SPACE-TIME DYNAMICS OF SINGLE FAMILY RESIDENTIAL WATER CONSUMPTION IN CHARLOTTE, NORTH CAROLINA

by

Yuhong Zhou

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Geography and Urban Regional Analysis

Charlotte

2016

Approved by:

____
Dr. Jean-Claude Thill

____
Dr. Harrison Campbell

____
Dr. Wenwu Tang

____
Dr. Sandra Dika

ABSTRACT

YUHONG ZHOU. Space-time dynamics of single family residential water consumption in Charlotte, North Carolina. (Under the direction of DR. JEAN-CLAUDE THILL)

Water availability has become a more significant economic and policy issue in contemporary America. Although emerging as an attractive tool to water authorities, demand-side water management in urbanized areas is more complicated due to the high complexity of coupled human and natural (mainly water and land) systems and great heterogeneity of households and neighborhoods in urban environments. A better understanding of how water is used by whom and in what ways water savings can be realized will be necessary and useful for planning, implementing, and evaluating demand-side alternatives. The purpose of this research is to understand the spatial and temporal dynamics of water use or demand, and of its relations with various factors in a fast-growing urban environment of Charlotte, North Carolina.

Using the water billing datasets over the 2000-2010 period for Mecklenburg County, North Carolina, this dissertation conducts a multidimensional investigation on the state, pattern, and process of the subject – single family residential (SFR) water consumption. It first explores the decennial evolution of SFR water usage and its association and sensitivity with historical weather conditions in Charlotte. Next, it examines the historically contingent effects of sociodemographic and housing factors on yearly SFR water consumption as well as the spatial heterogeneity and dependence in these effects. Lastly, monthly SFR water usage per household at the neighborhood level during the 2007-2009 period is explained by various factors (pricing, water usage

restriction, weather, sociodemographic and housing characteristics) within a spatial econometric modeling framework.

The results show that SFR water consumption is unevenly distributed across Charlotte and some neighborhoods (especially in southern Charlotte) consumed considerably more water in summer than winter; spatial variability in climatic sensitivity of neighborhoods is evident; the historical states of explanatory factors have more influence on SFR water usage in 2008 than their temporal change between 2000 and 2008 and those effects vary across space; the importance of price, non-price policy, mean temperature and precipitation in affecting monthly SFR water consumption during 2007-2009 is highlighted after spatial heterogeneity is accounted for.

This research is vital to the enhancement of the local community's knowledge. The multidimensional analyses will not only offer first-hand evidence for answering critical questions on weather sensitivity and driving factors related to local water usage, but also help identify new research questions, formulate novel hypotheses, and introduce more opportunities for disentangling the problems in water management in a comprehensive manner.

DEDICATION

To my grandparents and parents

and

those who value water and life

ACKNOWLEDGEMENTS

I cannot offer enough thanks to my advisor Jean-Claude Thill for his outstanding mentorship and continuous support during the long journey of pursuing my PhD. Without his endless inspiration, insightful suggestions and overall guidance, I would not have accomplished this dissertation. It has been my honor and pleasure to be his student and work under his guidance. His enthusiasm for new knowledge and ideas, critical thinking and diligent endeavor on scientific exploration has impressed me and will continue to motivate me to improve my research. I would also like to express my sincere gratitude to my other committee members: Harrison Campbell, Wenwu Tang, and Sandra Dika, for their valuable comments, attention to detail, great patience, and significant support.

I would like to thank Eric Delmelle for his instruction, collaboration, support and encouragement. I am appreciative of the instruction and support I received from Harrison Campbell, Wenwu Tang, Wei-Ning Xiang, Mike Duncan and Qingfang Wang during my professional development.

I sincerely acknowledge Janni Sorensen and Heather Smith for their great support as Program Coordinators and their excellent instructions as educators. I also thank Stephen Bowman and Laurie Garo for their kind help during my dissertation research. The communications with James Birkett and Julie Green from the Graduate School of UNC Charlotte are appreciated as well.

I am particularly grateful to Charlotte Water for supporting our research and providing valuable data. Barry Gullet (Director), Barry Shearin (Deputy Director), and Kim Eagle (Immediate Past-Deputy Director) facilitated this research initiative. Doug Groce (IT Manager) and Harold Deese (Sr. Business Analyst) have offered greatest assistance with delivering data and answering questions. The helps from Rob Bailey (Content Delivery Manager), Allen D. Brunson (Sr. Project Engineer), and David Czerr (Program Manager) are also appreciated. I would also like to acknowledge Wendy Mann (Head of Data Services Group) from the Fenwick Library of George Mason University and Lisa Sutton (Digital Spatial Data Librarian) from the American Geographical Society Library of University of Wisconsin-Milwaukee for their kind help with census data collection. Moreover, I thank Karen Isley from Mecklenburg County, Lori Quinn and Shad Spencer from the City of Charlotte, and many others who supported the research projects I participated in.

In terms of financial support, I am appreciative of the GASP funding from the Graduate School of UNC Charlotte and the BRIDGES Geography Fellowship as well as the research assistantship and conference travel funding from the Department of Geography and Earth Sciences at UNC Charlotte. I also acknowledge the teaching practice opportunity that the summer Teaching Program for Doctoral Students offered.

I value greatly the friendships established during my time at UNC Charlotte. I would like to thank Thi Hong Diep Dao, Mona Kashiha, Elizabeth Delmelle, Jae Soen Son, Zhaoya Gong, Kailas Venkitasubramanian, Xuchu Meng, Ran Tao, Min Chen, Min He, Chunchun Hu, Dingxiang Liu, Lirong Luo, Wenxin Jiang, Yuanyuan Zhang, Hongmei Zhu, and Yang Li for sharing their life experiences, knowledge and joy with me and providing their significant supports when I needed. I also appreciate the many moments shared with Paul McDaniel, Kara Tiller, Amy Davis, Thad Dixon, Tomas Vaclavik, Ryan James, Kunwar Singh, Monica Dorning, Ying Zhu, Shikai Tang, Yi Liu, Sisi Yan, Xiaojia Wang, Meng Jia, Rui Shen and Lu Zhao. Also, I am grateful to my new friends and colleagues Kirsten Beyer and Kelly Hoormann in Milwaukee for their support and encouragement.

Last but not least, my utmost appreciation goes to my family. I owe my parents Moyun Chen and Lamei Chen a debt of gratitude for their endless love and continuous support and encouragement. I would also like to thank my loving husband Xiangdong for his understanding, patience, care and support. Furthermore, I thank my lovely son Lele for his smile and love and my dear brother and his wife for their enthusiastic encouragement. Additionally, I wish for wind to help deliver my love and appreciation to my grandma who had raised me and passed away in the middle of this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Water availability has become a more significant economic and policy issue in contemporary America, especially in the areas where the demand for water is high and the supply of water is scarce. Urbanized areas with continuing and rapid growth face more problems and challenges related to water. On the one hand, continued growth demands more water and there is great pressure to exploit sources of water to the point of exhaustion and to make more costly investment to transfer water (Coyne 2003). On the other hand, the nature of urban/land development itself stresses water resources. Common problems with land use practices include wastewater discharge and run-off pollutants that degrade water quality, and impervious surfaces that impede the natural filtration of precipitation and water flows into the soils and thus affect water quality and water supply (Arnold 2005). The adverse impacts of urban development do great harm to aquatic ecosystems and water resources, which results in a decline in water supply and an inverse effect on water availability in those areas, such as the occurrence of water shortage or increased water scarcity (Shandas et al. 2015). In order to effectively use water, water (resource) management has been emphasized in the countries all over the world. The United States of America has devoted itself more to water (resource) management since the 1980s (Dziegielewski 1999).

Managing water demand and supply is part of water management practices. Demand-side management has emerged more recently as a complement of more traditional approaches and proves to be crucial for coping with water vulnerability in the face of climate change (Bates et al. 2008; Russell and Fielding 2010). For urbanized areas, this task becomes more complicated due to the high complexity in the coupled human and natural (mainly water and land) systems (House-Peters and Chang 2011) and great heterogeneity of households and neighborhoods in urban environments. In addition to management of water sources, water demand and supply in urban settings are associated with the land development process and the water infrastructure construction process, respectively. Water use is mainly driven by different consuming behaviors of diverse human groups, and sometimes urban land uses contribute to water consumption pattern. Hence the form of urban development and the socio-economic characteristics attached with land uses are the most significant factors of water demand.

There may also exist inequity among water users (Ingram et al. 2008). Actual water consumption can be different from the desired consumption, because it is constrained by both the users' access to water supply infrastructure and the affordability of water service. Water affordability is dependent on the price of the water service and the social and economic status of water consumers. Some social groups such as low-income households may face water poverty issue (poverty related to water availability and human welfare) (Porcher 2013) or be at a more disadvantageous position when coping with the change in water pricing (Agthe and Billings 1987).

Understanding the dynamics of water consumption at the city level is critical for local public water utility, planners and policy makers for several reasons. From an idiographic viewpoint, the uniqueness and complexity of a region/area could be best revealed by comprehensive empirical studies. The problems identified in water management in warm, humid cities like Charlotte, North Carolina, are less likely to be same as the ones in hot, dry cities like Phoenix, Arizona. Most of the case studies related to water use and demand in the United States focus on the cities or metropolis in the west such as Phoenix, Arizona and Portland, Oregon. Emerging concerns on the pressure of rising demand, future droughts and climate change (Henderson 2015) on water supply in the Charlotte-Mecklenburg area have called for a better understanding on local consumption.

From a systematic perspective, to come up with a work plan appropriate to the specific characteristics of water consumption, "to do some basic counting" will not be adequate. Instead, we have to grasp the multiple facets of water use dynamics. Thus a multidimensional investigation on the state, pattern, and process of the subject – water consumption, is essential. Among the many dimensions of water use, its relations to climate trend and change, and its driving factors are of interest to researchers and policy-makers. Given that the concepts location and time are involved in the process of consuming water, more attentions should also be paid to the spatial and temporal aspects of water use dynamics (a spatial science viewpoint/focus).

1.1 Statement of Research

The purpose of this research is to understand the spatial and temporal dynamics of water use or demand, and of its relations with various factors in a fast-growing urban environment. Using the water billing datasets over the period between 2000 and 2010 for the county of Mecklenburg, North Carolina (NC), this dissertation aims to identify patterns, contingency, dependence, associations, and effects related to water use at the selected temporal (monthly and yearly) and spatial scale (block group).

This research has three major objectives aimed at contributing to the understanding of the dynamics of single-family residential (SFR) water use at an intra-urban scale. The first objective is to explore the decennial evolution of SFR water use and its association and sensitivity with historical climate conditions in Charlotte, Mecklenburg County, North Carolina. The second component of this research is to investigate the historically contingent effects of associated factors on yearly SFR water consumption as well as the spatial heterogeneity and spatial dependence in these effects. The third objective seeks to identify the factors explaining the monthly water consumption per SFR household at the neighborhood level during the 2007-2009 drought within a spatial econometric modeling framework and evaluate the effects of pricing and water usage restrictions. We hypothesize that (1) weather factors would make a great contribution to local water use in Charlotte, (2) some household and housing characteristics such as household size, income, and lot size better explain monthly water consumption than the others (for example age factors and housing density), (3) price effect would be inelastic, and become smaller with the intervention of water usage restrictions which supposedly show negative effects.

Together these three objectives will contribute to the understanding of how SFR water use changes over time at a small temporal and spatial resolution and they approach the relationship of water use and its associated factors from different perspectives but within a spatio-temporal framework. These pieces of empirical work, built upon methodological advances in geographic information systems (GIS), spatial analysis, and statistical approaches, will collectively depict a relatively complete picture of the history and present of SFR water consumption in terms of its state, pattern and process. The evidences derived from the analyses on various issues from weather sensitivity, historical contingency, and relational assessment on water demand's determinants will serve our ultimate goal of deriving policy implications for water demand management.

1.2 Contribution and Significance

The case study of this research is vital to the enhancement of the local community's knowledge (in this case, Charlotte, NC). Public water utility, planners and decision makers will benefit from the research with regard to better understanding the complex urban dynamics of water consumption. The multidimensional analyses will not only offer first-hand evidence for answering critical questions on weather sensitivity and driving factors related to water use, but also help identify new research questions, formulate novel hypotheses, and introduce more opportunities for disentangling the problems in water management in a comprehensive manner.

In addition to the social values this study carries, the methodological highlight of our research is the development of the empirical spatial panel data models for water use research. The spatial econometric modeling approach is embraced to incorporate spatial dependence and heterogeneity in traditional panel data models. This framework not only helps reduce bias in coefficient estimates or standard errors from panel data models that do not account for spatial autocorrelation (LeSage and Pace 2009), but also is able to distinguish and quantify the direct and indirect effects of the household and housing factors on water use. The same specifications of the models will be tested against relatively small neighborhood unit - block groups.

1.3 Structure of the Dissertation

The dissertation is organized into seven sections. Chapter 2 provides a literature review of the determinants of water use or demand. Descriptions of the study area, data collection and processing are given in Chapter 3. Chapters 4, 5, and 6 will address the above three objectives respectively, and each of the chapters includes literature review on the topic in focus, methodology and results, and conclusion sections. The last chapter concludes the dissertation by an overall conclusion and limitations of the study.

CHAPTER 2: LITERATURE REVIEW

The widening gap between water availability and demand in urban environments has been calling for water management practices integrating both extensive (supply-oriented) and intensive (demand-centered) approaches. Governments and water providers who endeavor to satisfy the needs of their growing resident populations by bringing in more water from remote places have been challenged continuously by rising infrastructure costs, regional conflicts and competing uses in terms of water rights (Sewell, W. R. D. and Roueche 1974). The idea of managing urban water demands instead of expanding supply has become attractive to water authorities, due to its multiple (economic and ecological) benefits to the water supply system and the natural environment, and due to the improved feasibility supported by technological innovations, specialization and privatization in water services, and progress in novel practices (Dziegielewski 1999). The core objective of demand management is to make more efficient use of existing supplies. A better understanding of how water is used by whom and in what ways water savings can be realized will be necessary and useful for planning, implementing, and evaluating demand-side alternatives (Jorgensen et al. 2009). This chapter will start with a review of the factors determining residential water use/demand and the mechanisms linking them with household water use behaviors.

2.1 Determinants of Residential Water Demand

2.1.1 A Behavioral Framework

Quality water (mainly from water pipes for urban residents) is essential to human being's daily life. It serves our needs ranging from the basic ones such as drinking, cooking and cleaning, to the life-style-induced ones like toilet flushing, laundry and bathing, and to the discretionary uses of water for lawns, landscaping/gardening and pool. The quantity of water used by a household is a direct result of the aggregation of individual consumption behaviors that are conducted within and hence constrained by their environment (internal and external to household and/or its individuals).

