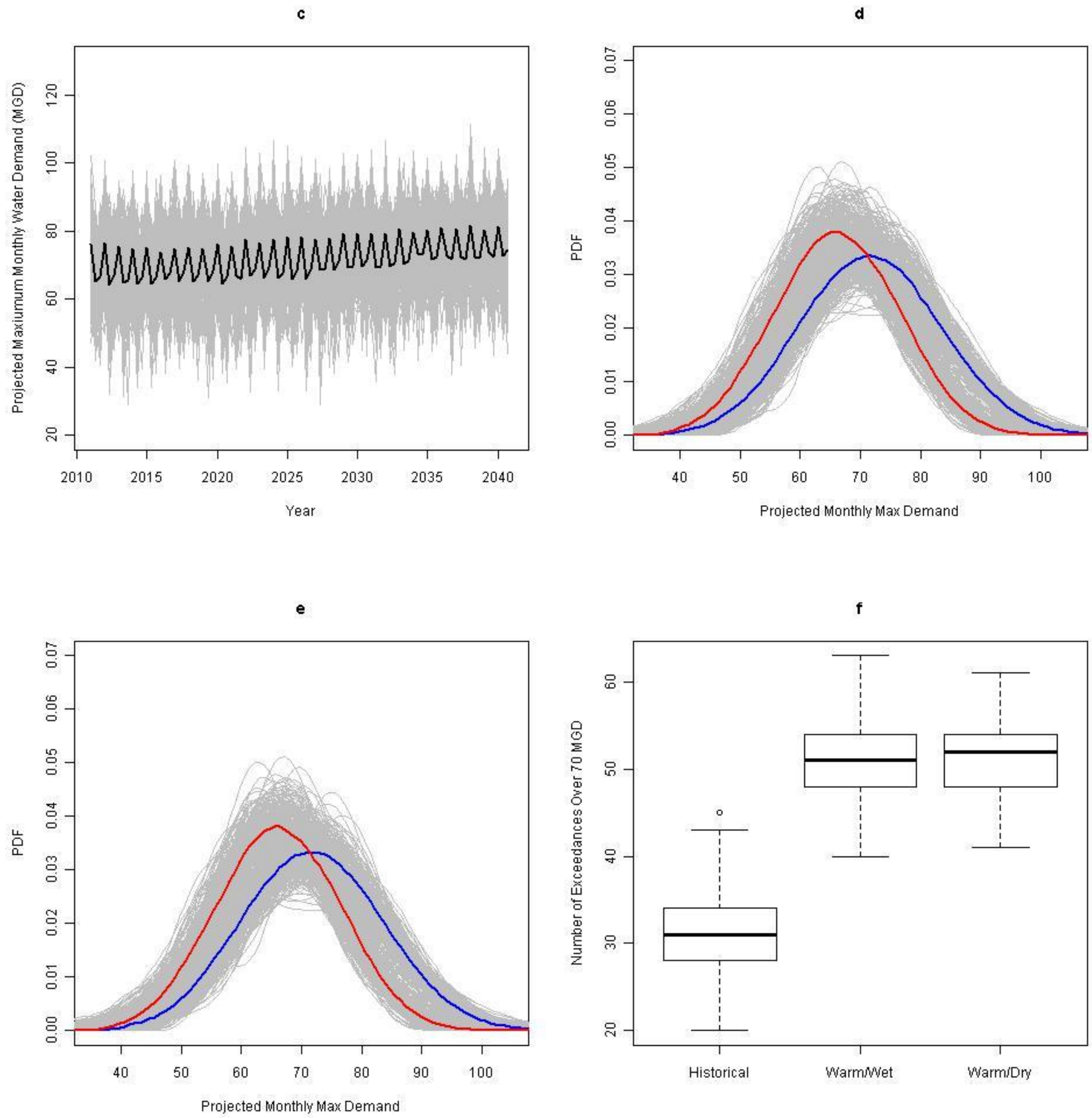


Figure 4.1 - Projections of monthly maximum water demand from nonstationary GEV model (a) based on natural climate variability (b) based on warm/wet future climate projection, (c) based on warm/dry future climate. (d) is the Probability density functions of monthly maximum water demand projections for warm/wet future climate (grey lines are from individual simulation and blue is the average) for natural climate variability (red), (e) is the same as (d) but for warm/dry climate projections and (f) is Boxplots of the number of monthly maximum demand exceedances of 85 MGD for the three scenarios.

increasing trend in temperature appears to negate the precipitation effect and increase the water demand. Increasing temperature leads to increased evaporative losses and thus creates demand for more water (*Arbués,* 2003). The probability density function (PDF) of the monthly maximum demand from the warm/dry scenario is shifted to the right relative to the PDF from natural variability (Figure 4.1d). This shift translates into increased risk of higher demands, which tends to stress the system and also impact infrastructure planning decisions. Consistently, the number of exceedances of 85 MGD for the three climate projections is shown in Figure 4.1f and here, too, the number of exceedances increases substantially under climate change relative to natural variability.

4.5.2 Water Demand Threshold Exceedance Projections

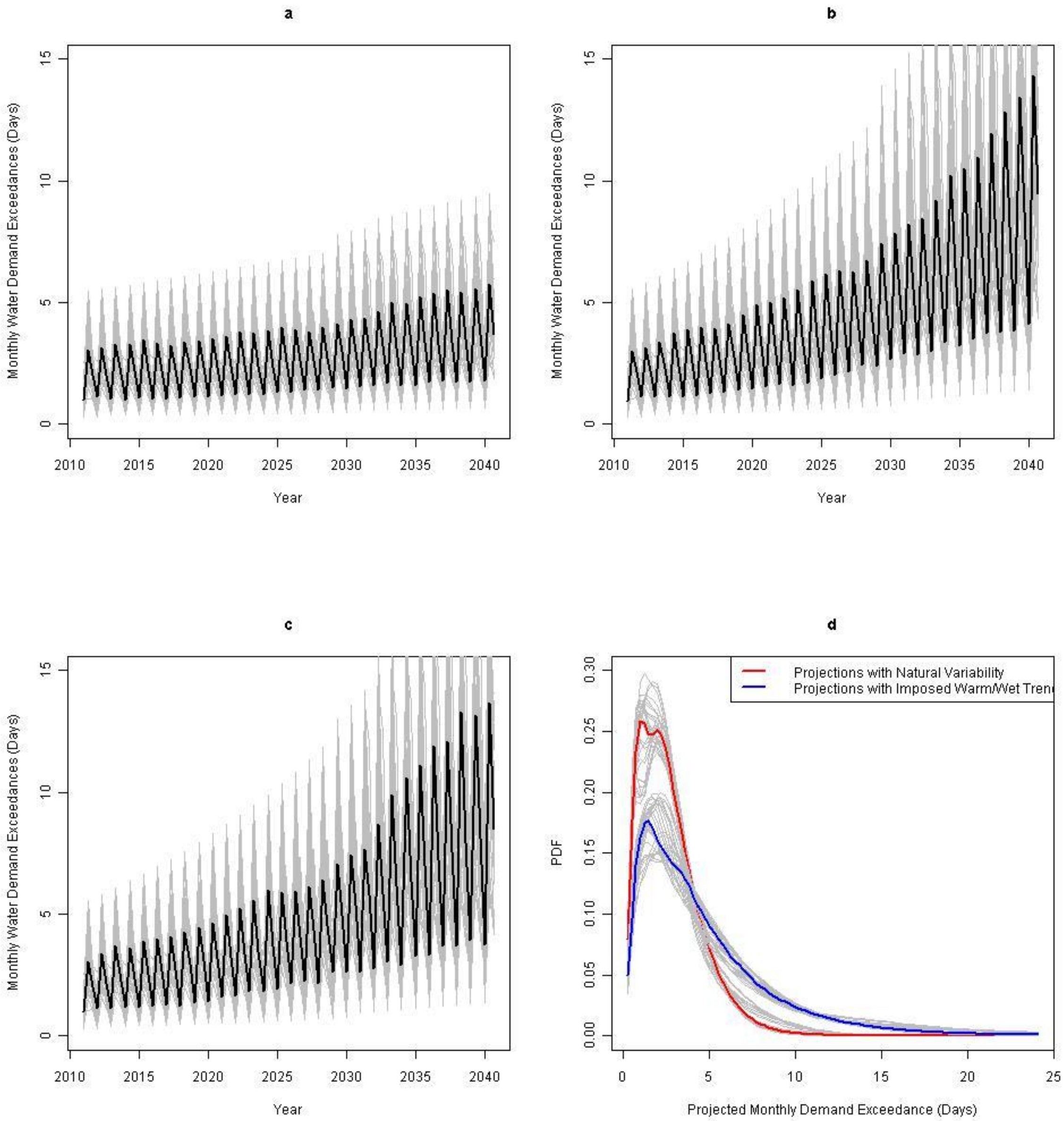
Here, the parameter of the Poisson Process is estimated based on the covariates and consequently the number of exceedances generated. The threshold for exceedance is selected

Variable	Nonstationary GPD Model *** $\sigma = \beta_0+\beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)$	Nonstationary Poisson Model** $\lambda = \beta0+\beta1(a2) + \beta2(b2) + \beta3(c2) + \beta4(d2)$
β_0	7.43	-13.119
β1	-0.28	0.111
β2	0.59	0.071
β3	-1.24	3.899
β4	-0.01	0.005
Scale Parameter (σ)	(nonstationary)	-
Shape Parameter (ξ)	-0.34	-
Negative Log Likelihood	623.45	-

*** (a) = Daily Maximum Temperature, (b) = Daily Minimum Temperature, (c) = Daily Precipitation, (d) = Time

**Where (a2) is the longest monthly spell with minimum temperatures above a threshold, (b2) is the longest monthly spell with precipitation above the threshold, (c2) is the monthly average maximum temperature and (d2) is the preceding month average water demand.

Table 4.5 – Generalized Parateo Distribution and Poisson models used for projections.


a
b
c
d
Monthly Water Demand Exceedances (Days)
Year
PDF
Projected Monthly Demand Exceedance (Days)
Projections with Natural Variability
Projections with Imposed Warm/Wet Trend

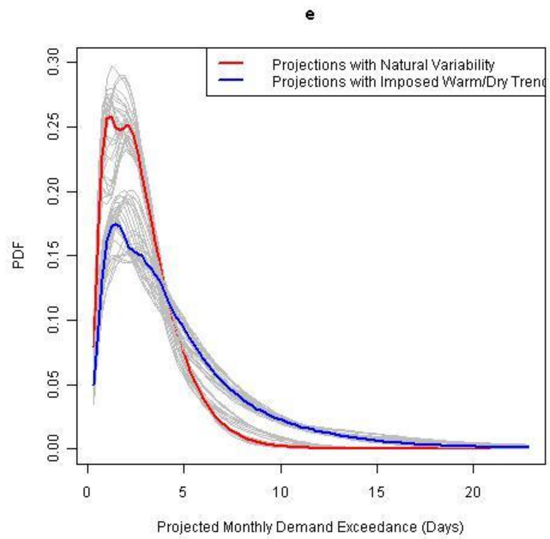


Figure 4.2 - (a) Projections of number of monthly exceedances of daily water demand over 90 MGD from the Poisson model (a) for natural variability, (b) for warm/wet climate projection, (c) for warm/dry climate projection. (d) Probability density functions of number of monthly exceedance for natural variability and warm/wet climate projection (dashed), its average in blue and the average from natural variability in red and (e) same as (d) but for warm/dry climate projection.

as 90MGD, then the magnitude of exceedances are generated from the GPD model. Figure 4.2(a,b,c) show the generated number of exceedances for the three projection scenarios – natural variability, warm/wet and warm/dry, respectively. As expected, the exceedances increase over time for the warm/wet and warm/dry case relative to natural variability. The PDF of exceedances from warm/wet and warm/dry scenarios (Figure 4.2d,e) is shifted to the right relative that from natural variability, indicating higher probability of increased exceedances. As mentioned above, the magnitude of exceedances is generated from the GPD model while the probability of experiencing an exceedance is generated from the Poisson Process model. Figure 4.3a shows the PDF of generated magnitude for the warm/wet climate projections and Figure