Behavioral and psychological research has offered some insights into the causes of behavior related to water use and conservation and their interrelations. Russell and Fielding (2010) categorized the causal factors of water conservation behaviors into five groups, including attitudinal factors, beliefs, habits or routines, personal capabilities, and contextual forces. We organized these factor groups and key factor elements, their relations to household water consumption, and their interrelations into a conceptual model (Figure 1).

In this conceptual model, rectangles in different colors represent different categories of factors and their key elements. Relations of attitude, belief and habit factor groups to water use behavior are symbolized by wide gray arrows. The factors from the capability and context categories encompass water consumption, implying their direct associations. Factor groups and/or their elements can interact, as indicated by various arrows. Solid lines with a single arrow represent direct impacts, while dashed arrows refer to mediation paths. For individual factors influencing each other (for example, price and non-price policies), double-arrows symbolize bi-directional impacts. The dotted line

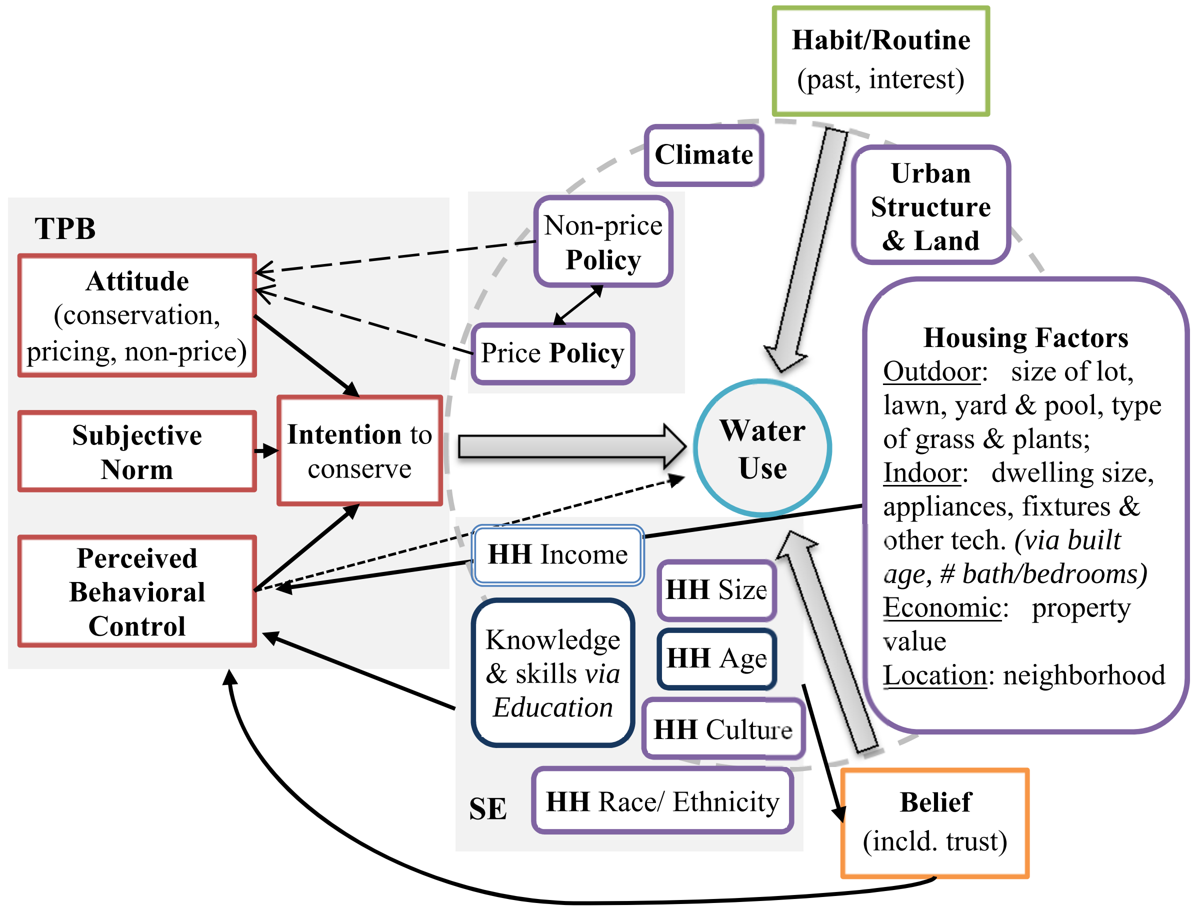


Figure 1: Household water consumption model from a behavioral perspective

Red: attitude, Orange: belief, Green: habit, Blue: capability, Purple: context; Rectangle with double outlines – Capability or context
Solid arrow: direct impact between factor groups and/or factor elements; Dashed arrow: indirect impact
(References: Russell and Fielding (2010) and Jorgensen et al. (2009))

with arrow represents non-substantive direct effects. The specific details about this model will be elucidated as follows.

The theory of planned behavior (TPB) (Ajzen 1999) suggests that a person's deliberate behavior is determined by his/her intention (i.e., a motivation or plan) to perform the behavior, and that attitude, subjective norms and perceived behavioral control lead to the formation of a behavioral intention. The three constructs (attitude, subjective norms and perceived behavioral control) are classified into one factor group named attitudes. Social norms denote the customary codes of behavior in a group or people or larger cultural context, and perceived behavioral control refers to a person's perception of the ease or difficulty of performing the behavior of interest. As a general rule, the more favorable the attitude and subjective norm, and the greater the perceived control, the stronger should be the person's intention to perform the behavior in question. Take the water conservation behavior as an example. If people have a positive attitude toward water conservation, if they perceive that important others in their life think that it is a good thing, and if they think that it is something they can easily do, then they will intend to engage in water conservation and their intentions should in turn translate into water conservation actions (Russell and Fielding 2010). The TPB research has important implications for policy-level water conservation efforts. Any conservation campaign or program merely providing information may not help change behavior very much. Instead, water conservation efforts should aim at influencing attitudes toward the behaviors of consuming and saving water, seek to gain widespread support in the community, and offer strategies to ensure people easily engage in water conservation behaviors.

Similarly, beliefs, often identified as significant drivers in the environmental psychology literature, are supposed to play an important role in predicting water (Russell and Fielding 2010). According to the TPB, there are three kinds of beliefs corresponding to the three attitudinal factors (attitude, subjective norms and perceived behavioral control). Beliefs may be conceived as a precursor to them (Eagly and Chaiken 1993; Ajzen and Fishbein 2000). When water conservation is concerned, water-specific beliefs such as viewing water as a scarce resource are regarded as the most immediate drivers, in contrast to generalized worldviews about the environment (Corral-Verdugo et al. 2003).

Recognizing that behaviors may not always be deliberative or rational, researchers (Aitken et al. 1994; Gregory and Di Leo 2003; Trumbo and O'Keefe 2005) have also examined the effect of habits on water use behaviors, and emphasized their importance for the design of policies and intervention strategies aimed at changing behaviors (Russell and Fielding 2010). Although a number of habit measures such as the number of showers and baths, clothes washing loads, and dish loads per week and the past water consumption and conservation behaviors were included in past studies in the domain of environmental psychology, habits relating to outdoor water use have not been assessed (Kenney et al. 2008).

All three categories of factors mentioned so far can explicitly characterize water use/saving behaviors. However, data or information on most of these factors are not readily available, and usually are collected by household surveys and interviews designed specifically for the purpose of attitudinal or behavioral analysis.

The determinants of water consumption commonly identified in the bodies of literature outside psychology and cognitive science largely belong to the causal categories of personal capabilities and contextual factors. Personal capabilities refer to knowledge and skills, availability of time, literacy, money, social status and power, while contextual factors can involve many different considerations including household characteristics and experiences, physical infrastructure, technical facilities and products in a household, pricing and non-pricing factors as well as other features of broader social, economic and political contexts (Stern 2000; Steg and Vlek 2009; Russell and Fielding 2010). A number of sociodemographic variables fit both categories, playing the roles of being proxies for personal capabilities as well as being contextual factors per se (income is a typical example). Since behaviors could be facilitated or constrained by personal capabilities and contextual factors, it is important to consider their impacts on water consumption. Next, we will review some popular determinants from these two groups with an emphasis on their connections with household intentions and behaviors toward residential water consumption.

2.1.2 Water Pricing and Rate Structure

North American economists first became interested in best price mechanisms to regulate water in the late 1960s and through the 1970s (Corbella and Pujol 2009), when water demand management started to take shape. The increasing applications of new rate and pricing structures during the later 20th century have provided many opportunities for water demand research (Kenney et al. 2008). Numerous econometric models for water use were developed to quantify price elasticity (the economic measure of how demand for water moves in response to price changes) and evaluate various pricing policies from flat rate to uniform, increasing or decreasing block (Arbués et al. 2003). These studies benefited the practices of urban water utilities mainly in forecasting water demand and performing benefit-cost analysis of demand management alternatives (Dziegielewski 1999).

Water pricing is regarded an important economic driver for water consumption because water as a commodity has its economic and market value, thus the price theory (that price can determine the quantity demanded of goods and services) would apply. Dozens of empirical pieces of research on residential water demand confirmed the negative effect of price, meaning that household water demand decreases when price increases. However, the estimates of the price elasticity vary widely (more lying between 0.25 and 0.75), and the general conclusion is that demand is largely price inelastic (less than 1) (Arbués et al. 2003; Kenney et al. 2008; Worthington and Hoffman 2008). There are different reasons to account for water price inelasticity. From an economic perspective, the price of water does not reflect its economic value since water is a public good and public goods lack substitutes for basic uses, are usually offered cheaply to urban residents; there also are other issues such as fairness and equity for water authorities to concern when setting water price. From a behavioral perspective, price plays the role of contextual factors, and its interactions with the other contextual factors (such as income, the number of occupants in household), with personality capabilities (income, typically knowledge about pricing or price change, understanding of the rate structure and water bill), and with psychological variables (e.g. the change in attitudes, intentions and beliefs during drought/water scarcity) complicate its net effect on water consumption (Jorgensen et al. 2009). The complexity in block rate structure (the increasing block is commonly employed by American public water suppliers) makes it difficult to separate the effect of marginal prices within each block of the pricing from the effect of the rate structure per se due to the increasing nature of the tiers (Olmstead et al. 2003; Carter and Milon J. Walter 2005). From the policy perspective, price elasticities can be influenced by the existence of non-price policies (such as water usage restrictions and other conservation programs). Moreover, price elasticity can vary significantly among seasons, uses, regions, and various social/economic conditions (Kenney et al. 2008).

Given the relatively low price elasticity, two different viewpoints hold with regard to the use of the pricing tool to reduce demand. One is that prices can play a crucial role in demand management as long as urban water prices reflect marginal costs and the elasticities are different from zero (Arbués et al. 2003). The opposite viewpoint is that price is an ineffective tool and other mechanisms would be more appropriate (Worthington and Hoffman 2008). For any specific local water authority, the choice of a position should be based upon a comprehensive understanding of the local pricing effect and its interrelations with other factors in determining water consumption.

Among the literature on estimating residential water demand using econometric techniques, price specification has been a focal issue with great controversy. Two mainstream specifications and their variations can be identified. We will briefly describe the two major price measures (marginal price and average price) and their disadvantages/problems here. For details, we suggest a reference to the two review papers by Arbués et al. (2003) and Worthington and Hoffman (2008) and the seminal papers on electricity demand that introduced the two original specifications.

Marginal price is defined as the price a customer paid for the last quantity he/she consumed (Taylor 1975). Marginal price is either a constant value for a flat rate (fixed fee no matter how much quantity) or a uniform water tariff (rate is quantity-independent). Given a non-uniform rate structure, marginal price varies by particular quantities. More volumes of water consumption are usually associated with a higher marginal price in an increasing block structure, vice versa for a decreasing pricing. Average price is a derived measure for price, simply equivalent to the division of total charge paid and the corresponding water quantity consumed. In comparison, marginal and average price are the same for simpler pricing structures, and marginal price is larger or smaller than average price for increasing and decreasing blocks respectively. In the price specification using marginal price, a difference variable was suggested to be included by Nordin (1976) to account for the income effect imposed by the block structure. Difference price is calculated as the difference between the total bill and what the user would have paid if all units were charged at the marginal price (Arbués et al. 2003).

The marginal price specification is theorized under the assumption that perfectly-informed consumer should react to marginal price and rate premium. However, the reality is that most consumers lack of knowledge about the pricing structure or intramarginal rates (difference in rates between blocks), thus the Nordin specification is contentious. Although average price has been widely accepted as an alternative measure of price, it is less efficient since marginal price is implicitly captured by this variable, and requires an appropriate translation for the purpose of pricing design. Some empirical research (Gibbs 1978; Dalhuisen et al. 2003; Schleich and Hillenbrand 2009) showed that price elasticity tends to be overestimated using average price rather than marginal price. Nieswiadomy and his co-workers used a price perception model (proposed by Shin (1985) to account for information imperfection in water use decision) to test customers' reactions to marginal, average and perceived price (Nieswiadomy and Molina 1989; Nieswiadomy 1992; Nieswiadomy and Cobb 1993), and their findings are not conclusive due to different case studies and data. The results from their panel model showed that customers react to marginal prices when facing increasing block rates and average prices when faced with decreasing block rates (Arbués et al. 2003), while the effect of average price is stronger for both decreasing and increasing structure as indicated in their models for the national cross-sectional data (Worthington and Hoffman 2008). To deal with the endogeneity [1] (sometimes called simultaneity) problem involving both specifications under non-uniform structures, advanced statistical methods have been employed including instrumental variables (IV) approach (e.g. two-stage or three-stage least square models (2SLS or 3SLS)) and simultaneous equation method (Arbués et al. 2003). Although researchers may prefer one specification to another, it is common to apply both average and marginal price in the same model and compare the parameter estimates, especially if price elasticity is of central interest.

In the earlier water demand models dominated by an economic perspective, only a few kinds of variables other than price were considered as control variables (maybe due to data availability and computing and software constraints), chief among them were climate/weather and income variables.

2.1.3 Weather Factors

Weather is the most dynamic factor (changeable at every moment) that impacts water consumption but is beyond the control of water authorities and households. Based on daily life experiences, people could come up simple and intuitive rules describing the

[1] Prices are endogenously determined by quantity demanded.

responses of their water use behaviors to weather condition and change. For example, water demands are higher in hot-dry weather than cold-wet conditions (especially in regions with distinct summer and winter seasons), because additional water is used for more frequent showering, bathing and laundering, irrigating vegetation (grass, tree, and other plants) and pursuing entertainment in pools and through fountains, etc. That great outdoor water use reduction occurs in the summers with abundant rainfall is another example. However, to fully understand the influences of climate factors on water demand is challenging for several reasons.

First, there are multiple weather proxies that potentially can explain water consumption patterns. Precipitation/rainfall, temperature, and evapotranspiration (ET) are commonly studied, and the underlying mechanism of how each of them is linked to water use differs. The values of these factors are recorded at a small temporal scale (daily and even hourly) and of ratio type, so we can characterize weather using their descriptive statistics. For example, average, maximum, minimum temperature and their changes have been included in the empirical models (Maidment and Miaou 1986; Griffin and Chang 1990; Stevens et al. 1992; Agthe and Billings 1997; Pint 1999; Martinez-Espiñeira 2002, to list a few), yielding different conclusions about temperature effects. Second, the assumption about the linear effect of the weather variables may not hold, and some empirical evidence revealed the diminishing nature of weather effects. In a case study in Australia, (Gato et al. 2007) found that, above a certain level, more precipitation would not reduce water use, and similarly there exists a threshold for temperature above which its effects are greatest. However, it lacks of easy criteria or guidelines for identifying these thresholds. Third, individual sensitivity to weather can be influenced by personal psychology. People are more aware if there is rain or not and for how long rather than the exact quantity of rainfall they receive, probably because the former is based on perceived information that is more easily available than the inquiry of actual information (Zhou et al. 2000; House-Peters and Chang 2011). Hence, the number of rainy or rainless or abnormally hot days (frequency and time between weather events) could be a better explanatory variable (Maidment and Miaou 1986; Martinez-Espiñeira 2002). Fourth, the complication of weather effects can be caused by the interactions between weather condition and personal knowledge or contextual factors related to a household. For example, the existence of rainfall tank or evaporative coolers will help reduce the use of tap water for irrigation; different types of grasses for lawn demand different amounts of water at different times; when abnormal weather conditions occur or persist (such as droughts), the use of pricing and non-price management tools could interfere with households' normal response to weather conditions (Kenney et al. 2008). It is also found that having specific knowledge about climate change impacts would encourage water conservation behavior (Clark and Finley 2007; Russell and Fielding 2010). Lastly, climate data sampled at the limited number of weather stations may not represent microclimates at the neighborhood level; household-level consumption data are usually available at best at a monthly scale while weather conditions change daily (Kenney et al. 2008). Thus temporal and spatial mismatch between data for water consumption and household may induce noise in modeling.

Despite the lack of preferred weather variables being suggested and the remaining complications, the intimate association between weather factors and water consumption is undeniable. More empirical research on this topic from multiple perspectives (climatologic, psychological and geographical) will be needed, given the challenge and uncertainty that climate change may impose on water consumption.

2.1.4 Sociodemographic Factors

A range of sociodemographic variables of households have been investigated in the residential water demand literature, probably due to the ease in collecting such data either from standard questionnaires on socio-demographics in survey studies or from census databases. Although all of them play the role of facilitator or constraints, their effects on water demand are not equally important, and sometimes the same single factor is found to impact water use in different ways.

2.1.4.1 Income

Similar to price, income elasticity of water demand has been widely studied with econometric models. Because income is believed to greatly affect the responsiveness to price mechanism (Corbella and Pujol 2009), it is essential to estimate the elasticities of both price and income for designing better pricing regimes (Worthington and Hoffman, 2008). Research looking at the responses of different household income groups to water pricing showed that both low and high income households may not respond to price for different reasons. The increase in price would not discourage low income families' basic water needs, while for well-off households the magnitude of price change is not large enough to put pressure on their disposable income and curb their consumption (Corbella and Pujol 2009).

From a behavioral perspective, household income reflects financial (in)capability for water spending. The positive relation between income and water consumption is consistently supported by various studies with different objectives and model specifications. The phenomenon that higher income households generally use more water was often explained by higher living standard (e.g. for hygiene and cleaning) and higher probability of the presence of water demanding outdoor uses such as lawn gardens, swimming pools, and even fountains (Cole 2004; Harlan et al. 2009). Although households with a higher income level can afford and may have intentions to install water efficient appliances and infrastructure (dual-flush toilet, front-load washing machine, garden sprinkle system, etc.) (Lam 2006), the amount of water supposed to be saved via these means is possibly offset by larger water consumption demanded by luxury indoor (e.g. spa) and outdoor activities (Ouyang et al. 2014). Therefore, income elasticity is thought to be larger in hot-dry seasons when outdoor water use is desired. Moreover, within affluent individuals and households, the group with higher education may have more knowledge about the environment and water saving products, and thus will exhibit greater conservation awareness and intentions (Russell and Fielding 2010).

As for the estimates of income elasticity, the literature almost universally reported low values (less than 1 in magnitude), and Worthington and Hoffman (2008) summarized possible explanations: income elasticity of residential water demand is actually low, sample or specification bias may take effect, increasing and decreasing block structures hypothetically capture income effects, or income may exhibit its effects over the longer term rather than in the short run.

2.1.4.2 Household Composition and Characteristics

Household composition is typically described by the number of occupants (household size) and their age and sex compositions in a household. Sex has not been recognized as an important determinant of water use in the American literature.

Household size and people's age are believed relevant to residential water consumption, but their effects are not simply linear.

Household size is a variant factor, especially for the western world in which a second demographic transition has been undergoing for decades. The trend of decreasing household size and increasing number of households in our modern society definitely has strong implications on water use patterns (Corbella and Pujol 2009), in spite of its various influences. In principle, more water is needed for larger household, however, economies of scale possibly lead to the negative association between household size and water demand when water use per capita is considered instead of aggregated demand (Höglund 1999; Arbués et al. 2003; Kenney et al. 2008; Worthington and Hoffman 2008). Arbués et al. (Arbués et al. 2004) argued that the effect of economies of scale becomes diminished when an optimum household size is reached, implying the influence of household size on residential water consumption per capita could be positive. Gilg and Barr (2006) provided psychological reasoning for this - it is more difficult for large households to establish conservation norms or have similar consumption behaviors, while households with fewer residents were more often categorized as committed environmentalists and thereby more likely to enact water conservation behaviors (Russell and Fielding 2010). Few researches have examined the interference of household size with price effect. It is found that smaller households in Zaragoza, Spain are more sensitive to price changes (Arbués et al. 2010).

As a factor that is part of capability, the age effect is relatively complicated. Retired people (often older) may consume more water due to more of their time spent at home and on gardening (Lyman 1992), or less water because they are more likely be thriftier

(Nauges and Thomas 2000; Shove 2003). More water could be expected in the families with children and teenagers, since their frequency of doing laundry, showing, and playing with water outdoor and indoor is higher (Hurd 2006; Balling et al. 2008) and children and younger persons are less careful when using water. The major problem with the studies addressing the age effect is that information on the age distribution of all the residents in a household is hard to obtain. Thus either the age of the household head or the census variables on percentage of the population over 64 years or under 19 years, the number of dependents per household, percentage of the household with own children under 18 years are used as proxies (Worthington and Hoffman 2008).

Regarded as a proxy for knowledge capability, the factor education is more often encountered in the literature on water conservation behaviors. It is assumed that people with higher levels of education may have greater awareness of water scarcity and environmental consciousness, which could be translated into water conservation actions such as purchasing water efficient appliances and planting drought-tolerant garden species (Geller et al. 1983; Gilg and Barr 2006). However, the findings from the behavioral studies showed contrasting results (Russell and Fielding 2010). Although the association between higher education and deeper conservation commitment has been supported (Gilg and Barr 2006; Lam 2006), some research has shown that less educated households demonstrate higher water conservation intentions and more water conservation behaviors (Gregory and Di Leo 2003; Clark and Finley 2007). The questions of why the conservation intentions of better educated people were not turned into actual water saving behaviors and whether other factors such as income confounded the influence of education on water conservation for lower educated households remain unanswered.

Race and nationality/cultural background composition of the household are occasionally discussed in water demand estimation. Griffin and Chang (1990) and Gaudin et al. (Gaudin et al. 2001) specified the percentage of the population of Hispanic origin as a determinant of water consumption in their Texas case study. It is argued that immigrant households may retain the habits of water usage from their country of origin where water is often a scarce resource and where there more consciousness of water saving (Pfeffer and Stycos 2002; Smith and Ali 2006).

2.1.5 Housing Characteristics

Physical features of houses can impose constraints on household water consumption. For example, the households with larger lawns are more likely to use more water for irrigation than the households with the same characteristics but smaller lawns. Other housing attributes similar to the size of lawn are lot size, yard size, pool size, dwelling size or physical size, and they are expected to positively influence water use. The information on the appliances, fixtures (e.g. faucets) and other water-using technologies featured in a house is supposed to be relevant to water consumption behavior. When such detailed information is not available, the number of bathrooms and the number of bedrooms are examined as proxies. These two variables can also substitute for household size, since they are recorded in typical cadastral databases which can be available to the public upon request. Another alternative to underscore the water efficiency potential of water fixture is age of house. The reasoning is that newer homes built after 1992 have more efficient water fixtures installed in order to comply with national minimum efficiency standards for all toilets, showers, urinals and faucets manufactured according to the U.S. Energy Policy Act of 1992(Harlan et al. 2009). Change et al. (2010) found a negative association between the age of building and water consumption and argued that it could because older houses are smaller in terms of physical and lot size and most of them were remodeled with more water-efficient fixtures and appliances.

Sometimes assessed property value is introduced in addition to income into water demand models because it may imply household preferences for home lifestyle and social status (Arbués et al. 2003), and the households living in an expensive house intend to maintain their property carefully (including irrigating lawns more frequently) to display status and convey social distinction (Askew and McGuirk 2004; Domene et al. 2005; Breyer 2014).

Sometimes the features of a lot such as landscaping type (e.g. grass and vegetation types), garden condition, and presence of a pool have been considered to scrutinize their effects on outdoor water use (Syme et al. 2004; Wentz and Gober 2007; Harlan et al. 2009; Ouyang et al. 2014). The type of dwelling (e.g., single family house vs. multi-family apartment) has occasionally been included when the study subject is urban or residential water demand in general (Mylopoulos et al. 2004; Domene and Saurí 2006; Hoffmann et al. 2006). Information about home ownership (owned vs. rented) will be useful in explaining some heterogeneity of households' water consumption and conservation behaviors. Hoffmann et al. (2006) found that residents of owner occupied houses have higher price and income elasticity of water demand than those living in rented dwellings. From the psychological perspective, home owners may be more likely to engage in efficiency behaviors, compared to tenants because tenants have less control over the installation of water efficient fixtures and appliances in their rented house and also do not necessarily receive a water bill as it is often a hidden cost of the rent (Randolph and Troy 2008; Russell and Fielding 2010). Another variable characterizing the social feature of a house is use of home (for vacation/seasonal vs. all-year-around residence), and is assumed helpful in identifying those communities where seasonal use can have a greater impact (Arbués et al. 2003).

A common concern with quantifying the influences of these housing characteristics is their correlations with household features, particularly income (Corbella and Pujol 2009). For several variables such as landscape type, pool size, ownership etc., the data sources are still limited.

2.1.6 Urban Structure and Land Use Patterns

There has been a growing interest in understanding the role of urban spatial structure in urban water management. The availability of Geographical Information System (GIS), spatial analysis techniques, and reliable spatial data for water consumption and land cover/use at fine spatial scales offer new opportunities for researchers to examine the spatial complexity of urban water consumption at the neighborhood scale as well as the spatial association between urban structure/land use pattern and residential water use (House-Peters and Chang 2011).

Several studies analyzed water use distribution across space and found significant clustered patterns of high/low water uses at multiple spatial scales (county, census tract, block group) (Balling et al. 2008; Franczyk and Chang 2008; House-Peters et al. 2010; Polebitski and Palmer 2010; Breyer et al. 2012; Ouyang et al. 2014; Gage and Cooper

2015). It is observed that the spatial concentration of water use is partially coincident with the land use patterns in a municipal area. For example, single family residential water use in Portland, Oregon tended to be lower and less weather-sensitive in older neighborhoods near the city center, which are featured with higher building densities/smaller average lot size (Chang et al. 2010; House-Peters et al. 2010; Breyer et al. 2012). In contrast, the high water users concentrated in more affluent suburban neighborhoods, and these neighborhoods typically have more elaborate water-intensive landscapes (e.g. grass lawns) and sometimes have swimming pools (Chang et al. 2010). Wentz and Gober (2007) reported similar findings in their study of Phoenix. This is not hard to interpret since both water consumption and land use/development patterns are the consequences of people's preferences and decisions on house and water use to a large extent. We have known that housing characteristics such as dwelling type and size, lot size, and landscape type are important determinants of water consumption, and these features are basic elements composing land use pattern which is one spatial representation of urban structure. There is no doubt that people who choose to live in places with large houses and larger lots are more likely to use more water, and those places are largely distributed outside urban centers.

Nevertheless, urban planning and design have the power to regulate land use type and development density and guide the formation of urban structure. After examining the influence of urban zoning (i.e., single-family residential or commercial), total building area, and the density of single-family residential developments on water consumption in Portland during the period 1999-2005, Shandas and Parandvash (2010) confirmed the significant relationship between land use and water consumption and elucidated the possibility to improve the effectiveness of water-conservation activities through the lens of urban planning. They also suggested planners to develop predictive models for assessing water demand given alternative scenarios of urban development. In line with this suggestion, Polebitski et al. (2011) coupled their statistical water demand model with an urban simulation model (UrbanSim) to evaluate regional water needs of the residents living in the Puget Sound region under different planning scenarios accounting for climate change and transitions. Embracing a similar idea, a couple of international researchers in Australia (Urich et al. 2011) and the Netherlands (Sanchez et al. 2011) have been working on the integration of urban environment dynamics and water supply systems for better water management. Fox et al. (2009) developed a methodology for statistically forecasting the amount of water that a new residential development would demand based on three housing characteristics, number of bedrooms, architectural type (i.e., detached or semidetached), and presence of a garden. Gage and Copper (2015) demonstrates the value of high resolution land cover and Lidar-derived vertical structure data for understanding urban water use patterns (especially outdoor) and emphasized their potentials in helping target water conservation efforts. The studies investigating the relationship of landscape types and water use can provide us insights into how urban design tools can facilitate water reduction.

All the existing research efforts concerning the influences of urban form and land use structures on water demand patterns highlight the need for coordinating urban land use planning and water management as well as the possibilities of utilizing land use planning and urban design as a potentially robust and equitable non-price mechanism to develop spatially explicit water conservation strategies (Chang et al. 2010; Breyer 2014).