4.3b shows the same for warm/dry scenario. It can be seen that in both cases the PDF from climate change projections are shifted to the right relative to natural variability, indicative of higher exceedances. The impact of climate change is more clearly experienced on the frequency of

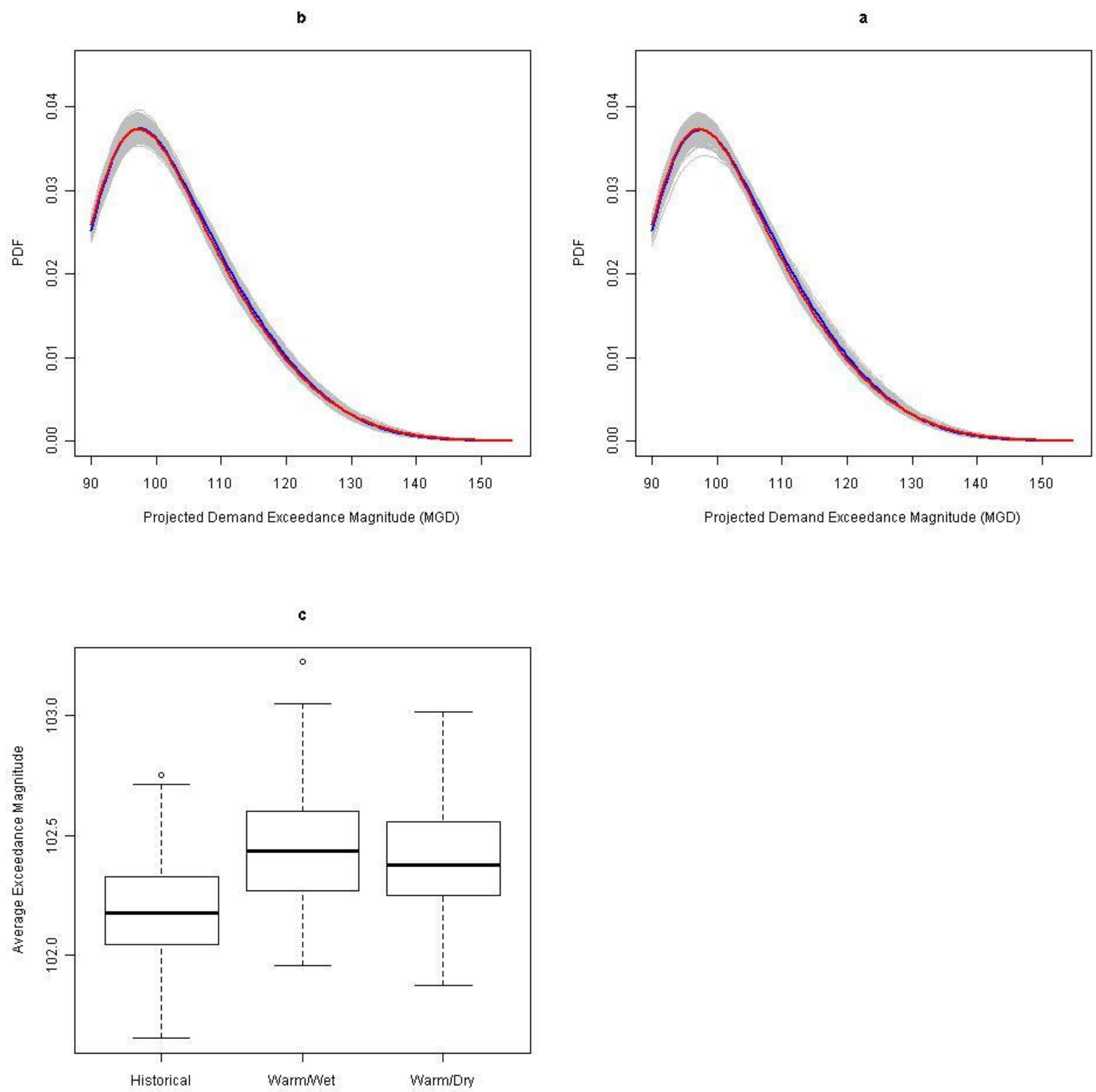


Figure 4.3 - (a) Probability density functions of the magnitude of exceedances for warm/wet climate projection (in gray), the average PDF is shown in blue and the average from natural variability in red, (b) same as (a) but for

warm/dry climate projection and (c) boxplots of average exceedances for natural variability, warm/wet and warm/dry climate projections.

exceedances than the magnitude of exceedances. This is clearly seen in Figure 4.3c which shows the boxplots of average magnitude of exceedances for the three climate scenarios and for the warm/wet and warm/dry case the average exceedances and variability is around 1MGD higher than that of natural variability.

4.6 Summary and Discussion

We proposed an integrated approach to projection of water demand extremes under climate change scenarios. The approach has two main components – (i) extreme value models for water demand maxima and threshold exceedances that are based on climate attributes are fitted to historic water demand data and weather attributes and, (ii) ensembles of future weather sequences and, consequently, weather attributes are generated conditioned on climate change projections using bootstrapping technique, which is combined with the extreme value models to generate ensembles of water demand extremes. We demonstrated this approach on water demand extremes data from water utility in Aurora, CO. Two climate change scenarios determined from Emissions path A2 (the middle path, IPCC, 2010) for this region – warm/wet and warm/dry (*Ray et al*., 2010) were used to conditionally generate daily weather sequences. Monthly maximum water demand and demand exceedances above 90MGD were modeled using extreme value analysis. The results for Aurora, CO suggest that both the frequency and intensity of extreme water demand events is likely to increase and, moreover, that the probability of experiencing more extreme events and even greater extreme events than have been previously experienced is increasing under climate change.

This methodology for projecting water demand extremes under climate change should be of immense use to water utility planners and water managers. The proposed approach provides a framework to translate uncertainties in climate projections to uncertainties in water demand extremes. Such a robust quantification enables better risk-based decision making of operational and infrastructure planning.

Furthermore, additional social and economic variables and their projections can be incorporated in the modeling framework, thus providing flexibility to integrate any other knowledge of the system. Sophisticated methods for stochastic weather generation (*Yates et al*., 2003; *Apipattanavis et al.,* 2010) can be employed to generate a rich variety of daily sequences and add further nuance to the approach presented here. Other simulation techniques for threshold exceedances such as those developed by Furrer et al., (2010) for heat wave spells can also be employed to add skill to these modeling techniques.

5 Conclusion

The current state of water demand modeling is very limited; only a cursory use of climate or other covariates are used and water demand models are either oversimplified linear regression models (as is the case with small utilities) or vastly complex models created by consultants who only deliver a single number (as is the case with large utilities). Either way, current water demand models lack the ability to incorporate climate information and do not carry a capacity to understand the risks and uncertainties associated with climate change.

The impacts of climate change on water supplies are extensively studied. However, water demand drives the need for supply and supply projections are not meaningful without a comprehensive understanding of water demand. The research presented here seeks to create a methodology and a framework for understanding the risk of climate change to residential water demand. The framework and modeling approach detailed gives an understanding of the uncertainty associated with climate change and how that uncertainty translates into uncertainty in water demand. An effort was made to show that the probability of seeing more extreme water demand events and greater water demand events than have been historically experienced. With that knowledge, one can be better informed to make decisions about utility operations, infrastructure updating, personnel management and additional supply acquisition.

The modeling results are useful for planning purposes and were created in an effort to add to the existent tools available to water utility managers. The modeling approach detailed can be modified, added to or subtracted from based on the particular needs, data and capacity of a specific utility. The real power of the outlined approach is the ability to modify according to specific situation at each individual utility. The results presented were of a single case study (with another case study in Appendix A) and illustrate the skill of the modeling approach. Discussions of the results from each chapter are subsequently presented.

5.1 Generalized Linear Modeling

It was shown how climate covariates can be easily incorporated into water demand projections using readily accessible tools and relatively simple methods using the Generalized Linear Modeling (GLM) framework. Using the GLM and climate attributes to model water demand provides multiple advantages to traditional methods. First, climate data is easily ascertained and from the data climate attributes can be simply calculated as opposed to socioeconomic variables, which can be difficult to define and the data hard to collect. Secondly, the inherent flexibility of Generalized Linear Modeling lends itself well to water demand modeling as the water demand variable of interest may be nonnormal or discrete. The GLM modeling results presented suggest that using climate covariates to describe monthly average and monthly maximum water demand is more parsimonious than using them to model monthly exceedance days. This shortcoming is fundamental to modeling monthly exceedance days, which is a particularly noisy water demand variable. Approaches to circumventing this issue are subsequently discussed.

The developed models can be applied for short term (seasonal) management by incorporating seasonal climate forecasts – such applications are common in water resources management (e.g., *Regonda et al.,* 2011) and recently for water quality management (*Towler et al.,* 2010).

While the demand attributes considered in this research are relevant for utility managers, a more important variable is the variability of 'extreme demand events'. In this, the managers would like to know the fluctuations in peak demand, threshold exceedance and length of exceedances. While individual components can be modeled using GLM as described above, extreme value analysis offers an attractive alternative to model these in an integrated manner incorporating predictor variables (e.g., *Towler et al.,* 2010; *Katz*, 2010; *Katz and Naveau,* 2002). Future research will broach the topic of water demand extremes modeling.

Water demand projections are important for efficient management and planning of water supply. Traditionally these projections are made using statistical and econometric models that account for socio-economic growth and limited climatic variability. Given the sophistication and uncertainty in economic projections these models are accessible only to relatively well-off utilities. We see two major gaps – (i) need for an approach to model a variety of demand attributes that are needed by utility managers such as – average demand, monthly peak and number of demand exceedances above a threshold and (ii) comprehensive use of climate information. We first identify a suite of climate variables that include high/low temperature and wet/dry spells and the more traditional average temperature and precipitation. We find that these variables capture nuanced connections between climate and water demand, especially the lagged relationships. Second we propose a GLM framework to model attributes of water demand by incorporating the climate variables. We demonstrate the utility of this approach to demand modeling for a water utility in Aurora, CO. The various dependent variables (monthly average demand, monthly peak demand and monthly number of demand exceedances above a threshold) are modeled quite well and the models have the ability to capture a variety of water demand attributes and distributions. Additional socio-economic predictors can be easily incorporated, although here we use 'time' as a proxy for this. Furthermore, the GLM based modeling approach with climate variables provides the capability to make future water demand projections under climate change. In this, the climate attributes can be derived under climate change scenarios which can then be used with the developed GLM models to make demand projections. Above all, this method is simple and can run with limited computational resources; thus, accessible to all utilities.

5.2 Extreme Value Analysis for Modeling Water Demand Extremes

We proposed two extreme value analysis techniques for modeling water demand extremes – one to model maximum demand values over a block of a month (i.e., monthly maximum) using GEV and the other to model exceedances above a threshold using POT. They are complementary and together provide comprehensive insights into the variability of extremes. We found that the extremes are modulated by climate attributes resulting in substantial temporal variability, to model this nonstationary versions of these methods are proposed. These powerful and emerging methods offer an elegant approach to modeling extremes in the face of variability from climate or other sources. We demonstrated the utility of these methods by applying them to water demand extremes from Aurora Water, a water utility in Aurora, CO.

The use of extreme value analysis to model water demand is motivated by the reality of the extreme nature of water demand events. That is, traditional modeling techniques—regression analysis or even Generalized Linear Modeling (see *Chapter 2*)—result in a thin tail and an underestimation of both the magnitude and the frequency of extreme water demand events. For this research, Extreme Value Analysis was used to specifically model the high impact or extreme events, resulting in substantially more robust estimation of the critical tail of the water demand distribution.

The utility of the Extreme Value Analysis approach extends beyond its ability to better model the tail of the water demand distribution; it is also particularly useful for water utility managers because it better addresses the questions that are pertinent to the planning process. For example, by using Extreme Value Analysis and the Points Over a Threshold approach, information about how often a critically important threshold may be exceeded can be easily ascertained. Similarly, information about how the frequency and magnitude of extreme water demand events is best determined from Extreme Value Analysis and the Block Maxima approach, which models specifically the water demand extremes.

Predictions of weather attributes can be translated to predictions of water demand extremes using these models. Furthermore, projections of weather attributes under climate change will enable the projection of water demand extremes directly – similar to the approach of Towler et al. (2011) – which is a powerful tool.

We used weather attributes as covariates but others such as socio-economic variables can also be include in the mix, enabling their integration with other modeling approaches.

5.3 Projecting Water Demand Extremes under Climate Change

We proposed an integrated approach to projection of water demand extremes under climate change scenarios. The approach has two main components – (i) extreme value models for water demand and threshold exceedances that are based on climate attributes are fitted to the historic data of water demand and weather attributes and, (ii) ensembles of future weather sequences and consequently, weather attributes are generated conditioned on climate change projections using bootstrapping technique, which is combined with the extreme value models above to generate ensembles of water demand extremes. We demonstrated this approach on water demand extremes data from water utility in Aurora, CO. Two climate change scenarios determined from Emissions path A2 (the middle path, IPCC, 2010) for this region – warm/wet and warm/dry (*Ray et al*., 2010) were used to conditionally generate daily weather sequences. Monthly maximum water demand and demand exceedances above 90MGD were modeled using extreme value analysis.

This methodology for projecting water demand extremes under climate change should be of immense use to water utility planners and water managers. Current approaches for water demand forecasting are full of uncertainties inherent in these forecasts, and climate change will likely add another layer of uncertainty to those forecasts. EVA is one approach to better understand the additional uncertainty, and then appropriately plan for future investments in new sources of supply, treatment plant capacity, transmission mains, storage tanks, and booster pumping stations. The potential costs for being wrong with future demand forecasts can be substantial.

Furthermore, additional covariates such as social and economic variables can be incorporated in the modeling framework, thus providing flexibility. Sophisticated methods for weather generation (*Yates et al*., 2003; *Apipattanavis et al.,* 2010) can be employed to generate more variety on daily time scales. Other simulation techniques for threshold exceedances such as those developed by Furrer et al., (2010) for heat wave spells can also improve the simulations. Our aim here was to develop an integrated approach that can is flexible to future enhancements.

5.4 Future Work

Climate change science is still in its infancy and the quality of models, projections and information from climate models is getting better with time. For this reason among others, there is a substantial amount of uncertainty that exists in climate change projections and, therefore, water demand projections under climate change. As climate modeling techniques improve it may be possible to used downscaled Global Climate Model (GCM) data at the specific utility location. Furthermore, more nuanced weather generation techniques can be employed to further bolster the weather projections. Extreme Value Analysis is a relatively new statistical tool to many fields, and this marks its first appearance in water demand modeling.

The tool is tremendously useful for modeling water demand extremes and the concepts introduced in this work should be built upon in the field of water demand modeling.

6 Bibliography

Agthe, D., Billings, B., Dobra, J., and Rafiee, K. 1986. *A simultaneous equation demand model for block rates.* Water Resources Research. Vol. 22, No. 1, pp. 1-4.

Akaike, H. 1974. *A new look at the statistical model identification.* IEEE Transactions Automation Control. Vol. 19.

Al-Quanibet, S. and Johnston, P. 1985. *Municipal demand for water in Kuwait: methodological issues and empirical results.* Water Resources Research. Vol. 2, No. 4, pp. 433-438.

Apipattanavis, S., Podesta, G., Rajagopalan, B. and Katz, R. 2007. *A semiparametric multivariate and multiscale weather generator*. Water Resources Research. Vol. 43.

Araujo, M., Pearson, R., Thuillers, W., and Erhard, M. 2005. *Validation of specie-climate impact under climate change.* Global Change Biology. Vol. 11, pp. 1504-1523.

Arbués, F., Garcia-Valinas, M., Martínez-Espiñeira, R. 2003. *Estimation of residential water demand: a state-of-the-art review.* Journal of Socio-Economics. Vol. 32, pp. 81-102.

Azzalini, A., Bowman, A., and Hardle, W. 1989. *On the use of nonparametric regression for model checking.* Biometrika. No. 76, pp. 1-11.

Babel, M., Gupta, A., and Pradhan, P. 2007. *A multivariate econometric approach for domestic water demand modeling: An application to Kathmandu, Nepal.* Water Resources Management. Vol. 21, pp. 573-589.

Beguería, S. and Vicente-Serrano, S. 2006. *Mapping the Hazard of Extreme Rainfall by Peaks over Threshold Extreme Value Analysis and Spatial Regression Techniques*. Journal of Applied Meteorology and Climatology. Vol. 45, pp. 108-124.

Billings, B. and Jones, C. 2008. *Forecasting Urban Water Demand*. American Water Works Association. Denver, CO.

Billings, B. 1982. *Specification of block rate price variables in demand models.* Land Economics. Vol. 58, No. 3, pp. 386-393.

Chen, M., Marquis, K., Averyt, M., Tignor M., and Miller H. 2007. *Climate change 2007: the physical science basis.* Cambridge University Press.

Christensen, N., Wood, A., Voisin, N., Letternmaier, P and Palmer, R. 2004. *The effects of Climate Change on Hydrology and Water Resources of the Colorado River Basin.* Earth and Environmental Science. Vol. 62, No. 1-3, pp. 337-363.

Coles, S. 2001. *An Introduction to Statistical Modeling of Extreme Values.* Springer Series of Statistics. Vol. 208.

Coles, S. and Simiu, E. 2003. *Estimating Uncertainty in the Extreme Value Analysis of Data Generated by a Hurricane Simulation Model.* Journal of Engineering Mechanics. Vol. 129, pp. 1288-1294.

Danielson, L. 1979. *An analysis of residential demand for water using micro time-series data.* Water Resources Research. Vol. 15, No. 4, pp. 763-767.

Eaton, J., McCormick, J., Stefan, H. and Hondzo, M. 1995. *Extreme value analysis of a fish/temperature field database.* Ecological Engineering. Vol. 4, No. 4, pp 289-305.

Efron, B. and Tibshirani, R. 1993. *An Introduction to the Bootstrap*. Chapman and Hall/CRC.

Embrechts, P., Resnick, S., and Samorodnitsky, G. 1999. *Extreme Value Theory as a Risk Management Tool.* North American Actuarial Journal. Vol. 3, No. 2.

Fiorentino, M. and Versace, P. 1984. *Two-Component Extreme Value Distribution for Flood Frequency Analysis.* Water Resources Research. Vol. 20, No. 7, pp. 847-856.

Foster, H. and Beattie, B. 1981. *On the Specification of Price in Studies of Consumer Demand under Block Price Scheduling.* Land Economics. Vol. 57, No. 4.

Furrer, E., Katz, R., Walter, M. and Furrer, R. 2010. *Statistical modeling of hot spells and heat waves.* Climate Research. Vol. 43, pp.191-205.

Gilleland, E.; Katz, R. 2005. *Extremes Toolkit (extRemes): Weather and Climate Applications of Extreme Value Statistics.* National Science Foundation.

Gumbel, E. 1941. The return period of flood flows. *Annals of Mathematics and Statistics.* Vol. 12, pp. 163-190.

Haagenson, E. 2012. *Statistical Modeling of Water Demand under Climate Change.* MS Thesis, University of Colorado, Boulder, CO.

Hanke, S., and de Maré, L. 1982. *Residential water demand: a pooled time-series cross-section study of Malmö, Sweden.* Water Resources Bulletin. Vol. 18, No. 4, pp. 621-625.

Harrell, F. 2001. *Regression Modeling Strategies – with Applications to Linear Models, Logistic Regression and Survival Analysis.* Springer-Verlag, New York.

Helsel, D. and Hirsch, R. 2002. *Statistical methods in water resources.* Techniques of Water-Resources Investigations Report. U.S. Geological Survey. Book 4, Chapter A3.

Hemson, G., Johnson, P., South, A., Kenward, R., Ripley, R., and Macdonald, D. 2005. *Are kernels the mustard? Data from global positioning system (GPS) collars suggests problems for kernel home-range analysis with least-squares cross-validation.* Journal of Animal Ecology. Vol. 74, No. 3, pp. 455-463.

Hewitt, J., Hanemann, W. 1995. *A Discrete/Continuous Choice Approach to Residential Water Demand under Block Rate Pricing.* Land Economics. Vol. 71, No. 2, pp. 173-192.

Holmes, J. and Moriarty, W. 1999. *Application of the generalized Pareto distribution to extreme value analysis in wind engineering.* Journal of Wind Engineering and Industrial Aerodynamics. Vol. 83, No. 1-3, pp. 1-10.

JFRCCVS. 2012. *Joint Front Range Climate Change Vulnerability Study.* Water Research Foundation.

Kachigan, S. 1991. *Multivariate Statistical Analysis: A Conceptual Introduction.* Radius Press.

Karl, T. and Knight, R. 1998. *Secular trends of precipitation amount, frequency, and intensity in the United States.* Bulletin of the American Meteorological Society. Vol. 79, pp. 231-241.

Katz R. 2010. *Statistics of extremes in climate change.* Special Feature.

Katz, R. and Naveau, P. 2002. *Statistics of extremes in hydrology.* Advances in Water Resources. Vol. 25, pp. 1287-1304.

Katz, R., and Brown, B. 1995. *Regional analysis of temperature extremes: Spatial analog for climate change?* Journal of Climate. Vol. 8, No. 1, pp. 108-119.

Katz, R.; Naveau, P. 2002. *Statistics of extremes in hydrology.* Advances in Water Resources. Vol. 25, pp. 1287-1304.

Katzman, M. 1977. *Income and price elasticities of demand for water in developing countries.* Water Resources Bulletin. Vol. 13, pp. 47-57.

Kenney, D., Geomans, C., Klein, R., Lowrey, J. and Reidy K. 2008. *Residential Water Demand Management: Lessons from Aurora, Colorado.* Vol. 44, No. 1, pp. 192-207.

Lahlou, M., and Colyer, D. 2000. *Water conservation in Casablanca, Morocco.* Journal of American Water Resources Association. Vol. 36, pp. 1003-1012.

Lane, M., Kirshen, P. and Vogel, R. 1999. *Indicators of Impacts of Global Climate Change on U.S. Water Resources.* Journal of Water Resources Planning and Management.

Leadbetter, M., Lindren, G. and Rootzen, H. 1983. *Extremes and related properties of random sequences and processes.* Springer, New York, NY.

Maidment, D. and Miaou, S. 1986. *Daily water use in nine cities.* Water Resources Research. Vol. 22, No. 6, pp. 845-885.

Malla, P. and Gopalakrishnan, C. 1997. *Residential water demand in a fast-growing metropolis: the case of Honolulu, Hawaii.* International Journal Water Resources Development. Vol. 13, pp. 35-51.

Martínez-Espiñeira, R. 2002. *Residential water demand in the Northwest of Spain.* Environmental and Resource Economics. Vol. 21, No. 2, pp. 161-187.

McCullagh P. and Nelder, J. 1989. *Generalized Linear Models.* Chapman and Hall, London.

McKee, T. and Doesken, N. 1995. *Drought monitoring with multiple time scales.* Ninth Conference on Applied Climatology.

McNeil, A. 1997. *Estimating the Tails of Loss Severity Distributions using Extreme Value Theory.* Astin Bulletin.

Miller, K. and Yates, D. 2006. *Climate Change and Water Resources: A Primer for Municipal Water Providers.* AWWA Research Foundation. Denver, CO.

Milly, P. 1994. *Climate, soil water storage, and the average annual water balance.* Water Resources Research. Vol. 30, No. 7, pp. 2143-2156.

Moncur, J. 1987. *Urban water pricing and drought management.* Water Resources Research. Vol. 23, No. 3 pp. 393-398.

Morrison, J. and Smith J. 2002. *Stochastic modeling of flood peaks using the generalized extreme value distribution.* Water Resources Research. Vol. 38, No. 12, pp. 1305-1314.

Myers, R., Montgomery., D., and Vining, G. 2002. *Generalized Linear Models: With Applications in Engineering and the Sciences.* Wiley Publications.

NCDC. 2012. National Climatic Data Center. <http://www.ncdc.noaa.gov/oa/ncdc.html>

Nieswiadomy, M. 1992. *Estimating urban residential water demand: Effects of price structure, conservation, and education.* Water Resources Research. Vol. 28, No. 3, pp 609-615.