However, the studies incorporating factors that represent urban structure and that can be influenced by planning policies are still limited, particularly in the literature from econometrics and psychology. More collaborative research across the disciplines such as water management and planning, urban planning and design, geography, and behavioral sciences are required before we translate significant and valuable findings into the practice of water management.

2.1.7 Non-Price Policies

Non-price policies are usually classified into three broad categories: informational strategies (such as public information campaigns, rationing, and increased billing frequency), technological change (such as installation of low-flow appliances) and quantitative restrictions (restrictions and bans on the quantity and timing of outdoor water use during periods of peak demand). Water utilities commonly implemented quantitative restrictions as short-run tools within their service territories/municipalities to limit the total water consumption and maintain the level of water source during droughts. A few studies have evaluated non-price policies and found the statistically significant effect (Renwick and Archibald 1998; Renwick and Green 2000; Timmins 2003; Olmstead et al. 2007; Ramachandran and Johnston 2011).

In summary, we summarized the key determinants widely examined in the literature of modeling residential water demand. They can be grouped into the broad categories including belief and attitudinal factors, price and non-price policies, climate factors, household sociodemographics, housing characteristics, and urban structure/land uses. The mechanisms explaining how they are correlated with water consumption were emphasized in this review. The combinations of these factors in the literature vary greatly, depending on research objective, the temporal and spatial scale of interest, data availability, and local contexts. For this dissertation, we compiled datasets covering some of the factors falling within the categories of capabilities and contextual factors (see the factors within the dashed circle of the conceptual model in Figure 1). There is no data readily available for understanding the effects of beliefs, attitudes and habits on water consumption behaviors. Next, we move to the introduction of the study area and datasets.

CHAPTER 3: STUDY AREA AND DATA

3.1 Study Area

3.1.1 County Facts and Water Use Profile

Mecklenburg County is located in the State of North Carolina, and it contains seven municipalities including the City of Charlotte and six towns of Huntersville, Davidson, Cornelius, Pineville, Matthews, and Mint Hill. According to the 2010 census, the total population in the county was 919,628 people, 79.5% of which live in Charlotte. In 2010, the city's population density was 2,457 people per square mile, which is about the average density for an urban area of the South census region (Cox 2014), while the county's density is relatively lower, only 1650 people per square mile. The county/city is still ranked as the most populated and densely populated county/incorporated area in the state of North Carolina. Also, the county/city has been growing fast in the past two decades (for the county, 35.2% in the 2000s and 36.6% in the 1990s; for the city, 32.2% and 36.0%, respectively). According to a report prepared by UNC Charlotte Urban Institute (2011) using Census data, the Charlotte Metro Area (consisting of City of Charlotte, two neighboring cities Concord and Gastonia) is ranked as the fourth fastest growing large urban region in the U.S., and notably it is the only Metropolitan Statistical Area (MSA) within the top ten fastest-growing MSAs during 2000-2010 that keeps an increasing growth pace. Other indicators for economy, society, transportation, and so on

(Charlotte Chamber 2011) depict the City of Charlotte (as well as Mecklenburg County) as a growing, diverse, more progressive (than before) and promising community. The population density and growing trend will undoubtedly add pressure to Charlotte's future water supply.

Charlotte Water (formerly Charlotte-Mecklenburg Utilities Department (CMUD)), established in 1972, is the largest public utility in the Carolinas today (Utilities 2011). It is the major water supplier for the Charlotte-Mecklenburg community, and also sells water to its neighboring counties and municipalities in North and South Carolina. In total, there are 267,664 active water service accounts being serviced and over 4,232 miles of pipe (water mains) being maintained by Charlotte Water today (Charlotte Water 2015).

As in most urbanized areas, a centralized water supply system was built in Charlotte gradually over a long time period to provide water to its urban population. Currently the system withdraws daily on average 103 Million Gallon per Day (MGD) of surface water from two reservoirs (Lake Norman and Mt. Island Lake), and water is treated at three water treatment plants (Franklin, Lee S. Dukes, and Vest WTPs) (Charlotte Water 2016a). The total capacity of the three WTPs is 242 MGD, and according to their historical records, the maximum daily water pumped over the past years was about 169 MGD and occurred in August, 2007, which was still relatively below the amount of water that can be treated every day by the existing plants. The two reservoirs store 163 MGD water for daily use, and an additional water supply of 25 MGD is supposed to be available in 2020. Thus it is estimated that the community can be supplied with around 188 MGD of water in the future. Assuming the per capita residential demand remains constant (5,236 gallons per month per home) and the household size is 2.5, based on the population projection from the Mecklenburg-Union Metropolitan Planning Organization (MUMPO) study, it is forecasted that the community may not face water shortage problem (demand greater than supply) until 2060 (Division of Water Resources 2014).

In terms of metered daily average water use (Figure 2), residential use (around 58 MGD in 2010) ranks first and is more than twice as much as commercial use and surpasses the sum of all the other types of water use. Assuming the service population were the entire census population in 2010 (788,000 people), the metered consumption would be 73.65 gallons per capita per day (gpcpd). Compared to the amount of water for fundamental and standard human requirements, 36 and 50 liters per capita per day (lpcpd) (or around 9.51 and 13.21 gpcpd) (Gleick 1996), the water consumption standard of the residents living in the Charlotte-Mecklenburg community is relatively high. In the past five years, the per capita residential consumption has been decreasing from 74 to 69 gallons, however.

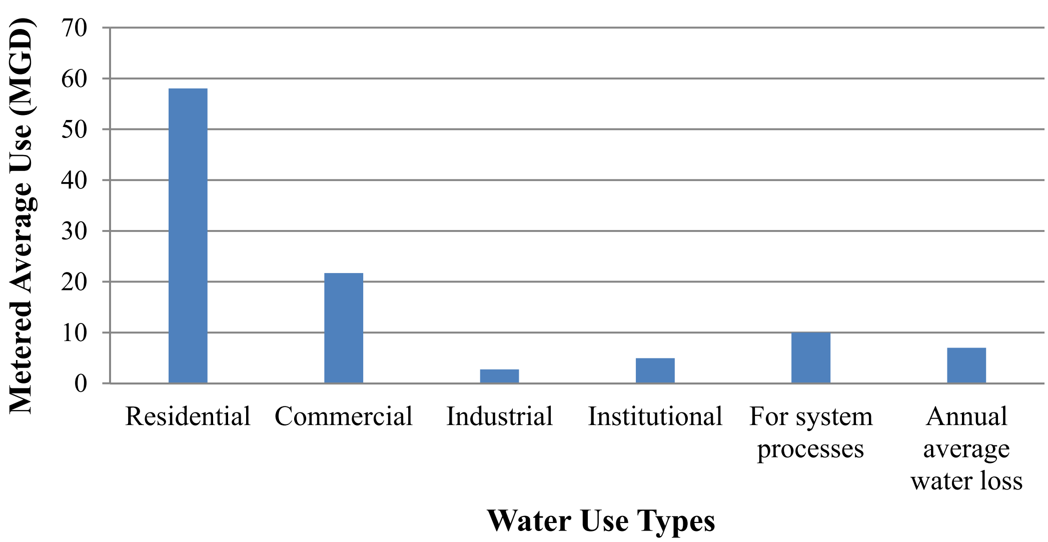


Figure 2: Metered daily average water usage by types in 2010

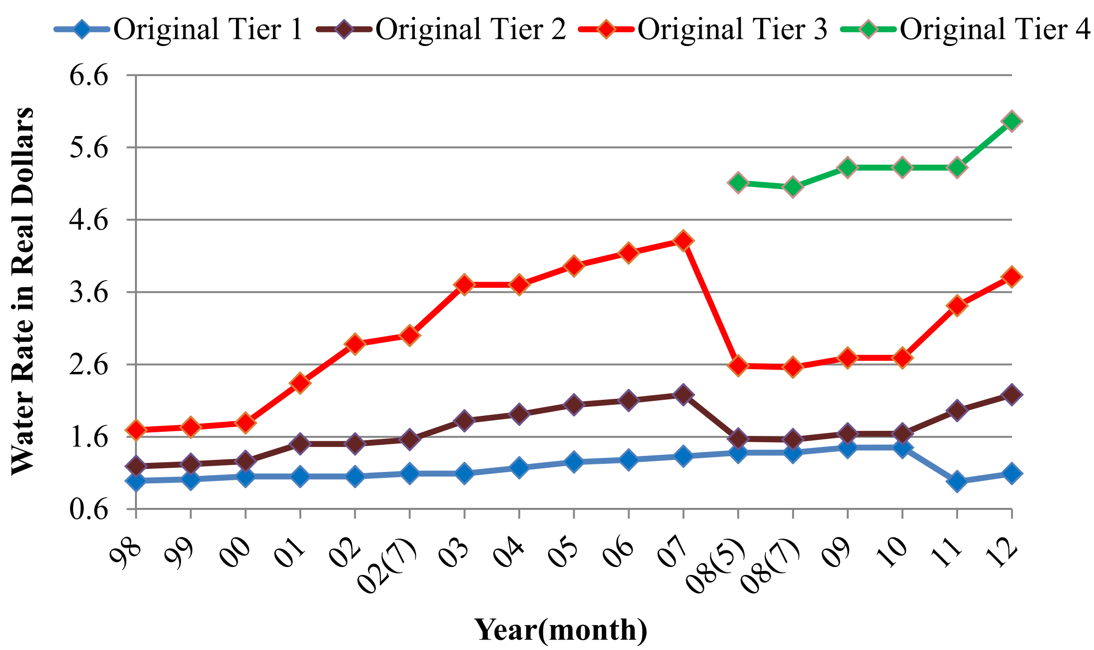
Source: North Carolina Division of Water Resources

3.1.2 Pricing

Charlotte Water is a community-owned enterprise fund of the City of Charlotte, not operating on a for-profit basis. Use fees support the full cost to run and maintain water and sanitary sewer operation. The residential water bill structure is relatively complicated, consisting of flat monthly charges (fixed fees and availability fees (effective after July 1, 2013) and usage charges for water and sewer together with storm water fees. The sewer rate by usage is of uniform structure, while the rate structure for water has been increasing block since 1998. The sewer charge is based on metered water usage each month with a sewer cap, meaning sewer fees are only charged by up to a maximum amount of water usage.

Charlotte Water classifies its customers' service accounts by rate type (17 in total). For example, residential water use involves three types, multi-family (WA), single family detached (WAD) and residential (WR). The rate structures for the WAD and WR rate types are exactly the same. For the WA rate type, the marginal prices are same as the ones for WR/WAD, but the threshold levels of the tiers are different for the years before May 1st, 2008. Actually, the tiered usages for the WA type remained same over time (10 and 12 CCFs for the first and second tiers). Next, we mainly discuss the changes in pricing for the WR/WAD types.

Since 1998, water and sewer pricing have been adjusted every fiscal year. For most of those years, at least one of the tier rates of water increased (Figure 3), meaning water rate structure has been changing continuously. The most significant change is that a four-tier rate structure replaced the existing three-tier one in May 2008, when the most severe drought in history was ongoing. The general longitudinal trends of the 3-tier structures are 1) the lowest tier rate grew at the slowest speed (except for a faster drop in 2011); 2) the highest usage charges were increased by a much higher percentage, and experienced a major boost (slightly over 100%) between 2000 and 2003 when the first severe drought in the twenty-first century occurred. For the 4-tier rate structures, the increments were small for the years from 2008 to 2010. A moderate drought hit Charlotte in 2011. After that drought the third tier had a sharp growth and the second tier also increased faster than before, while the lowest tier rate had a drop. The tier with the highest rate remained relatively stable but started to grow after 2011. Note that when the real dollar rates were adjusted by the annual Customer Price Index (CPI), only the trend line for the lowest tier became virtually flat with a subtle decrease at a few time points, the rising trend in the other tiers was still evident though.



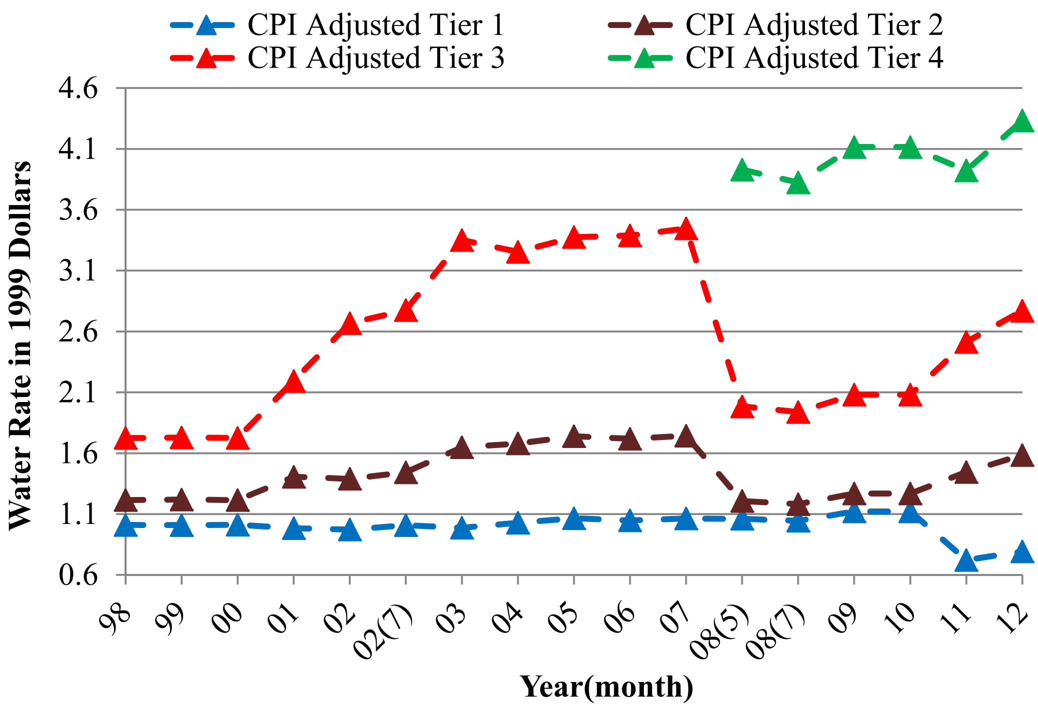


Figure 3: Primary rate structure for water

(Upper: in real dollars; Lower: in 1999 dollars)

The value of rates has to be combined with the quantity of water usage defined in the corresponding tier (termed usage threshold) to reveal the full picture of water pricing. The usage thresholds were less frequently changed over time but generally decreased in all the tiers (Figure 4). Such decreases indicate that the bar defining higher water user is lowered. In other words, more customers would be paying more, assuming the same amount of water was consumed and the rates for each tier were same. In this sense, local water pricing has been encouraging water conservation behavior, and at the same time gaining more revenues from the enlargement of customer base with relatively high water use. Another possible reason is that the overall population may have been reducing their water consumption, thus the adjustment to usage levels would be made to accommodate such a trend.

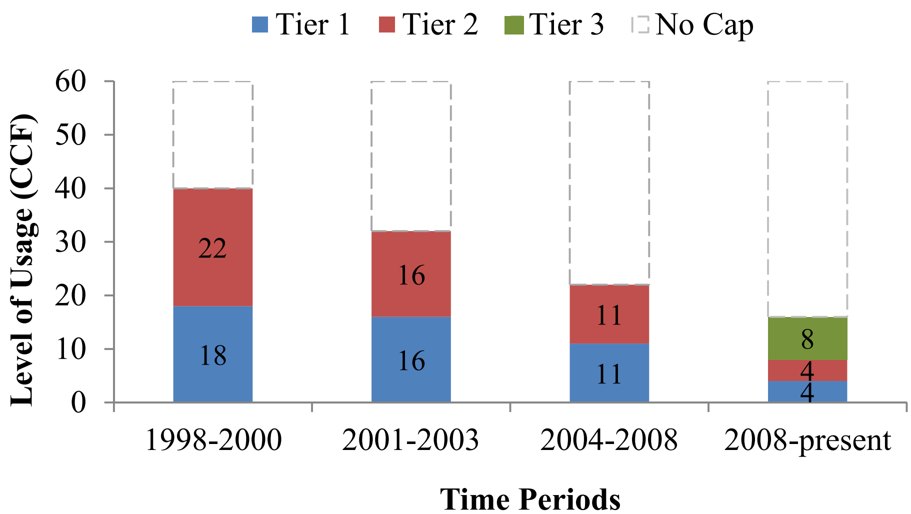


Figure 4: Usage tiers defined for water rate structure from 1998 to present

We also observe that consumption ranges are shifted downward more with the larger increases in rates in the 4-tier structure. This is due to the huge revenue reduction resulting from water usage restrictions implemented and the residents' cooperative efforts of conserving and protecting the water supply during the severe drought in 2007 and 2008. Such a pricing policy change was intended to recover and ensure the revenues in future years. It also split the lower water user groups into two, signaling Utilities' appreciation for aggressive conservation practices.

3.1.3 Historical Droughts and Local Water Management in Practice

While close attention has been paid to water shortages in the Western areas with arid conditions for some time, shortages are an emerging phenomenon in the Southeast of the United States (Henderson 2015). In fact, droughts are not new to the Carolina region. By the early 1990s, North Carolina had recorded seven major droughts occurring in 1925-29, 1930-35, 1950-57, 1965-1971, 1980-82, and 1985-88 (Weaver 2005). It was not until 2002 that threats of water scarcity became more pronounced. Attributed to several years of below-normal precipitation since 1998, river and reservoir levels became critically low in the summer 2002. The 2002 drought severely impacted water systems and numerous water users across the state (Lackstrom et al. 2011). Mandatory water restrictions, which usually impose rules for lawn watering, residential car-washing and pool/fountain filling and are called when the utility reaches Drought Stage 2 or higher, was implemented between August 21 and November 1, 2002.

After half a decade, Charlotte experienced another drought of record in 2007-2009. Characterized by its rapid and intense onset in summer 2007 and prolonged hot and dry conditions through 2008 and into 2009 (Charlotte Water 2016b), this drought was much more severe and many community water systems in North Carolina, including Charlotte-Mecklenburg Utilities were vulnerable to running out of water had the drought continued (Lackstrom et al. 2011). With the escalation of the drought classification for the Catawba River Basin from a Drought 2 stage (severe drought) to a Drought 3 stage (extreme drought) in 2007, tougher mandatory water restrictions were announced in the county. Taking lawn watering as an example, the restriction was changed from limiting lawn watering to two days a week to a complete ban on it. The tough restrictions were enacted from September 26, 2007 to April 5, 2008, and then got amended twice when water conditions were improved gradually.

In addition, there were several short periods between 2002 and 2015 (specifically in 2006, 2011, and 2015) when the Charlotte-Mecklenburg Utilities declared a Stage 1 drought (meaning moderate drought) and voluntary water conservations were recommended. Such moderate droughts are usually due to Dry weather conditions and above-average/warm temperatures.

With the continuous growth of population and urban development and the aging water supply system in Charlotte, and climate change at global and local scales, the residents of Charlotte (“Charlotteans”) may face more droughts and water shortages in the future. Although the water restrictions guided by the regional drought response plan have been able to help limit local water demands under the safe level, such a solution is reactive and temporary instead of proactive and sustainable. We lack a basic understanding of local water consumption dynamics and its sensitivity to droughts and climate conditions.

The water management efforts made from the utility side in Charlotte-Mecklenburg are two-sided. On the one hand, the utility has been implementing capital improvement projects and water and wastewater rehabilitation/replacement projects to ensure water supply. It also strived to reduce water loss via active leak detection, pressure management, advance metering infrastructure, district metered areas and/or meter right-sizing programs(Division of Water Resources 2014). Additionally, the utility is participating in the Partnership for Safe Water Distribution System Optimization program and the Non-revenue Water and Loss Mitigation program, which are all currently active and funded. On the other hand, a number of water conservation education campaigns and programs were organized in the past (especially during drought periods), and the WaterSmart program (in the form of a website), the educational facility (Blue Planet Water Environmental Center), and community awards program have been available throughout the years. Other efforts such as a meter replacement program, a plumbing retrofit program and non-revenue water audits are also part of demand management programs of the utility.

However, it seems that the effectiveness of these non-price approaches to urban water conservation has never been formally evaluated, and there are no educational campaigns and programs specifically designed for certain consumer groups with potentials to save more water. To support demand-side management in a southeastern urban area like Charlotte, disentangling the associations between water consumption and its determinants specific to Charlotte is essential. We could get to know what kinds of households and neighborhoods consume more (or less) water and how much effects household and neighborhood characteristics have via modeling water demand. It may also offer some insights on the possible ways in which land use planning facilitates or even reinforces urban water conservation efforts.

Similarly, no studies reported the exact impacts of Charlotte's prescriptive water consumption programs such as mandatory restrictions during droughts, although their implementation did result in evident reduction of water consumption and keeping water demands manageable even at times when droughts persisted. For example, water use in total was cut by 15 percent during the first month when mandatory restrictions were called in August 2007 and by a total of 35 percent (53 million gallons per day) within 23 weeks from that August (Charlotte Water 2016b). Such a huge reduction presumably is mainly attributed to the lawn watering ban since it is regarded as the most effective means by the Charlotte-Mecklenburg Utilities. Putting assumptions aside, it is possible for us to elucidate the effects of water restrictions and investigate the spatial variations of residential consumption reduction resulted from the prescriptive mechanism, using analytical methods.

Another common water conservation mechanism is price-based. Increasing tiered water pricing was found to be effective to demand reduction as well as being less costly and more equitable in comparison with prescriptive and market-pricing programs (Baerenklau et al. 2014). Charlotte water has executed such a pricing for more than fifteen years. Although its block rate structure has been changed almost every year, the needs for recovering revenues that dropped due to the consumption cut during droughts and supporting the increasing operating budget and capital improvement projects due to growing system (for growing population) overweigh the conservation motivation.

In summary, this study has important social values aiming to empower our community with better knowledge and possible preventive strategies to cope with the uncertainty related to water shortages.

3.2 Data Collection and Processing

3.2.1 Water Consumption Data

It is widely acknowledged that water consumed by households living in single family residential (SFR) housing units accounts for a large portion of total water usage in an urban area (Wentz and Gober 2007; Balling et al. 2008). This is similar to land development in American cities where land developed for SFR dominates urban land uses. A focus on SFR water use has great implications for planning water supply systems and designing demand management programs since the single family residential sector exerts the strongest influence on water consumption dynamics (such as peak and seasonal demand due to outdoor water needs) (Day and Howe 2003). Hence SFR water consumption is the subject of interest in our study.

We obtained 1999-2013 water billing records for all types of water uses of all the customers served by Charlotte Water. Due to the nature of water bills, the usage and charge (only for water) in each record is for the corresponding billing cycle, which usually does not match calendar months. We have to estimate the daily consumption and charge based on the consecutive billing records, then multiple the number of days in calendar months (this number differs by month and by leap/regular year) to derive monthly consumption and charge. We also prepared annual consumption data according to our research needs. There are generally two ways in the literature to aggregate monthly water use. One is to sum up the water usage of the twelve months in a calendar year. Another way is to group months into seasons, and then to add up seasonal uses to get yearly use. The usages in November and December of the previous year are counted into the yearly usage of the current year, since these two months are for cold (winter) season. To differentiate the computed values from these two ways, we refer the first one as calendar yearly use, and the second as season-based yearly use here.

The billing records over time are attached to the unit called "premise" (a name given by Charlotte Water). The premise dataset has the billing address attribute as well as latitude and longitude; therefore it can be geocoded and linked to parcel data via a spatial relationship (point within polygon). A premise is not equivalent to a household because households can move, while a premise is immobile. Also household is a concept in sociology and premise is merely created for and exist in the computer system that is used to manage water services. Additionally, a premise is not equivalent to a parcel. One parcel could have zero, or one or more premises associated with it. Since our studies mainly focus on SFR water use and most of the time only one or two premises are identified for an SFR parcel, we could regard premise as residential unit. Furthermore, if we assume that each residential unit accommodates one and only one household, a premise could be conceptualized as an abstract household (thus premise and household are used interchangeably when discussing data processing). Thus, we term the average of the water uses across multiple premises on an SFR parcel as the average water consumption per household. Because no household survey data is available to retrieve household size information, we use *average SFR water consumption per household* (monthly or yearly) as the dependent variable. The shorter term *average water consumption* will replace the full name of this variable in the following chapters, and we will specify when referring to its actual meaning.

In addition to temporal aggregation, we conducted spatial aggregation to derive average SFR water consumption per household for the analyses at the neighborhood level (including census tracts and block groups). We decided to address our research objectives using census geographies as the analysis unit for several reasons. The unavailability of household-level data especially for sociodemographic variables and for a long time span and availability of census variables at the same geography for multiple years is the first reason. Second, compared to household and city/county, census geographies are more appropriate spatial scales for the representation (as continuous surface) and investigation of spatial pattern or spatial dependency in water use within a municipal or metropolitan area (Chang et al. 2010; Polebitski and Palmer 2010; Ouyang et al. 2014). It is neither too rough nor too detailed to discover and generalize spatial variation. Third, we are interested in examining the influence of urban structure and land use patterns on water consumption, and the structural variables such as SFR housing density (current or zoned) are usually calculated at the neighborhood level. This spatial scale is useful for water managers and urban planners to translate research along this line to spatially explicit water-conservation-oriented land use policy and water demand management initiative and policy (Chang et al. 2010). Moreover, it is argued that econometric models developed at the household level cannot directly capture the inter-household interactions or social norms that influence individual water use behaviors (Edwards et al. 2005; Ouyang et al. 2014), though little research progress has been achieved in elucidating this concern. However, we should be aware of the disadvantages of using aggregated data, including impossibility of analyzing behavioral and attitudinal drivers, overlook of variations across households (Worthington and Hoffman 2008), unavoidable modifiable areal unit problem (MAUP), ecological fallacy, and uncertain geographic context problem (UGCoP) (Ouyang et al. 2014).

The data processing procedure can be summarized into four steps, calculating monthly consumption for each premise, geocoding premise's address, filtering premises, data aggregation. There are a few challenges we have dealt with when preparing SFR water consumption data. It is due to the fact that the firsthand data from Charlotte Water was not collected for our specific research purpose and the reality captured by that data is much more complicated than we would think of.

We have roughly discussed the problem that water consumption was recorded on a basis of billing period instead of calendar month and our solution. Although the regular billing period is supposed to be approximately one month, there are cases in which the time interval of the two consecutive bills is more than two months. This situation occurs because the bills with zero usage sometimes are not recorded in the database made available to us. For such cases, if we divide the consumption by the number of days in a billing period, we probably will end up with underestimated daily consumption. So we make the assumption that a billing period will be at maximum the number of days between the billing date and the first day of the immediate month prior to the month of the current bill. Another observation is that negative consumption as well as charge could exist in the records, representing the correction made to the previous bill. In this case, we will need to combine the two billing records into one with the time stamp of the earlier bill before calculating daily consumption/charge.

Address geocoding is always a time-consuming process, which is particularly true when the number of total premise records is huge (around 337,000 in total and 248,000 involving residential water types). Although the majority of the premise addresses was georeferenced either directly using the coordinates information from the original data table (almost 95%) or using street network (almost 4%), there were still a few thousands addresses to be processed one by one. By joining the geocoded points with parcel data, we can not only retrieve housing characteristics but validate the correctness of those locations. Adding up the unsuccessfully geocoded and wrongly geocoded premises, unless you decide to ignore them, the manual work required is fairly intensive. This geocoding step also serves as the preliminary filtering process. To choose the premises associated with SFR water use, we used the water rate type information in the billing dataset together with the land use type of the parcel to which a premise is matched. A premise located on a single family residential parcel and with majority of bills charged by residential water rate will be a qualified candidate for analysis.

The purpose of the filtering step is to keep the premises with unique residential water use, legitimate consumption and charge, relatively complete time series, and being matched with parcel data. The whole process is complicated and iterative, and various assumptions and rules were established. We will only describe the major issues here. The original billing data is extracted from the computer system that Charlotte Water uses to manage the water service accounts, billing transactions and more. Although the CW IT staff has been working carefully to only retrieve the information we requested, there still exist some irregular records that result from the manipulation of water bills but are counterintuitive. For example, sometimes a billing record has zero consumption but non-zero water usage charge, or non-zero consumption but zero usage charge, or zero consumption and zero charge. Occasionally either consumption or charge recorded is negative. We simply exclude the above-mentioned cases. It is often observed that the quantity of water consumed and the dollars charged do not follow the tiered rate structure. Thus, we make use the average price variable derived to identify the billing records with 'abnormal' average price (defined by the values falling outside of the normal range of tiered rate – lowest and highest tier rates). There are special cases that have a very large value (>9800) for either usage or charge, though the calculated average price is reasonable. We finally exclude the records whose values in usage or charge are larger than two times the standard deviation of the billing record population being kept, since we regard them as the outliers.