Nieswiadomy, M., and Molina, D. 1988. *Urban water demand estimates under increasing block rates.* Growth and Change. Vol. 19, No. 1, pp. 1-12.

Parlange, M. and Katz, R. 2000. *An Extended Version of the Richardson Model for Simulating Daily Weather Variables.* Journal of Applied Meteorology. Vol. 39, pp. 610-622.

Pielke, R., Doesken, N., Bliss, O., Green, T., Chaffin, C., Salas, J., Woodhouse, C., Lukas, J. and Wolter, K. 2005*. Drought 2002 in Colorado: An Unprecedented Drought or a Routine Drought?* Pure and Applied Geophysics. Vol. 162, No. 8-9, pp. 1455-1479.

Rajagopalan, B. 2010. *Hydrologic Frequency Analysis in a Changing Climate.* Workshop on Nonstationarity, Hydrologic Frequency, and Water Management. Boulder, CO. pp. 107.

Rajagopalan, B. and Loll, U. 1996*. A Nonparametric Wet/Dry Spell Model for Resampling Daily Precipitation.* Water Resources Research. Vol. 32, No. 9, pp. 2803-2823.

Rajagopalan, B., Nowak, K., Prairie, J., Hoerling, M., Harding, B., Barsugli, J., Ray, A. and Udall, B. 2009. *Water supply risk on the Colorado River: Can management mitigate?* Water Resources Research. Vol. 45.

Ray, A., Barsugli, J., Averyt, K., Wolter, K., Hoerling, M., Doesken, N., Udall, B. and Webb, R. 2008. *Climate Change in Colorado: A Synthesis to Support Water Resources Management and Adaptation.* Feature Article From Intermountain West Climate Summary.

Regonda, S., Zagona, E., and Balaji, R. 2011. *Prototype Decision Support System for Operations on the Gunnison Basin with Improved Forecasts.* Journal of Water Resources Planning and Management. ACSE September 2011.

Richardson, C. and Wright, D. 1984. *WGEN: A Model for Generating Daily Weather Variables By.* Earth Sciences. Vol. 8, pp. 1-16.

Sankarasubramanian, A. and Loll, U. 2003. Flood quantiles in a changing climate: Seasonal forecasts and causal relations. Water Resources Research. Vol. 39, pp. 1134-1146.

Schneider, M. and Whitlatch E. 1991. *User-Specific Water Demand Elasticities.* Journal of Water Resources Planning and Management. Vol. 117, No. 1.

Smith, R. 1989. *Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone.* Statistical Science. Vol. 4, pp. 367-393.

Smoyer-Tomic, K. 2003. *Heat Wave Hazards: An Overview of Heat Wave Impacts in Canada.* Natural Hazards. Vol. 28, pp. 463-485.

Stehpenson, A., Gilleland, E. 2005. *Software for the analysis of extreme events: The current state and future directions.* Springer Science. Vol. 8, No. 3, pp. 87-109.

Swift, L. and Waide J. 1989. *Refinements in the ZT method of extreme value analysis for small watersheds.* Sixth Conference on Hydrology.

Tampa Bay Water's Long-term Water Demand Forecasting System and Base-Year-2009 Point Demand Forecast. 2010. Hazen and Sawyer.

The R Project for Statistical Computing. 2012. <http://www.r-project.org/>.

Towler, E., Rajagopalan, B., Gilleland, E. Summers, S., Yates, D. and Katz, R. 2010. *Modeling hydrologic and water quality extremes in a changing climate: A statistical approach based on extreme value theory.* Water Resources Research. Vol. 46. 2010.

Trenberth, K. 1998. *Atmospheric moisture residence times and cycling: implications for rainfall rates and climate change.* Climate Change. Vol. 39, pp. 667-694.

Trenberth, K. 1999. *Conceptual framework for changes of extremes of the hydrological cycle with climate change.* Climate Change. Vol. 42, pp. 327-339.

Venables, W. and Ripley B. 2002. *Modern Applied Statistics with S.* 4th edition. Springer, New York, NY.

Vorosmarty, C., Green, P., Salisbury, J. and Lammers, B. 2000. *Global Water Resources: Vulnerability from Climate Change and Population Growth.* Journal of Science. Vol. 289, No. 5477, pp. 284-288.

Weirich, S., Silverstein, J. and Rajagopalan, B. 2011. *Effect of Average Flow and Capacity Utilization on Effluent Water Quality from US Municipal Wastewater Treatment Facilities.* Water Resources Research. Vol. 45, pp. 4279-286.

Wu, R.S. and Haith, D. 1993. *Land Use, Climate, and Water Supply.* J. Water Resources Planning Management. Vol. 119, pp. 685-699.

Yates, D., Gangopadhyay, S., Rajagopalan, B., and Strzepek, K. 2003. *A technique for generating regional climate scenarios using a nearest-neighbor algorithm.* Water Resources Research. Vol. 39, No. 1199, pp. 15.

Yates, D., Sieber, J., Purkey, D. and Huber-Lee, A. 2005. *WEAP21—A Demand-, Priority-, and Preference- Driven Water Planning Model.* Water International. Vol. 30, No. 4.

Zachman, B., Balaji, R., and Summers, R. 2007. *Modeling NOM Breakthrough in GAC Adsorbers Using Nonparametric Regression Techniques.* Environmental Engineering Science. Vol. 24, No. 9, pp. 1280-1296.

7 Appendix A – Additional Case Study

Appendix A consists of modeling results for Tampa Bay, FL using the same techniques that are used on the Aurora, CO case study in Chapters 2, 3 and 4. Appendix A illustrates that the presented approaches can be applied to widely ranging areas.

A.1 Tampa Bay, FL Modeling Results: Statistical Modeling of Seasonal Water Demand Attributes Using Climate Variables

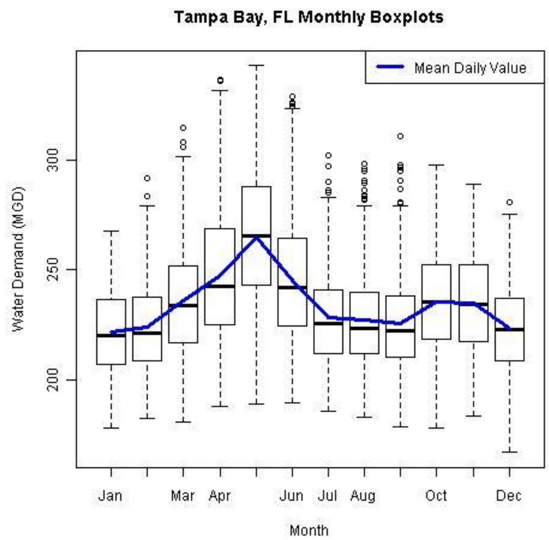


Figure A.1.1 – Monthly boxplots of daily water demand values from 1992-2010 of a water utility at Tampa Bay, FL – boxes represent the interquartile range, the whiskers are 5th and 95th percentiles and the points are beyond.

Weather Statistic Thresholds	
Minimum Temperature	
June	68°F
July	73°F
August	76°F
Maximum Temperature	
June	85°F
July	90°F
August	92°F
Precipitation	
June	0 in
July	0 in
August	0 in

Table A.1.1 – 75th percentile values for weather in Tampa Bay, FL. These values were used as threshold for

developing spell statistics.

Covariate	Definition	Correlation with Monthly Average Demand	Correlation with Monthly Maximum Demand	Correlation with Monthly Demand Exceedances
Average Monthly Max Temperature	Daily maximum temperature values averaged over one month.	0.38	0.20	0.36
Average Seasonal Minimum Temperature	Daily minimum temperature values averaged over one month.	0.33	0.13	0.30
Average Seasonal Precipitation	Daily precipitation values averaged over one month.	0.06	-0.16	0.03
TIME	A vector of values progressing chronologically.	0.23	0.00	0.19
Preceding Month Average Water Demand	Average water demand from preceding month (lag 1).	0.55	0.63	0.53
Average Monthly Daily High Hot Spell in Days	The average length in days of a spell with daily high temperatures over a given threshold for each month.	0.05	-0.02	0.02
Total Monthly Days with High Temperature Above Threshold	Total days the maximum daily temperature exceeds a given threshold in a month.	0.20	0.23	0.19
Maximum Monthly Daily High Hot Spell	The longest spell with daily high temperatures over a given threshold in a given month.	0.12	0.11	0.11
Average Monthly Nightly Low Hot Spell in Day	The average length in days of a spell with nightly low temperatures above a threshold for each month.	0.26	0.13	0.25
Total Monthly Nights with Low Temperature Above Threshold	Total days with the nightly low temperature above a given threshold for each month.	0.16	0.08	0.16
Maximum Monthly Nightly Low Hot Spell	The longest spell in days with nightly low temperatures above a given threshold for each month.	0.27	0.17	0.27
Average Monthly Precipitation Spell in Days	The average length in days of a spell with precipitation above a given threshold for each month.	-0.32	-0.19	-0.27
Total Monthly Days with Precipitation Below Threshold	Total days with the precipitation above a given threshold each month.	-0.49	-0.28	-0.44
Maximum Monthly Precipitation Spell	The longest spell in days with precipitation above a given threshold each month.	-0.44	-0.27	-0.40

Table A.1.2 – Climate Attributes for Tampa Bay, FL; the correlation with water demand attributes shown in the last three columns.

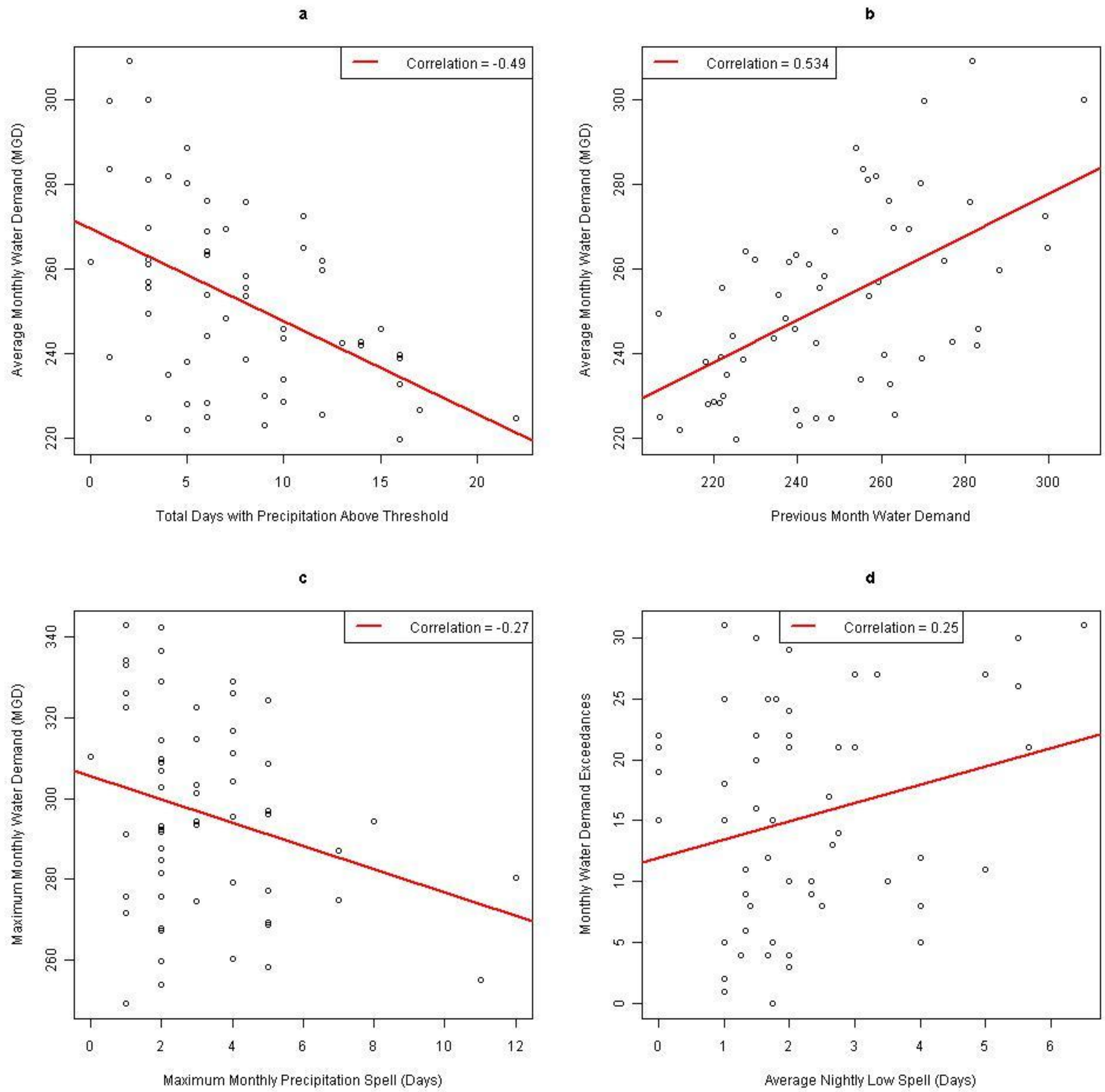


Figure A.1.2 – Scatterplot of selected spell statistics and water demand. (a) average water demand in a month versus total days with precipitation, (b) previous month water demand (antecedent conditions) versus monthly average water demand, (c) Longest monthly precipitation spell versus monthly maximum water demand and (d) Average spell in days with nightly low temperatures above the threshold versus number of days of water demand exceeding 250 MGD water demand value for each month.