Another common issue is that a premise may miss data for one or multiple billing cycles in a year or have zero values for water consumption recorded. There could be many good or 'bad' reasons for this. For example, the water service is suspended for a short or long time due to being absent from the house (in holidays, on vacation, or vacation/seasonal house), or the sale or demolition of the house or any other reasons for no water use for that time period. Sometimes a new premise is created to replace the old one, thus they store the billing records for different time periods. It is easily understood that the premise for the houses newly developed would only have the records from the date of opening service and on. The zero consumption may be an indication of erroneous bills or system errors. There could be unknown reasons too. Note that this issue may not be important if the analysis is based on household-level data combining with appropriate modeling methods that could tolerate missing data. However, for our analysis purposes (with a focus on the neighborhood level), we have to take care of those premise records when compiling time series datasets for consumption or aggregating all the available premise-level observations by geographical unit. The general and simple rule is to ensure the completeness of the records for a specific temporal scale and/or spatial scale. Specifically, for average monthly/yearly water consumption per household, we excluded the premises with any billing record that has missing data or zero value for a specific month or an entire year.

Aggregating data by a hierarchy of spatial units is the last step. For the parcels with multiple premise records, we first sum up or calculate the average of the usages depending on the rate types of premise and the relationship between premises and buildings on the same parcel. Then we aggregate parcel-level average water consumption to the selected geographic unit (census tract or block group). With regard to temporal aggregation, there are two options. One is to start with the month-based filtered data and add up monthly water consumption per household across geographical units to get average annual water consumption per household. Another way is to first filter the premises on a yearly basis then get the summation of the usages within the year. The former option will include more premises in the dataset, while the latter will more likely yield smaller yearly values. We tried both approaches and concluded that the choice will not meaningfully affect the spatial pattern of SFR water consumption.

3.2.2 Data for the Selected Determinants

Although driven by convenience, our selection of determinants of water consumption covers all the other five categories of factors discussed in the early section, except attitudinal and behavioral groups. The variable definitions and their data sources are presented in Table 1.

Table 1: Definitions of variables and their data sources

Variables	**Temporal scale**	**Geographical units**	**Unit**	**Direction**	**Data source**
Average SFR household water use	month / year	household	centum cubic feet(CCF)		Charlotte Water
Average price	month / year	household	US dollar	-	
Mandatory water usage restriction enacted	month / year	county	n/a	-	
Accumulative precipitation	month / year	county/ climate division	tenth of millimeter	-	NCDC USHCN/ GHCN
Average maximum temperature	month / year	county/ climate division	tenths of degrees (°C)	+	
Palmer hydrological drought index	month / year	climate division	n/a	+	
Household size	year	CT/BG	person	+/-	2000 SF1/SF3 census, Geolytics Estimates 2001- 2008, ACS 5-year data for 2009, 2010
Median annual household income	year	CT/BG	US dollar in 1999	+	
Median number of rooms	year	CT/BG	room	+	
Percent of population 19 years and under	year	CT/BG	%	+	
Percent of population 65 years old and over	year	CT/BG	%	+/-	
Percent of population whose ethnicity is Hispanic	year	CT/BG	%	+	
Percent of owner-occupied housing units	year	CT/BG	%	+	
Per capita income	year	CT/BG	US dollar in 1999	+	
Percent of population 25+ obtaining a college degree	2000	CT/BG	%	-	Census 2000
Ratio of own children under 18+ over number of households	2000	CT/BG	%	+	
Population density	2000	CT/BG	person per mile		
Livable area	2008	parcel	square feet	-	Mecklenburg County Assessor Database in 2008
Area of lot	2008	parcel	square feet	+	
Assessed property value	2008	parcel	US dollar in 1999	+	
Number of bathrooms and bedrooms	2008	parcel	room	+	
Area of irrigable land	2008	parcel	%	+	
Percentage of irrigable land	2008	parcel	%	+	
Area of swimming pool	2008	parcel	square feet	+	
Average house age	2008	parcel	year	+/-	
Number of single-family houses built after 1992	2008	CT/BG	house	+/-	
Percentage of houses built after 1992	2008	CT/BG	house	+/-	
Average percent of households with a pool	2008	CT/BG	%	+	
SFR housing density	2008	CT/BG	%	-	

Note: the data originally at household/parcel level is averaged to block group (BG); 1 CCF=748.05 Gallon = 2831.68 Liter

We choose three climatic measures for our analyses, including average maximum temperature, cumulative precipitation, and the Palmer Hydrological Drought Index (PHDI). The temperature and precipitation variables are the most commonly used measures in the literature on climatology and water research. The PHDI variable is mainly employed for addressing the research objective related to climate sensitivity (Chapter 4). PHDI was developed by Palmer (1965) to measure hydrological (long-term cumulative) drought and wet conditions. Because it accounts not only for precipitation totals, but also for temperature, evapotranspiration, soil runoff and soil recharge, and more accurately reflects groundwater conditions, reservoir levels, etc., it is useful in measuring the abnormality of recent weather for a region and representing historical droughts spatially and temporally (National Oceanic and Atmospheric Administration 2013).

There are two public data sources for climate data, both available through the National Climate Data Center. The United States Historical Climatology Network (USHCN) provides high-quality daily and monthly data for the three variables of our interest. The USHCN records are calculated from many long-term observing stations (the NOAA Cooperative Observer Program (COOP) Network) within relatively homogeneous climate divisions. Charlotte belongs to the fifth climate division of North Carolina (NC). USHCN has daily and monthly weather records for a long term (data from 1988 to present is available for the climate division where Charlotte is situated). USHCN data have undergone extensive quality control and corrections to remove biases (identified as systematic, nonclimatic changes), and the detailed station history information and the new, high quality U.S. Climate Reference Network data are made use of in the processing of USHCN data (National Oceanic and Atmospheric Administration 2016). Another source of weather data is based on the Global Historical Climatology Network (GHCN), which was established to serve research communities across the globe. In Mecklenburg County, the observing station of GHCN that can provide a relatively long-term dataset is located at the Charlotte-Douglas-International-Airport (USW00013881). This station started collecting daily records from July 1st, 1939. Both GHCN-Daily and GHCN-Monthly databases are subjected to a suite of quality assurance reviews. The data were collected locally (instead of regionally) and may capture the coarse trend of local climate in a better way. No drought indices are derived from the GHCN data, but yearly summaries are made available, thus there is no need for users to derive them additionally. Thus, we employ both USHCN and GHCN datasets to conduct our analyses.

Our initial goal is to collect sociodemographic variables representing household characteristics for the time period 2000-2010. Although these variables seem easy to get from the Census Bureau, census data are only available for selected years before 2010. For 2000, 2009 and 2010, we retrieved data from decennial census and American Community Survey (ACS) 5-year estimates. For the other years, we obtained a list of census variables in the CD “Annual Estimates for 2001 through 2008” published by the commercial data provider Geolytics Inc. Different estimation procedures together with different data sources were developed for population, household, housing and income estimates. Here we simply summarize Geolytics’ methodology for deriving population estimates, and readers can refer to the company’s webpage for the other procedures in details (Geolytics Inc. 2016). The basic idea in estimating population at the block level is, first to distribute the County level annual estimates in 2001 reported by the Census

Bureau to blocks, based on a 2000 Short Form (SF1) block level dataset and the race and age distribution coefficients derived from the county level totals, then developed a linear regression death-birth model to estimate the number of population who died or were newly born during the one-year interval, and relied on data from the US Postal Service to estimate immigrants and emigrants. In this way, an estimation base of population was created for 2001, which was employed to get next year's estimation by repeating the procedure.

Based on the data sources described above, we derived the variables on income, household size, age composition, ethnicity (Hispanic), owner-occupancy, education, and population density at the 2000 block group level. The potential issue with this multiple-year dataset is data (in)compatibility, due to the differences in the nature of the data sources (population-based or sampled) and the estimation methods.

The determinants in the housing characteristics category are mainly calculated from the 2008 parcel and building data, based on tax data generated from the Mecklenburg County Assessor Database. The variables we compiled include the physical attributes of SFR buildings (livable area, number of bathrooms and bedrooms, age of building - derived from the built year attribute) and parcels (lot size), and assessed property value. The information about the size of swimming pool was retrieved from the original Mecklenburg County Assessor Database. We do not have the attribute for the size of lawn in the parcel and building data, but we have information about building stories. Assuming all the SFR buildings have the same area for each of its stories and assuming that the impervious area accounts for ten percent of lot area, we estimated the irrigable lot size with the following formula (Harlan et al. 2009; Ouyang et al. 2014):

$$ILS = LS - LA/S - PS \qquad \text{(Eq. 2.1)}$$

where ILS is the irrigable lot size, LS is the lot size, LA is the livable area, S is the number of stories, and PS is the pool size. The percent of irrigable land is computed by dividing irrigable lot size by lot size (equal to ILS/LS). In addition to the variable age of building, we also derived another variable on the number of SFR dwellings built after 1992 at the neighborhood level to estimate the water efficiency potential of neighborhood housing stock.

The last variable in the list is housing density, a measure for the urban structure factor. It is defined as the ratio of the number of total SFR residential units within a geographical unit and the total area of the corresponding unit.

Unfortunately, the data for the housing and structural variables is only for the year 2008, since we did not have the readily available parcel and building data for the entire study period. If we assume the parcels/buildings developed before 2008 do not change their physical features (for many cases, it is a reasonable assumption), we could either use the same values for the majority of the variables or infer their values of the variables such as age of building and housing density for the years between 2000 and 2007. This solution does not count the buildings demolished before 2008. However, the way we filtered data for the completeness reason may to some extent lessen this problem. The problem we would be more concerned with is that it is hard to generate data of these housing and structural variables for the year 2009 and 2010, because there should be new parcels and buildings emerging after 2008. This means the new development pattern emerging in those two years will be neglected from our analyses.

In summary, we introduced the variables that are thought to be useful in addressing our different research objectives. Note that not all of the variables are employed and reported in each of the following chapters, although we did examine each of them during the whole process. We did not provide descriptive statistics for these variables here because the dataset used in each chapter varies slightly; thereby we prefer to report them in each individual chapter. In the next section, we present a few maps and descriptive statistics generated from the residential water consumption data.

3.2.3 Water Consumption in Charlotte in 2000-2013

Based on the derived monthly water consumption data, we calculate average annual water use for each water rate type (Table 2). The procedure is to first sum up monthly water use for each year, then get the mean of the annual water use for the period 2000-2013. We roughly exclude the records with negative value in consumption, and keep the premises that were successfully matched to parcel data. As seen from the table, the customers paying for residential, commercial and multi-family water rates rank the top three largest water user based on total average annual water consumption. The total volume of water used for lawn purpose ranks the fourth. When the mean or median of average annual water use across the premises/service accounts is considered, the water consumption per user is the smallest for the residential rate type (WR) (excluding water used for fire protection). The mean/median volumes of average annual water use in the multi-family and single family detached rate categories are very high, which is reasonable since each service account records the aggregated water uses in multi-family dwellings. The boxplot (Figure 5) illustrates the variation of average annual water use by rate type. For each of the rate categories WA, WC, WG, WL, WN and WR, there are some customers using much more water.

Table 2: Average annual water usage by rate type

Water rate code	Description	User	Descriptive statistics (unit: CCF)					
			Min	Max	Sum	Mean	Median	Std. dev.
WA	Multi-Family	2,338	1	113496.2	**7558755**	**3233**	**872.9**	6287.4
WAD	Single Family Detach	39	50	192656	167105.2	**4284.7**	**2793.2**	4775.7
WB	Sale for Resale	7	4	40149.3	125483.1	17926.2	17288.6	14388.8
WC	Commercial	13,177	0.118	157556.3	**9592275**	**728**	**121.5**	3351.1
WF	Fire Line	2,450	0.035	19998	131537.6	53.7	4.41	517.9
WG	Governmental	364	0.13	108717.2	436350.2	1198.8	156	7138.6
WH	Health Facility	7	53.65	73185.7	114455.8	16350.8	380.2	26788
WI	Industrial	65	1.5	19969.2	107127.7	1648.1	498.2	3435.7
WL	Lawn Meter	9,305	0.035	18785.6	**3340824**	359	**156.3**	715.3
WN	Institutional	483	0.258	129885.4	1117750	2314.2	406.9	9808.3
WO	Other Non-Profit	728	1	11578.8	206200.0	283.2	109.4	718.6
WR	Residential	243,640	0.115	33449.7	**2.32E+7**	**95.4**	**82.0**	143
WS	Swimming Pool	167	1	3064.2	66169	396.2	265.5	462.5
WTUC	Union County Residential	16	57.77	320.8	2122.8	132.7	118.3	67

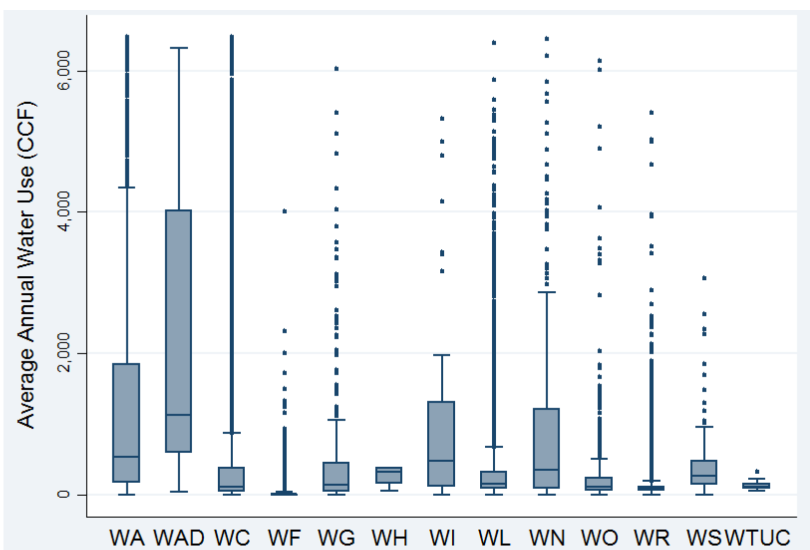
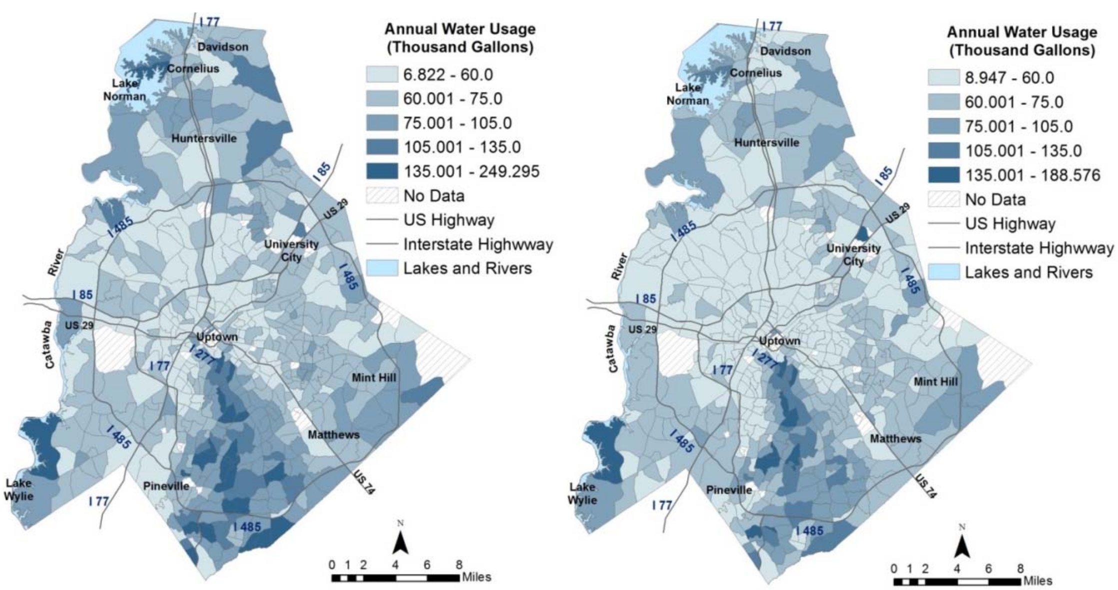