Tampa Bay, FL					
Monthly Average Demand					
Independent Variable	Coefficient Estimate	Standard Error	t Value	Pr(>|t|)	R^2

Independent Variable	Coefficient Estimate	Standard Error	t Value	Pr(>|t|)	R^2
Y Axis Intercept	6.15E-03	2.35E-04	26.12	2.00E-16	7.80E-01
Longest Spell with Min Temp Above Threshold	-2.06E-05	6.84E-06	-3.01	4.01E-03	
Total Days with Precip Above Threshold	6.82E-05	1.08E-05	6.33	5.85E-08	
Average Water Demand from Preceding Month	-9.77E-06	9.48E-07	-10.30	3.73E-14	
Monthly Maximum Demand					
Independent Variable	Coefficient Estimate	Standard Error	t Value	Pr(>|t|)	R^2
Y Axis Intercept	5.10E-03	2.40E-04	21.25	2.00E-16	6.55E-01
Average Spell with Max Temp Above Threshold	7.33E-05	2.97E-05	2.72	8.84E-03	
Total Days with Max Temp Above Threshold	-2.31E-05	6.81E-06	-3.39	1.35E-03	
Total Days with Precip Above Threshold	2.49E-05	4.89E-06	5.09	4.95E-06	
Average Water Demand from Preceding Month	-7.72E-06	9.50E-07	-8.12	8.11E-11	
Number of Days per Month with Demand Exceeding a Given Threshold					
Independent Variable	Coefficient Estimate	Standard Error	z Value	Pr(>|t|)	R^2
Y Axis Intercept	-3.53E-01	3.58E-01	-0.99	3.24E-01	6.27E-01
Average Spell with Max Temp Above Threshold	-1.27E-01	4.63E-02	-2.74	6.08E-03	
Total Days with Max Temp Above Threshold	2.63E-02	1.18E-02	2.23	2.55E-02	
Longest Spell with Min Temp Above Threshold	3.06E-02	1.26E-02	2.44	1.49E-02	
Total Days with Precip Above Threshold	-7.31E-02	7.79E-03	-9.39	2.00E-16	
Average Water Demand from Preceding Month	1.42E-02	1.38E-03	10.32	2.00E-16	

Table A.1.3 – Best GLM models selected using BIC for the three demand attributes. The Pr(>|t|) value notifies significance for the coefficient estimate.

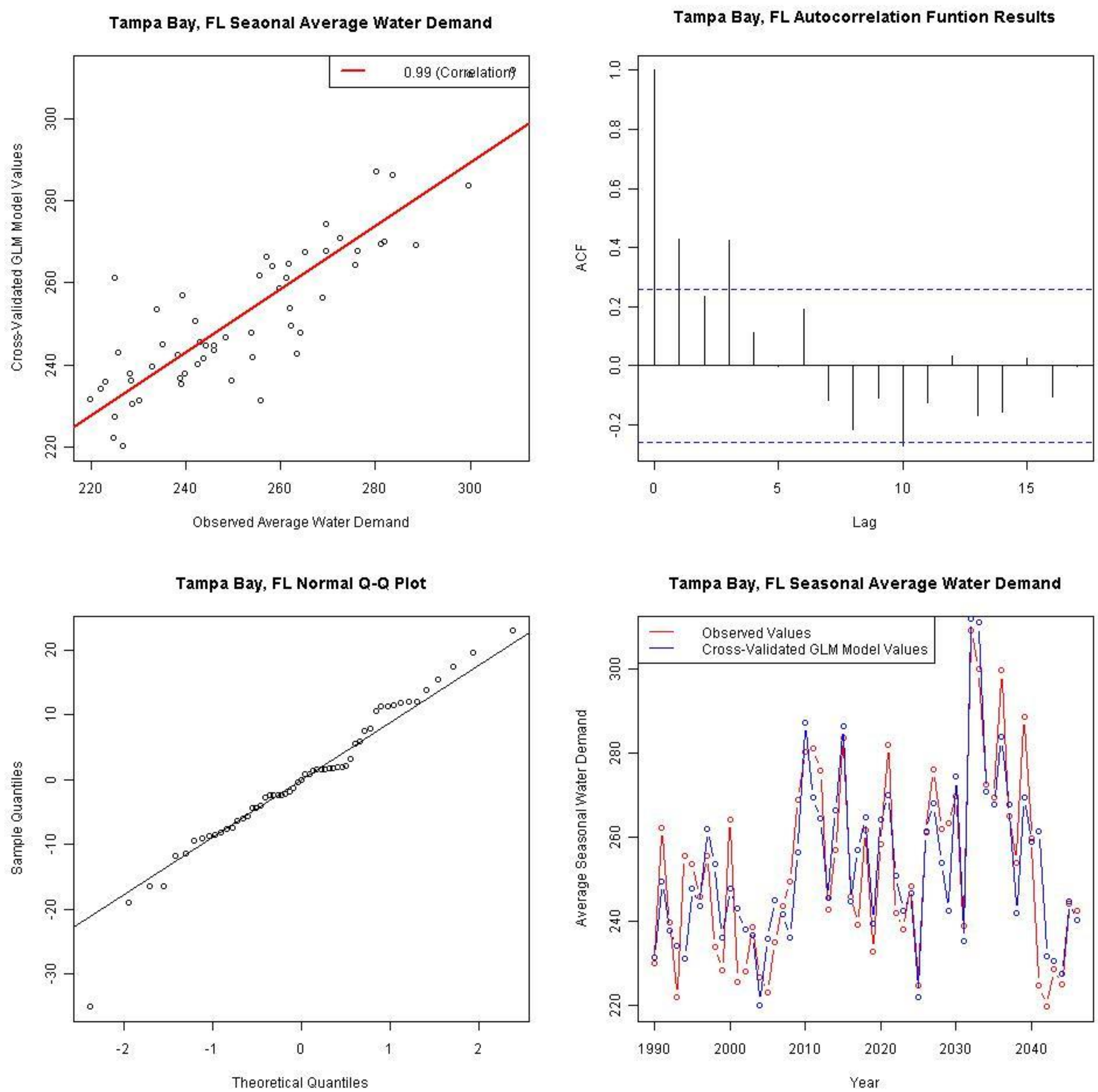


Figure A.1.3 – Diagnostics plots for monthly average demand – (a) Scatterplot of observed vs. model estimates with the linear fit line, (b) autocorrelation of model residuals, (c) Q-Q plot of model residuals and (d) leave one out cross validated estimates versus observed value.

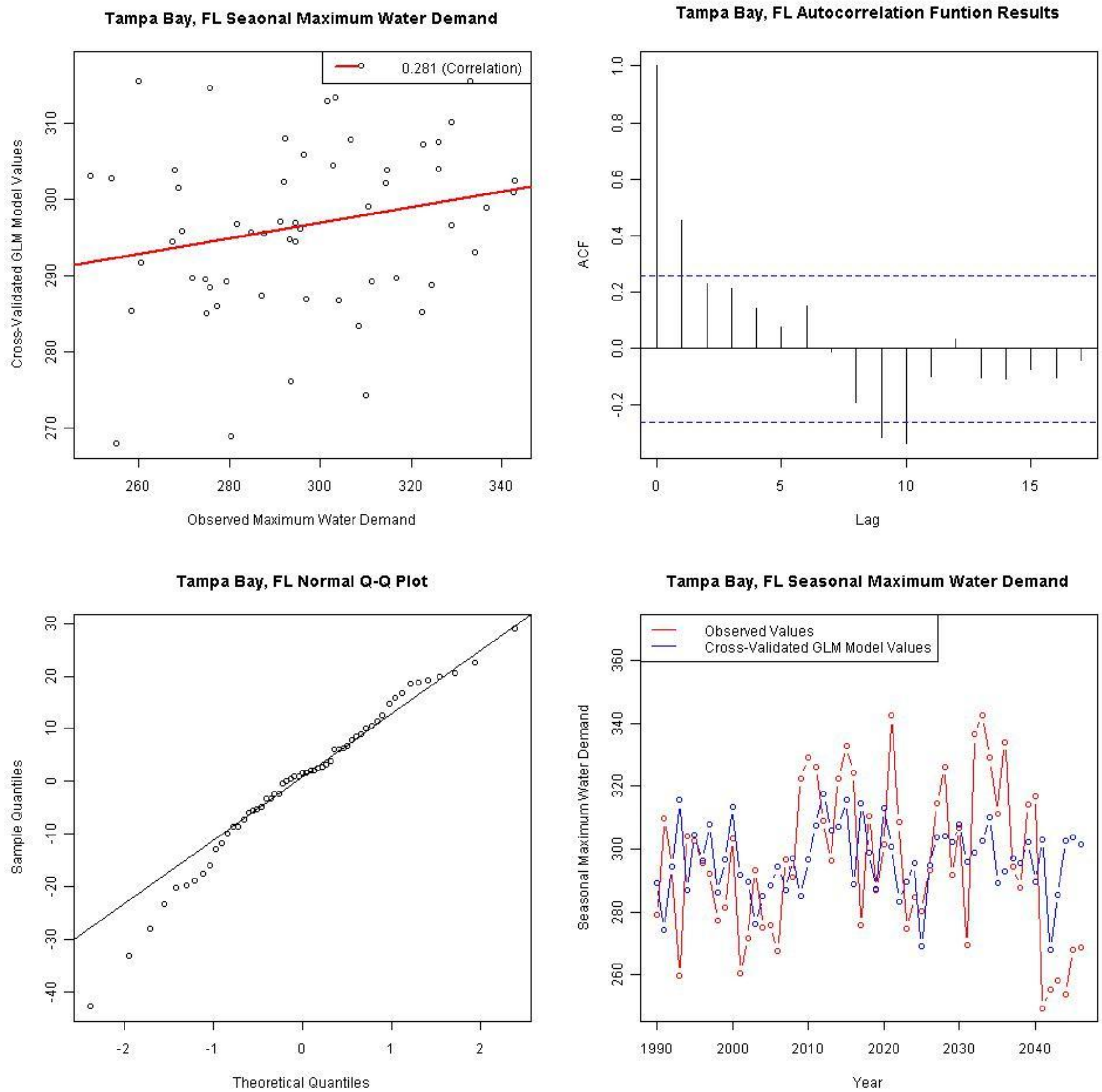


Figure A.1.4 – Same as Figure 5 but for monthly maximum demand.

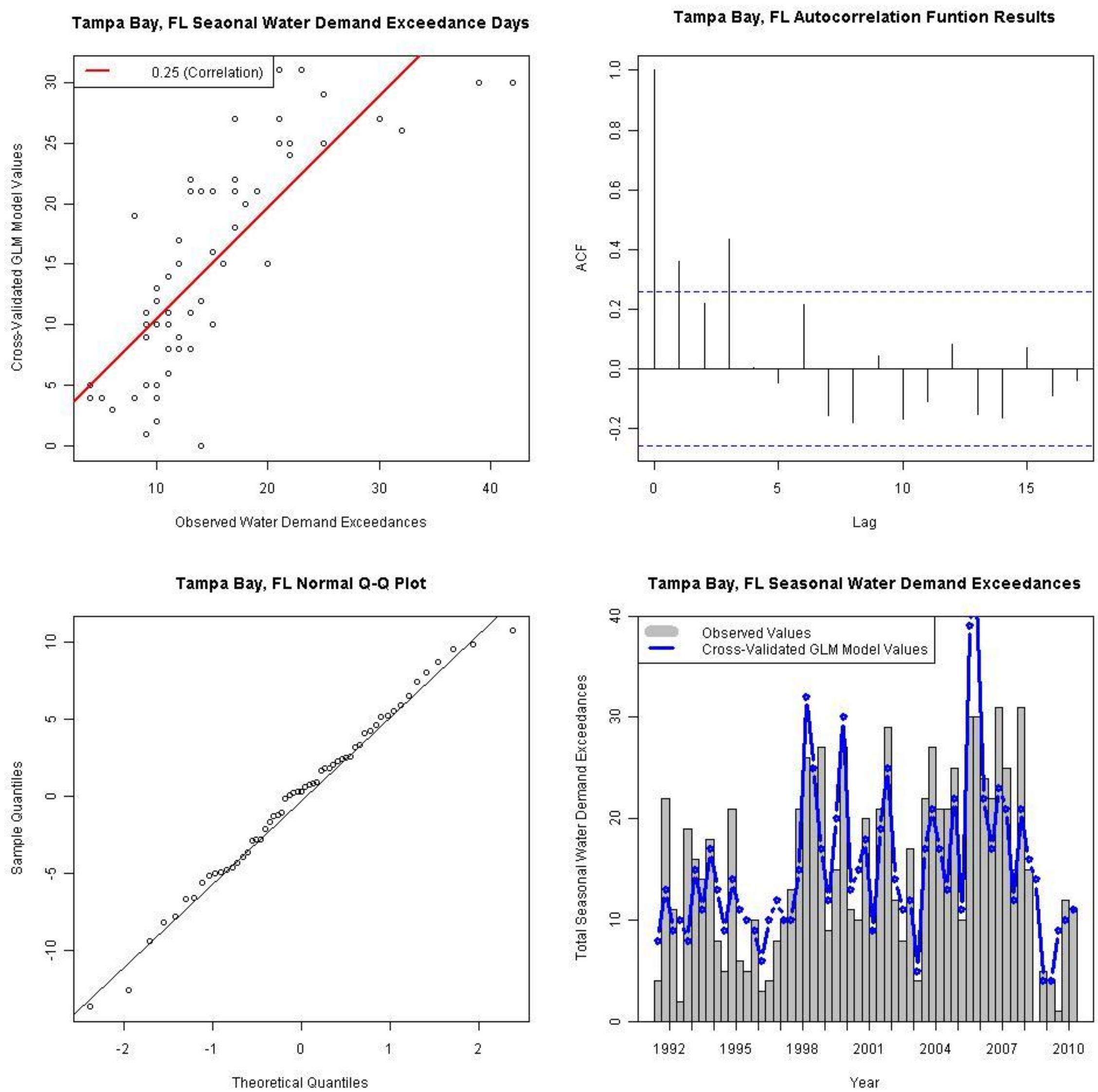


Figure A.1.5 – Same as Figure 5 but for number of days of demand exceeding a threshold of 81.3 MGD (75th percentile value).

A.2 Tampa Bay, FL Modeling Results: Modeling Water Demand Extremes Using Extreme Value Analysis

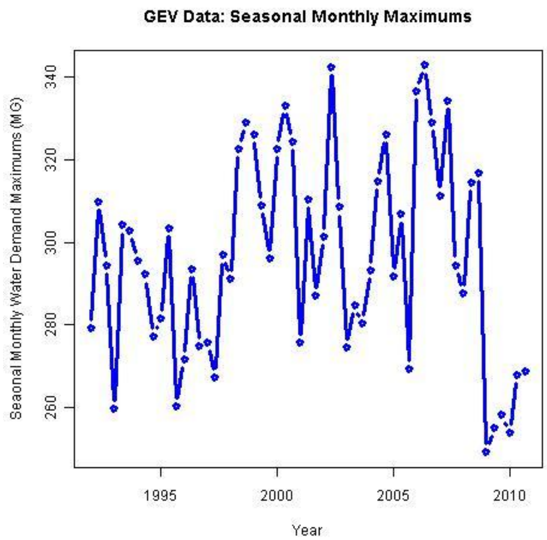


Figure A.2.1 –Monthly maximum water demand (MGD) (Jun-Aug).

Generalized Extreme Value Model

Variable	Stationary Model	Non-Stationary Models		
		$\sigma = \beta_0+\beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)$	$\boldsymbol{\mu = \beta_0+\beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)}$	$\sigma = \beta_0+\beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)$ & $\mu = \beta_5+\beta_6(a) + \beta_7(b) + \beta_8(c) + \beta_9(d)$
β_0		18.66	**294.92**	289.54
β_1	-	0.27	**0.78**	1.45
β_2	-	-0.85	**0.34**	-0.54
β_3	-	-0.61	**-1.44**	0.01
β_4	-	0.09	**-**	21.47
β_5	-		**-**	0.06
β_6	-	-	**-**	-0.71
β_7	-	-	**-**	-1.30
β_8	-	-	**-**	0.19
β_9		-	**-**	-
Location Parameter (μ)	288.16	283.16	**-**	-
Scale Parameter (σ)	24.47	-	**23.91**	-
Shape Parameter (ξ)	-0.33	-0.04	**-0.40**	-0.95
Negative Log Likelihood	261.99	260.89	**257.97**	
Significance (p-value)		Fail	**Yes**	Fail

**Where (a) is the longest monthly spell with maximum temperatures above a threshold, (b) is the longest monthly spell with minimum temperatures above a threshold, (c) is the total days with precipitation above a threshold and (d) is TIME.

Table A.2.1 – Generalized Extreme Value Distribution model results. The column in bold is the best fit model.

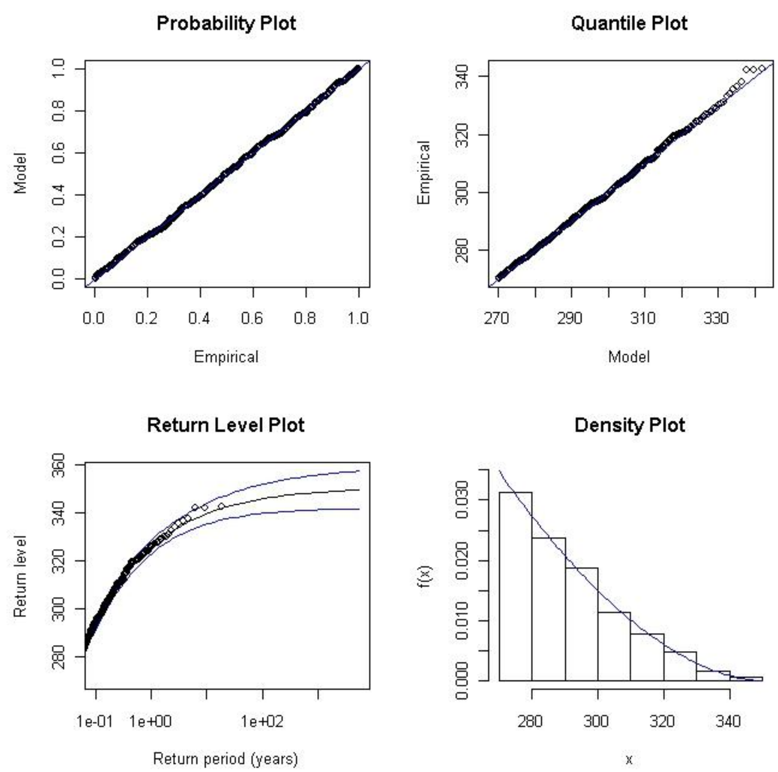


Figure A.2.2 - Diagnostics for the stationary GEV model include (a) probability plot showing empirical and model probabilities with the 1:1 line, (b) quantile plot showing observed and model quantiles with the 1:1 line, (c) water demands of various return periods from the data (shown as dots), model (solid line) and the confidence intervals (dotted) and (d) histogram of the data with the fitted GEV probability density function.

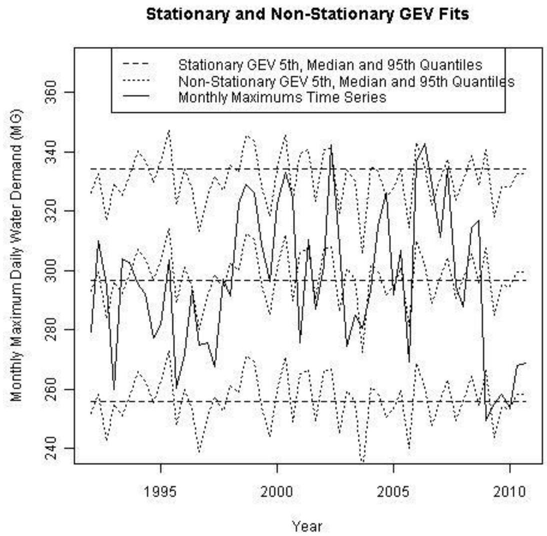


Figure A.2.3 – Quantiles from stationary and nonstationary GEV models. Solid line shows the monthly maximum water demand, dotted lines are the 0.05 and 0.95 quantiles from the nonstationary GEV, and dashed lines are the quantiles from stationary GEV model.

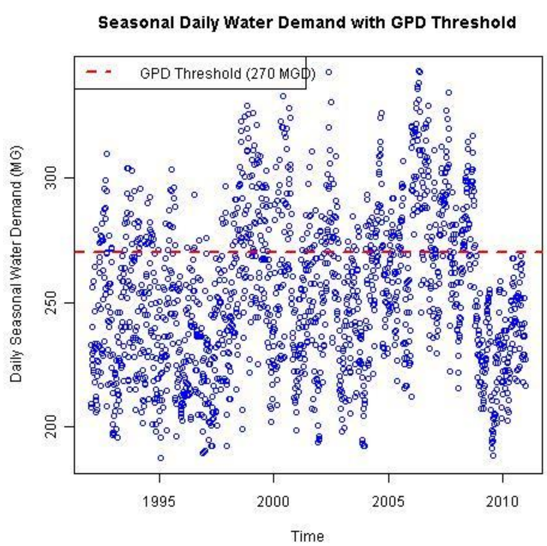
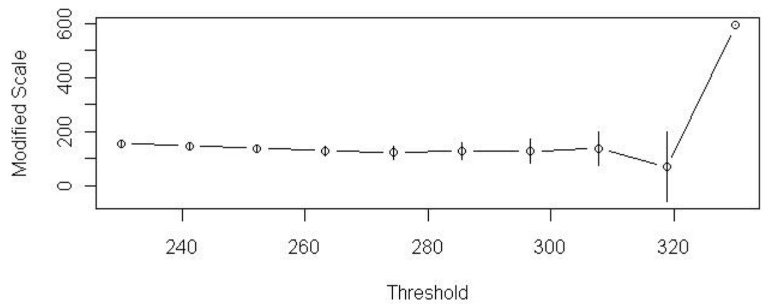


Figure A.2.4 – Daily water demand during Jun-Aug of each year. The dotted line is the threshold corresponding to 90 MGD, considered to be high enough that exceeding it would stress the water supply of the utility.



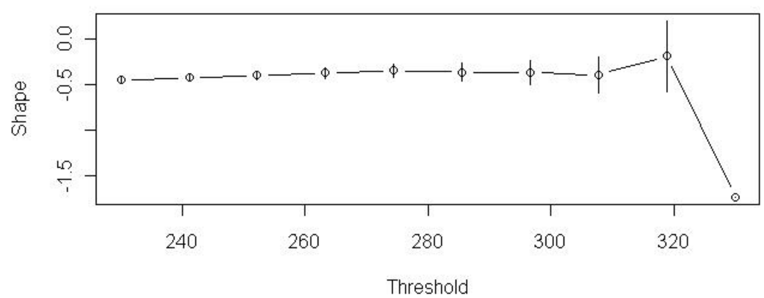


Figure A.2.5 – Sensitivity of the scale parameter to choice of threshold. A threshold selection of 90 MGD is in the stable region and an appropriate choice for GPD modeling.