Figure 5: Average annual water usage by rate yype in 2000-2013

(excluding the uses larger than 6500 CCF for illustration purpose)

We are interested in the spatial distribution of annual household water use. We identify the premises/service accounts that are associated with the parcels used only for single family housing, multi-family housing and condos and are mainly charged by residential-related water rates (WA, WAD, WR, and WTUC) and the lawn and swimming pool water rates. Then the monthly records are aggregated to obtain annual totals for each year in 2000-2013. The annual uses are averaged across the time for each premise, and finally the values are aggregated to the neighborhoods (represented by the block groups in 2010). The average annual water consumption per service account shows the concentration of higher water users mainly in the wedge of affluence south of Uptown (Figure 6a). There are a few neighborhoods in the north, south-eastern and south-western corners of the county consuming more water as well. Lower water usage is found pretty ubiquitously in closer-in neighborhoods.

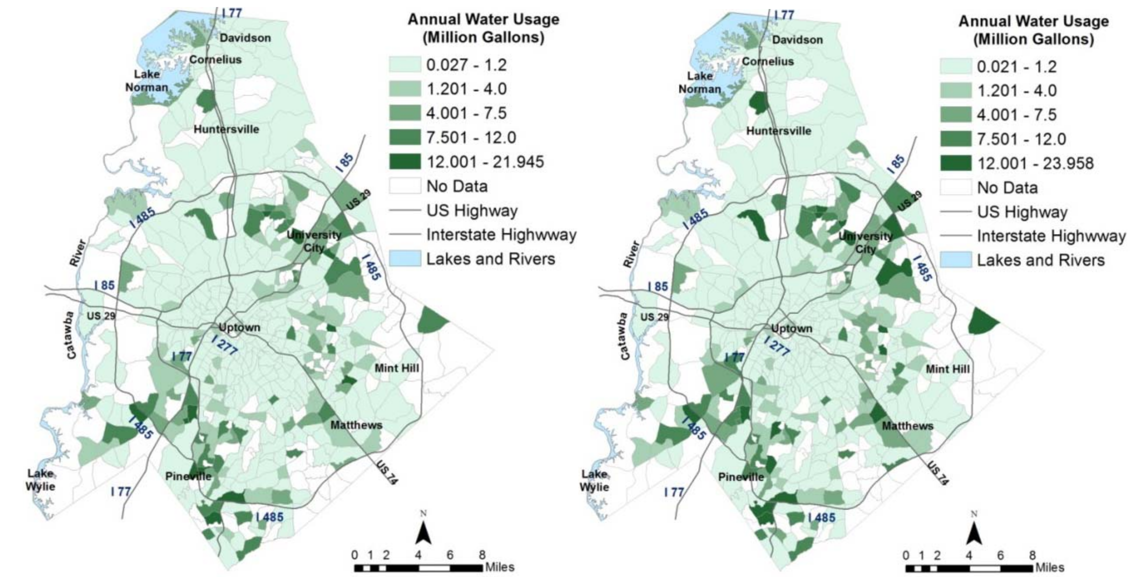


(a) 2000-2013 (b) 2013

Figure 6: Single family residential average annual water use per service account

Due to the two severe droughts in Charlotte's recent history, the average annual water use could be influenced by the peak annual uses in certain years (e.g. 2007 and 2008). Thus, we separately mapped the annual use per premise in 2013 when the most recent data is available and had a normal weather. Although the general pattern of annual water use in 2013 is similar to the averaged trend during the time period 2000-2013 (Figure 6b), its range of the annual water usage across the block groups is smaller and fewer neighborhoods are associated with higher water consumption.

Similarly the average annual water use per premise by multi-family or condo households in 2000-2013 and in 2013 is presented in Figure 7. The spatial patterns and the value ranges are similar for the 14-year period and the single year. Higher water usage primarily concentrates in the northeastern and southeastern areas along I-485, including the University City area. Surprisingly the center of the city where a few apartment/condo buildings are located does not have any higher water users. The possible reason is that the apartment/condo buildings near the periphery of the city have large areas of lawn and pools demanding water since we included the water volumes recorded from the water meters for lawn and pools.



(a) 2000-2013 (b) 2013

Figure 7: Multi-family residential/condo average annual water usage per service account by 2010 block group

The descriptive statistics for the single family and multi-family residential/condo average annual water usage per service account during 2000-2013 are listed in Table 3. The value range for the multi-family/condo-related usage is very large, indicating the mixture of building with various residential units.

Table 3: Average annual water usage by residential land use types

Land use type	**Users**	**Descriptive statistics (unit: CCF)**					
		Min	**Max**	**Sum**	**Mean**	**Median**	**Std. dev.**
Single Family Residential	227,554	0.129	33449.7	22700000	99.833	84.743	152.394
Multi-Family Residential/Condo	15,897	0.903	113496.2	6197934	389.881	58.167	2283.434

CHAPTER 4: SENSITIVITY OF SINGLE FAMILY RESIDENTIAL WATER CONSUMPTION TO WEATHER VARIABILITY

4.1 Introduction

Urban water resource management in Charlotte has been challenged by continuing population growth, urbanization, and potential climate change since the beginning of the twenty-first century. Although the estimated metered average water usage per capita per day within the service area of Charlotte Water showed a (slightly) declining trend (Division of Water Resources 2014), its variation over the years is evidently large. Such a variation could be attributed to the combined effects of climatic variability, water conservation efforts and household behavioral changes of water consumption (Balling and Gober 2007).

The average SFR residential water usage per premise (equivalent to households to a large extent) in Charlotte is approximately 75,000 gallons per year over the time period 2000-2013. This amount is similar to the ones reported for Seattle, Washington (Polebitski and Palmer 2010) and Hillsboro, Oregon (Chang et al. 2010). Among Charlotte's customers living in a SFR house, it is estimated that around 5 percent use at least as much as the water (around 171,712 gallons) consumed by the residents in Phoenix, Arizona (a city in a semi-arid climate). Although we do not have much knowledge about the outdoor portion of SFR residential water usage, considering the distinct four seasons in Charlotte and the possible need for irrigation and outdoor activities such as the use of swimming pool, outdoor usage may account for a large proportion of the average residential household water consumption, and we know that outdoor water usage is usually responsive to weather variation.

Given Charlotte's steady 2 percent or above annual population growth rate during the 2010-2014 period (Charlotte Observer 2015), the projected population of Charlotte will reach 1.3 million by 2040. The potential water demand of such a huge population will undoubtedly impose pressure to local water supply and its sources, and the climate change related to global warming (such as more extreme daily temperatures and more intense precipitation events) may intensify the threats of water shortage in future. Therefore, despite the humid subtropical climate, it is still relevant and important to understand the sensitivity of local water usage in Charlotte to weather variations from a demand management perspective. There lacks empirical research (of any kind) dedicated to Charlotte regarding this topic. This study as the first step will bring some insights on the relationship between weather variability and Charlotteans' water consumption, and hopefully contribute to local water conservation efforts particularly related to outdoor water use reduction.

4.1.1 Research Statement

Research relating water use to climatic conditions has mostly been conducted on the cities in the southwest. Although an obvious link between weather and climate conditions and water consumption would be anticipated, research findings have been disparate and inconsistent (Gutzler and Nims 2005). Some studies discovered significant associations between temporal variations in water consumption and variations in weather, while others found no link at all (Balling et al. 2008). The mixed results about weather effects may be in part because of differences in measures of climatic factors and other variables, temporal scales (yearly vs. monthly vs. daily), and methods (Table 4), but more importantly, each case has a unique context including water pricing, environmental circumstances, urban lifestyles, and individual and household preferences toward water use in general and the relative importance of outdoor versus indoor use (Balling et al. 2008). Since the cities in the southeast have distinctive climate conditions and water ethics, it is interesting to investigate how the mechanism linking climatic variation and water consumption pattern in the case study of Charlotte disagree or agree with the existing literature, in what aspects, and maybe why.

Previous research rarely examined the relationship between climatic variability and residential water consumption with a dedicated focus on weather conditions (rather than price) and at a fine spatial scale. Balling et al. (2008) proposed a study to explore the intraurban spatial variations of the sensitivity of residential water usage to climatic conditions in Phoenix, AZ. They found that different neighborhoods (represented by census tracts) show sensitivity of variable magnitude (from zero up to 70%) to climatic variation during the period of 1995-2004. The neighborhoods with greater sensitivity share common characteristics including large lots, many pools, a high proportion of irrigated mesic landscaping, and a high proportion of high-income residents, while the neighborhoods accommodating large families and many Hispanic population are less sensitive to the same climatic conditions. Based on these findings, they suggested specific ways to encourage the development of low sensitivity for the purpose of enhancing the resilience of urban water use to climatic change. We will follow their approach to examine the climatic sensitivity of single family residential water consumption in Charlotte at a smaller geography of neighborhood (block group).

Table 4: A summary of the literature on the relationship between climatic variation and water consumption

Studies	Data	Cities	Dependent variables	Independent variables	Mo-del	Scale	Sig.	Reasoning
Berry and Bonem (1974)	Cross-sectional, 16 samples in 1970	16 cities/ towns in New Mexico	Municipal water use per capita per day (gal)	Per capita income, *moisture deficit*	OLS	municip al	No	Relatively smaller geographic area; controlled for sizable climatological variation
Mayer and DeOreo (1999)	Cross-sectional, 12 samples derived from 1,000 SFR HH samples per study site, the most recent available complete year (1996/1997) of historic billing data	12 cities in US	Avg. annual outdoor consumption (kgal) avg. seasonal / outdoor use component (%)	*Net E*	OLS	national	Yes	Seasonal water use
Michelsen et al. (1999)	Panel, 7 observations over the period bt. 1984 and mid-1995	7 south-western cities	City total monthly water consumption	Avg. price, HH income, *Avg. monthly T, total monthly P*, $of SFR accounts, # of non-conservation programs in effect, drought	Panel model	regional	Yes for T, No for P	A focus of Nonprice conservation programs, as control variables
Maidment and Parzen (1984)	Time series, 1961-78 each for a study city	6 cities, 3 each in High Plains and East Texas	Monthly water withdrawal, after detrending, deseasonalizing, and autoreg-ressive filtering	*Precipitation, evaporation, maximum air temperature*	OLS	municip al	Yes for P, E only High Plains	Time patterns
Wilson (1989)	Cross-sectional 1963-65	10 sub-urban area in the west, plus Fort Worth	Average summer sprinkling demand per day per dwelling unit	Marginal price, home value, *moisture deficit*	OLS	regional	Yes	forecasting

Studies	Data	Cities	Dependent variables	Independent variables	Mo-del	Scale	Sig.	Reasoning
Rhoades and Walski (1991)	Time series for each of the twelve months in a year, 1968-89	Austin, TX	Pumpage (MG) for each month	Population, *R*, CDD	OLS	municipal	Yes, summer	Forecasting
Billings and Agthe (1998)	Panel, 1974-88	Tucson, AZ	Avg. monthly water per HH	Marginal price, block rate subsidy, real income per capita, *Avg. T, P*	OLS	HH	Yes	Forecasting, comparison
Gutzler and Nims (2005)	Time series, 1980-94	Albuquerque, New Mexico	Total annual, Per capita, or year-to-year change in summer-season residential water demand	Avg. max. T, avg. daily P rate	OLS	municipal	Yes to both	Interannual variability of water demand and summer climate Groundwater supply insensitive to short-term climate variability; summer-season 1/3 residential
Young (1973)	Time series, 1946-64, 1965-71	Tucson, AZ	Annual water pumped per active service	Avg. charge per 1000 gallons, R, T, E, retail sales per capita per unit time	OLS	municipal	Yes to R, No to T & E;	Price elasticity, for control
Billings and Day (1989)	Panel, 1974-80	Tucson, AZ	Avg. monthly water use by district	Price, rate premium, income, T, High T, Summer rain, Winter rain, publicity, HH size, age 55-64, age 65+, new HH, growth	OLS	municipal	Yes to all, esp. high T	Price elasticity, for control
Billings and Agthe (1980)	Time series, 1974-77 (September)	Austin, TX	Monthly water consumption per HH (all residential types)	Marginal price, difference price, implicit marginal sewer charge, personal income per HH, *ET minus rainfall*	OLS	municipal	Yes	Focus on price elasticity, used for accounting for weather and sprinkling demand variations

CDD: cooling degree days; E: evaporation; ET: evapotranspiration; P: Precipitation; R: Rainfall; T: Temperature; Moisture deficit = Thornthwaite potential evapotranspiration minus 60% of the rainfall; Avg: Average; HH: household; Max: Maximum; # number

In this chapter, we will address three objectives and their related research questions.

(1) Explore the temporal trend and spatial patterns of water consumption during 2000-2010 and examine their association with historical climatic conditions in Charlotte. This objective will focus on the question that *"Does water consumption and its change before, during, and after drought represent spatial variation?"*

(2) Determine whether there were evident geographic patterns in climate sensitivity of SFR water consumption (measured by the ratio of summer versus winter water use) and their changes before, during, and after drought. The corresponding research question is: *"Is there an evident geographic pattern in climate sensitivity (measured by the ratio of summer versus winter water use)?"*

(3) Determine how much variations of water consumption can be explained by climatic factors, whether the neighborhoods (block groups) in Charlotte have different levels of sensitivity and respond to climatic variation differently across space. We are interested in answering the questions *"Are some neighborhoods more climatically sensitive than others?"* and "what are the socio-demographic and housing characteristics of the neighborhoods with lower or higher sensitivity?"