Variable	Stationary Model	**Non-Stationary Model** $\sigma = \beta_0+\beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)$
β_0	-	**29.60**
$\beta 1$	-	**-0.15**
$\beta 2$	-	**0.06**
$\beta 3$	-	**-11.53**
$\beta 4$	-	**0.01**
Scale Parameter (σ)	28.73	**-**
Shape Parameter (ξ)	-0.36	**-0.39**
Negative Log Likelihood	2037.37	**2024.16**
Significance (p-value)	-	**Yes**

*** (a) = Daily Maximum Temperature, (b) = Daily Minimum Temperature, (c) = Daily Precipitation, (d) = Time

Table A.2.2 – Generalized Parateo Distribution model results. The column in bold is the best fit model.

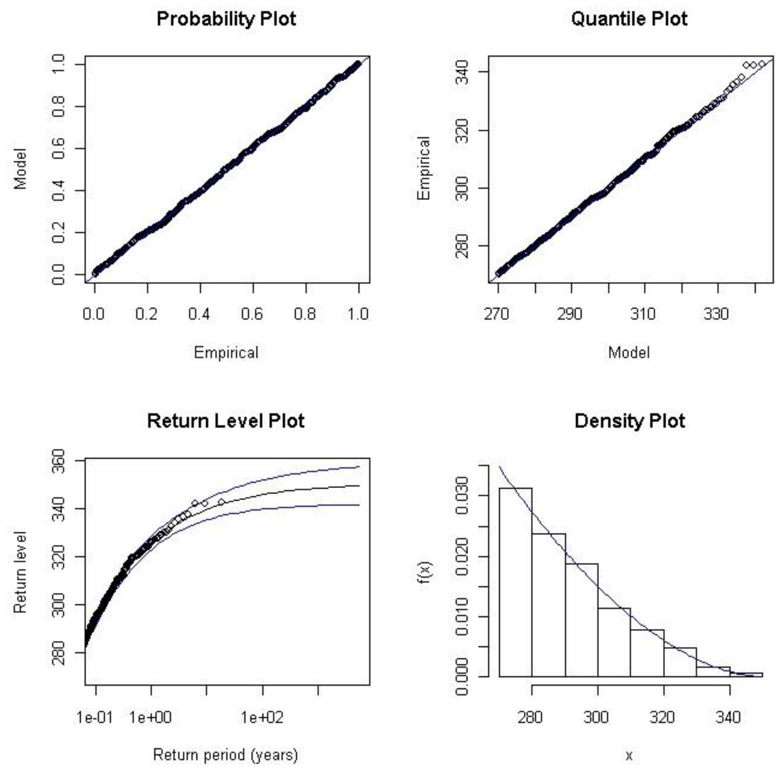


Figure A.2.6 - Same as Figure 2 but for the GPD model.

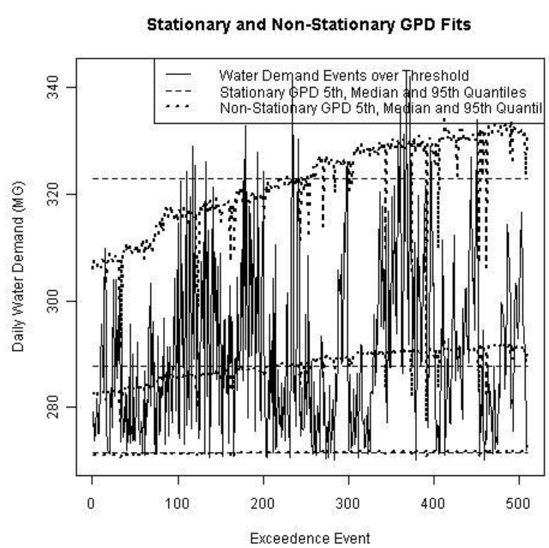


Figure A.2.7 – Quantiles from stationary and nonstationary GPD models. Solid line shows the daily water demand, dotted lines are the 0.05 and 0.95 quantiles from the nonstationary GPD, and dashed lines are the quantiles from stationary GPD model.

A.3 Tampa Bay, FL Modeling Results: Projecting Demand Extremes under Climate Change Using Extreme Value Analysis

Weather Statistic Thresholds	
Minimum Temperature	
June	68°F
July	73°F
August	76°F
Maximum Temperature	
June	85°F
July	90°F
August	92°F
Precipitation	
June	0 in
July	0 in
August	0 in

Table A.3.1 –Threshold of daily minimum and maximum temperatures and, precipitation, to compute weather attributes for use as covariates in the extreme value analysis models.

Water Demand Covariates: Definitions and Correlations				
Covariate	**Definition**	**Correlation with Monthly Average Demand**	**Correlation with Monthly Maximum Demand**	**Correlation with Monthly Demand Exceedances**
Average Monthly Max Temperature	Daily maximum temperature values averaged over one month.	0.38	0.20	0.36
Average Seasonal Minimum Temperature	Daily minimum temperature values averaged over one month.	0.33	0.13	0.30
Average Seasonal Precipitation	Daily precipitation values averaged over one month.	0.06	-0.16	0.03
TIME	A vector of values progressing chronologically.	0.23	0.00	0.19
Preceding Month Average Water Demand	Average water demand from preceding month (lag 1).	0.55	0.63	0.53

Average Monthly Daily High Hot Spell in Days	The average length in days of a spell with daily high temperatures over a given threshold for each month.	0.05	-0.02	0.02
Total Monthly Days with High Temperature Above Threshold	Total days the maximum daily temperature exceeds a given threshold in a month.	0.20	0.23	0.19
Maximum Monthly Daily High Hot Spell	The longest spell with daily high temperatures over a given threshold in a given month.	0.12	0.11	0.11
Average Monthly Nightly Low Hot Spell in Day	The average length in days of a spell with nightly low temperatures above a threshold for each month.	0.26	0.13	0.25
Total Monthly Nights with Low Temperature Above Threshold	Total days with the nightly low temperature above a given threshold for each month.	0.16	0.08	0.16
Maximum Monthly Nightly Low Hot Spell	The longest spell in days with nightly low temperatures above a given threshold for each month.	0.27	0.17	0.27
Average Monthly Precipitation Spell in Days	The average length in days of a spell with precipitation above a given threshold for each month.	-0.32	-0.19	-0.27
Total Monthly Days with Precipitation Below Threshold	Total days with the precipitation above a given threshold each month.	-0.49	-0.28	-0.44
Maximum Monthly Precipitation Spell	The longest spell in days with precipitation above a given threshold each month.	-0.44	-0.27	-0.40

Table A.3.2 – Weather attributes and their correlation with monthly maximum water demand. Bolded values are statistically significant at 95% confidence level and, bolded rows are variables used in the nonstationary GEV as covariates.

Scenario	GCM/Ensemble	30 Year Precipitation Change (%)	30 Year Average Temperature Change (°F)
Warm/Wet	ncar_ccsm3_0.2	3.77	3.40
Warm/Dry	Micro3_2_medres.1	-8.51	3.40

Table A.3.3 – Two climate change scenarios used in stochastic weather generation.

Generalized Extreme Value Model

	Non-Stationary Models

Variable	Stationary Model	$\sigma = \beta_0+\beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)$	$\mu = \beta_0+\beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)$	$\sigma = \beta_0+\beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)$ & $\mu = \beta_5+\beta_6(a) + \beta_7(b) + \beta_8(c) + \beta_9(d)$
β_0		18.66	**294.92**	289.54
β_1	-	0.27	**0.78**	1.45
β_2	-	-0.85	**0.34**	-0.54
β_3	-	-0.61	**-1.44**	0.01
β_4	-	0.09	**-**	21.47
β_5	-		**-**	0.06
β_6	-	-	**-**	-0.71
β_7	-	-	**-**	-1.30
β_8	-	-	**-**	0.19
β_9		-	**-**	-
Location Parameter (μ)	288.16	283.16	**-**	-
Scale Parameter (σ)	24.47	-	**23.91**	-
Shape Parameter (ξ)	-0.33	-0.04	**-0.40**	-0.95
Negative Log Likelihood	261.99	260.89	**257.97**	
Significance (p-value)		Fail	**Yes**	Fail

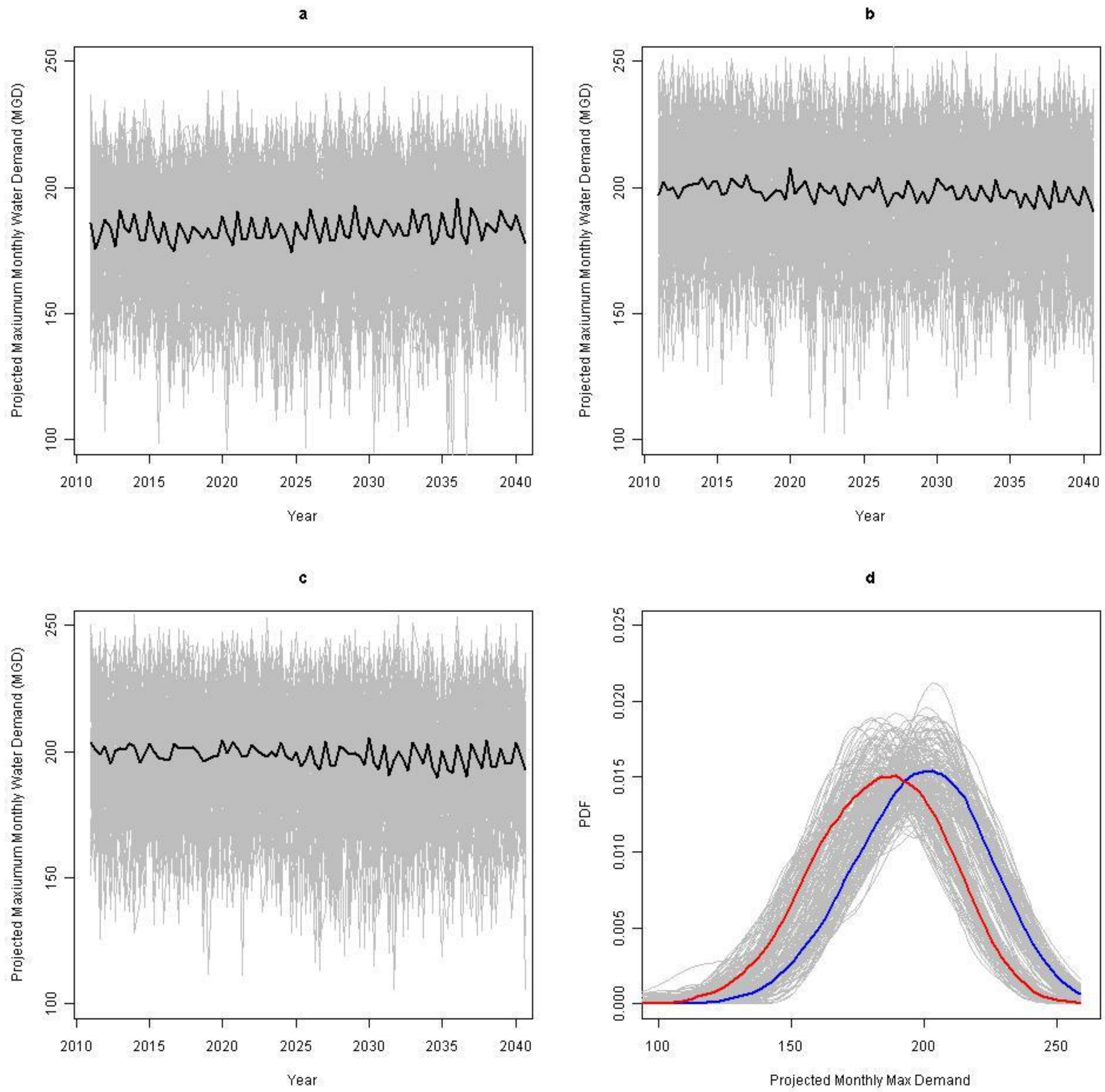
**Where (a) is the longest monthly spell with maximum temperatures above a threshold, (b) is the longest monthly spell with minimum temperatures above a threshold, (c) is the total days with precipitation above a threshold and (d) is TIME.

Table A.3.4 – Generalized Extreme Value distribution model

Variable	Stationary Model	**Non-Stationary Model**
		$\sigma = \beta_0+\beta_1(a) + \beta_2(b) + \beta_3(c) + \beta_4(d)$
β_0	-	**29.60**
β1	-	**-0.15**
β2	-	**0.06**
β3	-	**-11.53**
β4	-	**0.01**
Scale Parameter (σ)	28.73	**-**
Shape Parameter (ξ)	-0.36	**-0.39**
Negative Log Likelihood	2037.37	**2024.16**
Significance (p-value)	-	**Yes**

*** (a) = Daily Maximum Temperature, (b) = Daily Minimum Temperature, (c) = Daily Precipitation, (d) = Time

Table A.3.5 – Generalized Parateo Distribution and Poisson models used for projections.


a
b
c
d
Projected Maxiumum Monthly Water Demand (MGD)
Projected Maxiumum Monthly Water Demand (MGD)
Projected Maxiumum Monthly Water Demand (MGD)
Year
Year
Year
PDF
Projected Monthly Max Demand
100
150
200
250
2010
2015
2020
2025
2030
2035
2040
0.000
0.005
0.010
0.015
0.020
0.025

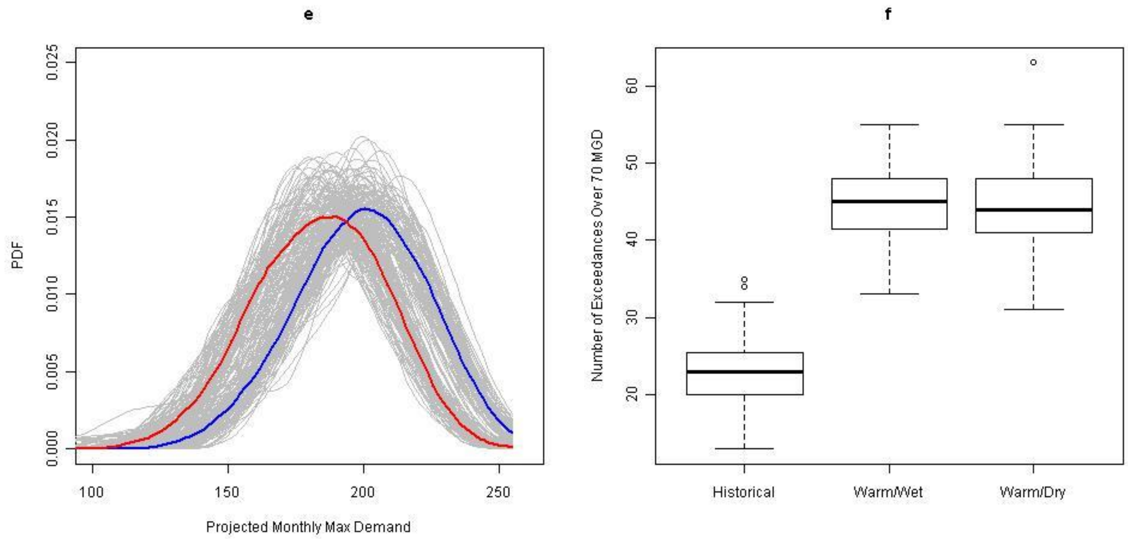


Figure A.3.1 - Projections of monthly maximum water demand from nonstationary GEV model (a) based on natural climate variability (b) based on warm/wet future climate projection, (c) based on warm/dry future climate. (d) is the Probability density functions of monthly maximum water demand projections for warm/wet future climate (grey lines are from individual simulation and blue is the average) for natural climate variability (red), (e) is the same as (d) but for warm/dry climate projections and (f) is Boxplots of the number of monthly maximum demand exceedances of 85 MGD for the three scenarios.

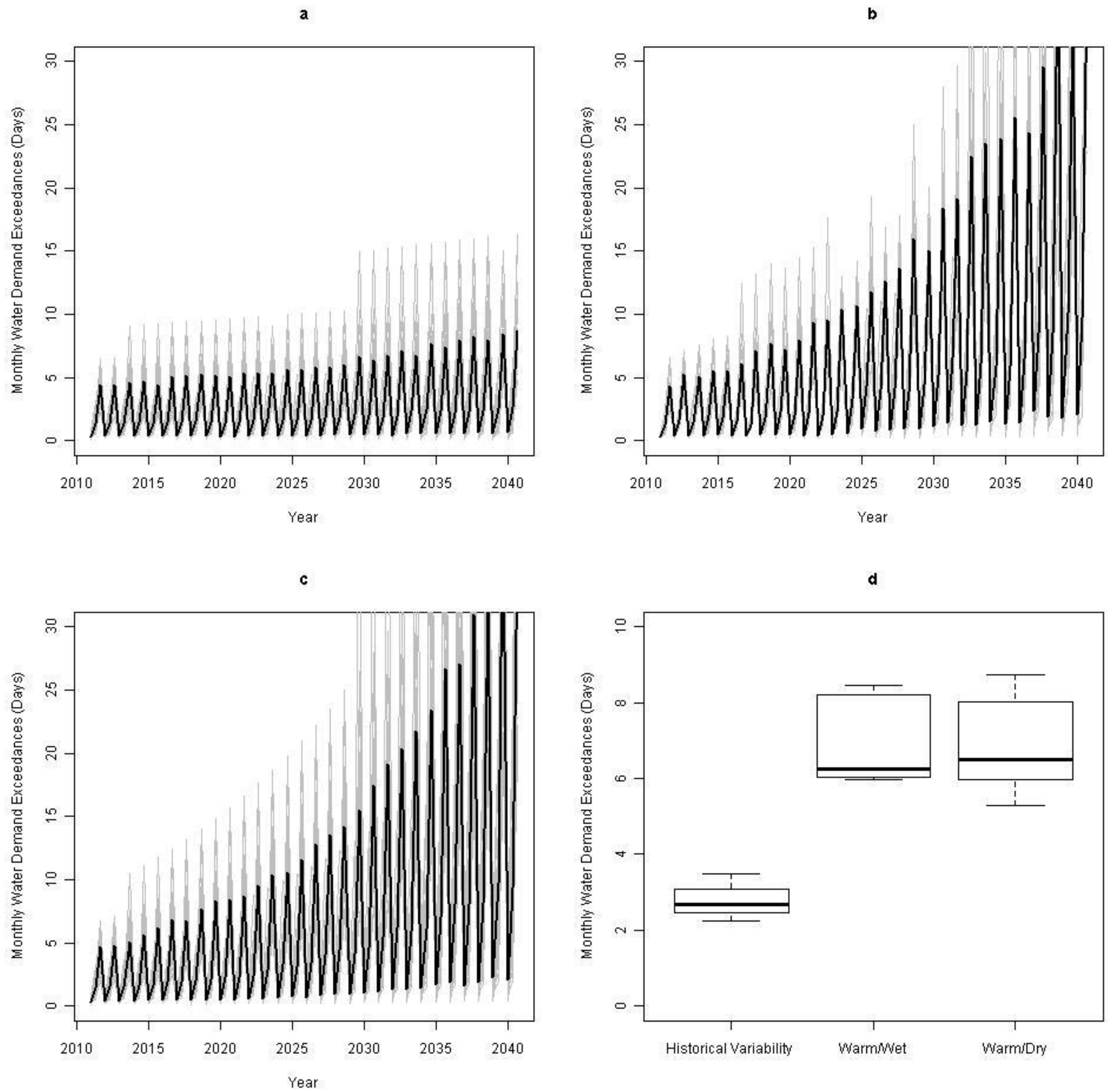


Figure A.3.2 - (a) Projections of number of monthly exceedances of daily water demand over 90 MGD from the Poisson model (a) for natural variability, (b) for warm/wet climate projection, (c) for warm/dry climate projection. (d) Probability density functions of number of monthly exceedance for natural variability and warm/wet climate projection (dashed), its average in blue and the average from natural variability in red and (e) same as (d) but for warm/dry climate projection.

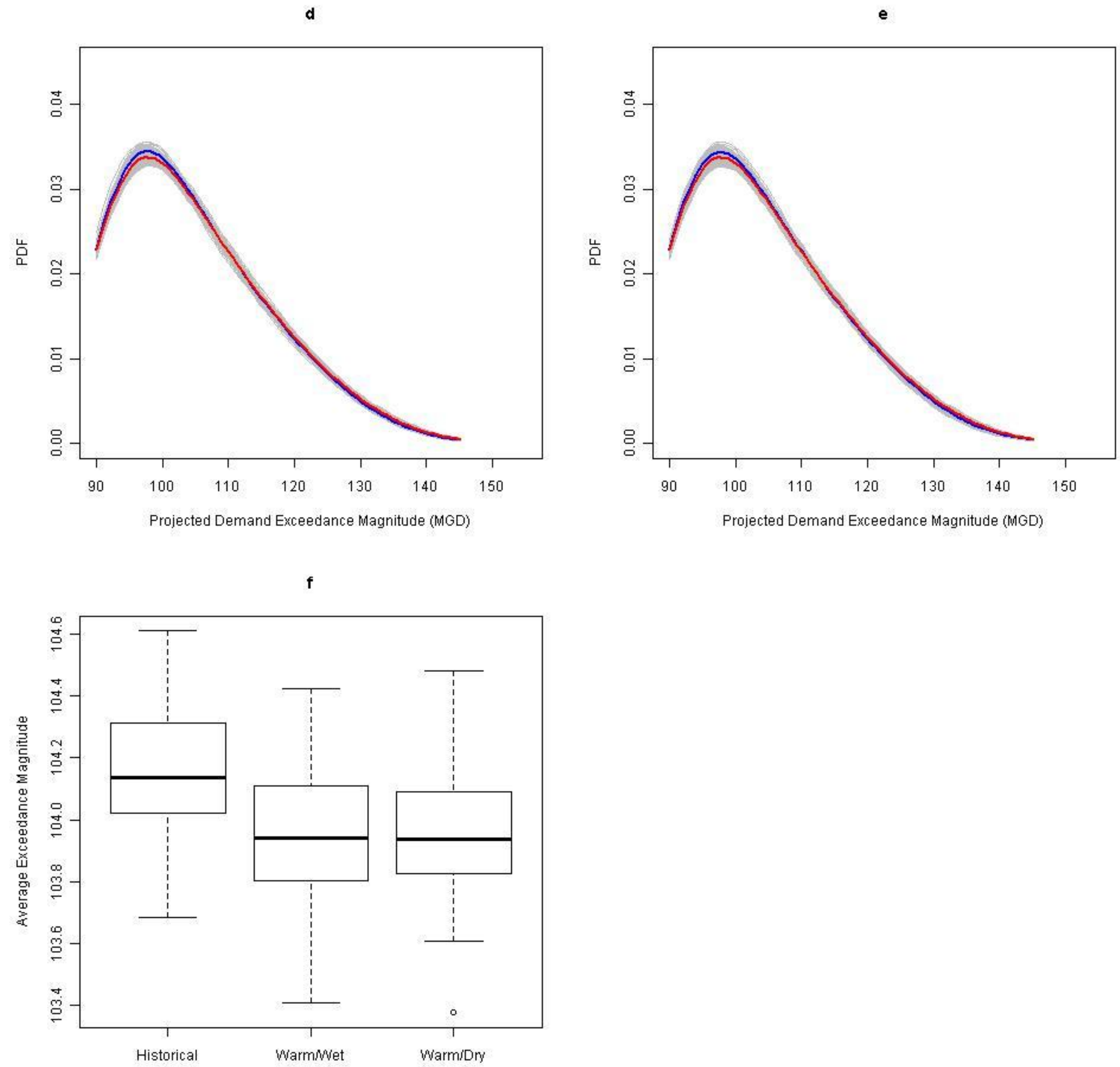


Figure A.3.3 - (a) Probability density functions of the magnitude of exceedances for warm/wet climate projection (in gray), the average PDF is shown in blue and the average from natural variability in red, (b) same as (a) but for warm/dry climate projection and (c) boxplots of average exceedances for natural variability, warm/wet and warm/dry climate projections.