We first examine the spatial patterns of average per-household annual water consumption for single family homes over the 11-year study period 2000-2010. The ratio of summer versus winter water usage within each year is also mapped to reveal geographic patterns in climate sensitivity. Using the same spatial data, spatial clusters are identified for the change rate of winter, summer and annual water usage from the consecutive years of drought periods in Charlotte in the first decade of the twenty-first century. Next, we construct a time series of monthly SFR water use anomalies and explore their variation with three climatic measures (temperature, precipitation, and the Palmer Hydrological Drought Index); with this series, we identify social-demographic variables and housing characteristics that are associated with the explained variations.

We hypothesize that among the household and housing determinants of residential water consumption discussed in chapter 2, those that are closely linked to outdoor water usage will contribute to the intraurban spatial variations of the sensitivity being investigated in this case study. We also anticipate that the neighborhoods in Charlotte would respond differently to atmospheric conditions, but the relationship between temporal variations in SFR water consumption at the neighborhood level and weather variations may be weaker than in other studies given Charlotte's humid climate and relatively cooler and shorter summer compared to Phoenix.

4.1.2 Methods

Three basic methods are applied to address our objectives. One is to use GIS mapping method to present the spatial distribution of SFR water consumption in different years. Correlation and regression analysis which we apply to deal with the third objective are common methods for studying the associations between/among variables. We also conduct spatial clustering analysis to assess global and local spatial clustering in the percent change and the climatic sensitivity of SFR water consumption. A brief introduction is given to this analytical approach since it is originated from geography and regional science and has been gaining popularity in other domains in the past decade.

Global patterns can be assessed by Moran's I test (Moran 1950)[2]. The null hypothesis of this test in our application is that the spatial pattern in the SFR water consumption in Charlotte is random. Rejection of the null hypothesis indicates a spatial autocorrelation in the distribution of SFR water consumption. The Moran's I statistic is given as:

$$I = \left(\frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}\right) \times \left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i-\bar{y})(y_j-\bar{y})}{\sum_{i=1}^{n}(y_i-\bar{y})^2}\right) \qquad \text{(Eq. 3.1)}$$

where n is the number of spatial units (block groups in our case); i and j are indices of spatial units; y is the outcome variable of interest; $\bar{y}$ is the mean of y; and w_{ij} is the spatial weight between units i and j. The spatial weights for each pair of units compose a $n \times n$ spatial weight matrix W, quantifying the conceptualized spatial relationships that exist among the spatial units being investigated. There are various ways to specify a spatial weight matrix, and contiguity-, k-nearest-neighbor- and distance-based spatial weights are commonly used. As for contiguity-based spatial weight matrix, the spatial weight between two units is assigned to one when the units share a common border or corner, otherwise the spatial weight is zero. The other two spatial weight matrix specifications are more intuitive, basing on a threshold of the number of nearest neighbors or the distance between two spatial units (centroid distance for polygon features). The latter two are better options when there are holes in the distribution of spatial units. As a common practice, the spatial weight matrix is row-standardized to get the sum of the weights for each row (each unit i) equals to one (Anselin et al. 2008).

---

[2] There are other statistics measuring global spatial autocorrelation, such as Geary's C and Getis-Ord G statistic.

The Moran's *I* statistic ranges from -1, denoting the highest negative autocorrelations (dispersion), to +1, reflecting the highest positive autocorrelations (clustering). The expected value of Moran's *I* under the null hypothesis of no spatial autocorrelation (randomness) is $\frac{-1}{n-1}$.

As a measure of global spatial autocorrelation, Moran's *I* does not identify where the significant clusters are, nor does it indicate what type of autocorrelation is occurring spatially (Anselin 1995). Therefore, local indicators of spatial association (LISA) were proposed to test the null hypothesis of spatial randomness (or no local spatial autocorrelation) by comparing the values in a given location with values in neighboring regions (Anselin et al. 1996).

A local measure of Moran's *I* is a special case of a LISA, and it decomposes spatial autocorrelation patterns into four types of clusters, high-high, low-low, high-low, and low-high (Anselin et al. 1996). Positive spatial autocorrelation refers to a geographical distribution of values when an above-average value is surrounded by the neighbors with above-average values (high-high, HH) and vice versa (low-low, LL). If a location with a high value (above-average) is encompassed by low-value neighbors, or inversely, a location with a low value (below-average) is encompassed by high-value neighbors, a negative spatial autocorrelation is present (Messner and Anselin 2004). The local Moran's *I* statistic is calculated as:

$$I_i = \frac{\sum_{j=1}^{n} w_{ij}(y_i-\bar{y})(y_j-\bar{y})}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \qquad \text{(Eq. 3.2)}$$

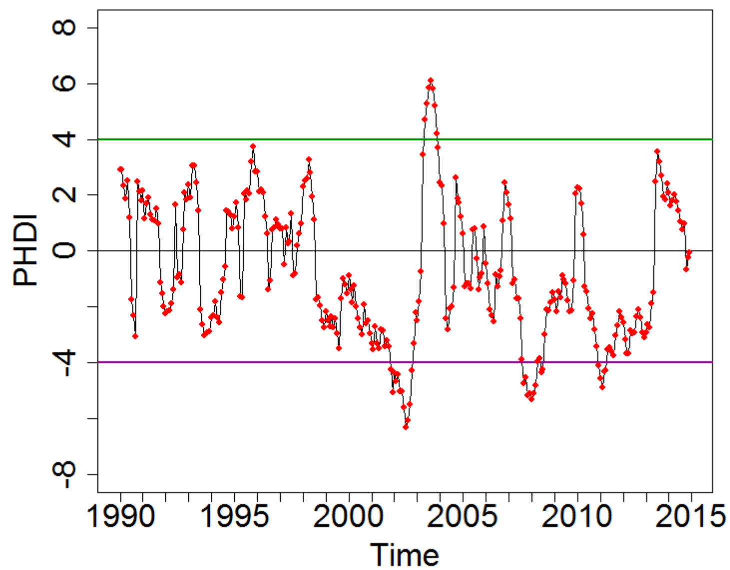
We calculate the global and local Moran's *I* in GeoDa™, a spatial analysis software. A randomization approach is used to generate a spatially random reference distribution to assess statistical significance (GeoDa Center).

4.2 Overall Trends in Meteorological Conditions and SFR Water Consumption

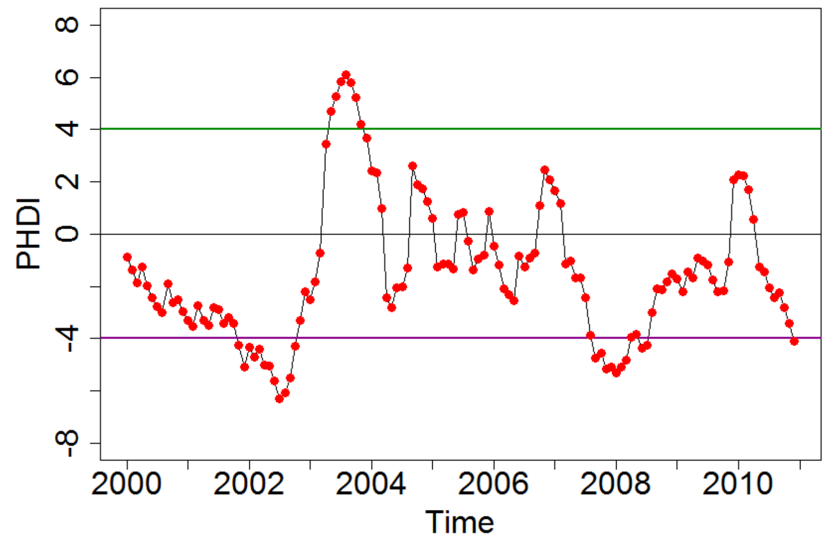
4.2.1 Historical Meteorological Conditions in Charlotte

We follow the extend literature on weather and water research and choose the three typical weather measures including average temperature, accumulative precipitation, and the Palmer Hydrological Drought Index (PHDI) to analyze the temporal variability of weather in Charlotte. Temperature and precipitation are supposed to impact outdoor water usage (mainly for watering lawn and swimming pool) more than indoor usage (mainly showering, laundering and drinking). The PHDI is a composite indicator of meteorological conditions, and has been used as one of the state-level drought indicators of South Carolina (Mizzell Hope et al. 2010) and as an important index included in the Catawba River Basin Drought Model (North Carolina Division of Water Resources and Department of Environment and Natural Resources 2005). Due to the different nature of the meteorological data sources (GHCN and USHCN) described in chapter 2, we employ both GHCN and USHCN datasets, especially for the third objective.

Figure 8 presents the trend of the monthly PHDI of North Carolina's fifth climate division (referred as the Charlotte region in the following discussion) for the periods of 1990-2015 and 2000-2010. The National Oceanic and Atmospheric Administration (NOAA) classifies the values of various Palmer Drought Indices including PHDI into seven categories. Values less than -4 or more than +4 indicate extreme drought or extremely wet conditions; values near zero (between -2 and 2) indicate normal conditions for a region. The value ranges [-4, -2] and [2, 4] are divided into two classes each (-3 and 3 are the break points), representing extreme drought and moderate drought, and moderately wet and very wet conditions. Figure 9 shows the seasonal transitions of the drought conditions in the region, based on NOAA's classification and Table 5 lists the PHDI values over time.



(a) 1999-2015



(b) 2000-2010

Figure 8: Monthly palmer hydrological drought index for the fifth climate division of North Carolina

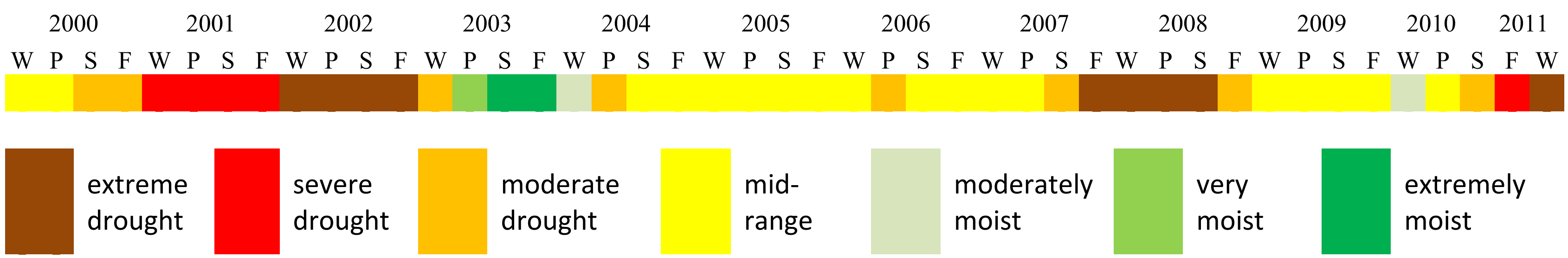


Figure 9: The classified drought conditions by seasons for the fifth climate division of North Carolina, 2000-2010

(W: Winter, P: Spring, S: Summer, F: Fall)

Within the 25-year time span of 1990-2015, the Charlotte region has experienced extreme drought conditions three times. All of the extreme drought periods occurred after 2000. However, their characteristics differ significantly. The first drought starting in the summer of 2000, gradually developed over 2001, then worsened, and reached the worst situation (in history) during summer 2002. Such extreme drought condition persisted until the beginning of 2003, and reversed to extremely wet quickly, in less than a year. Within the next three and a half years, the region's drought condition stayed at a more or less normal status. However, without a clear warning sign, the second drought threatened the region starting in September 2007, continued and peaked in January 2008, and retreated gradually. Compared to the first drought, the drought between 2007 and 2008 was more acute, was a little shorter, and had a lower peak value. The acuteness and persistence of the 2007-2008 droughts jeopardized the water supply and usage in the county. A series of water use restrictions were enacted to limit local water demands under the safe level and to ensure the uninterrupted water supply for the entire river basin that might face continuous drought conditions. Among the restrictions, the toughest ones were announced at the end of September 2007, banning lawn watering and the operation of any sprinkler system and disallowing the operation of ornamental fountains (without fish), residential car washing and refilling of swimming pools. These rules were enforced for seven months, and then amended to allow one-day-a-week outdoor water use until September 2008, as drought conditions improved. It took the region a longer time to arrive at a wetter condition after the 2007-2008 droughts. The development of the third drought in 2011 was also quick, but not as extreme as the second one. After the greatest extreme level of the third drought was reached, the drought condition kept at the severe level for almost a year. The mild wet condition came back in winter 2013.

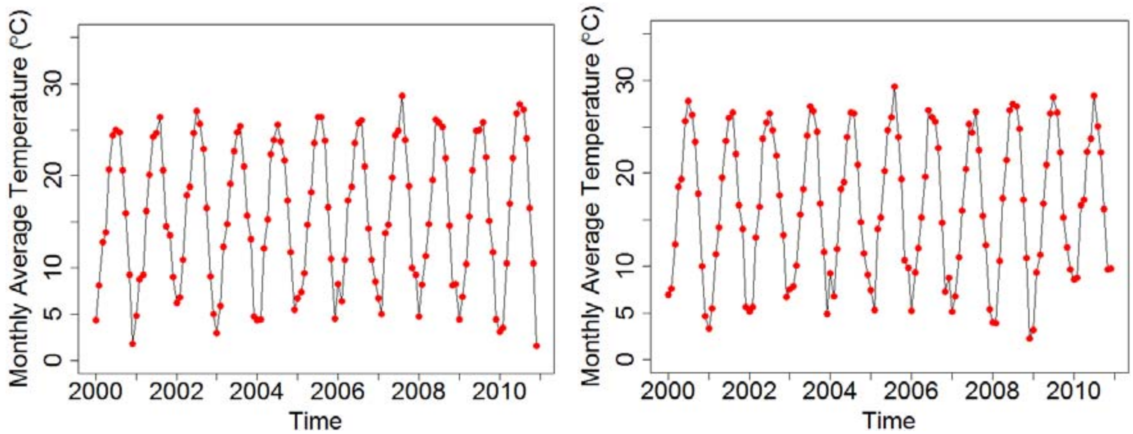
Table 5: The monthly PHDI for the period 2000-2010, 5th climate division of NC

Yr/M	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2000	-0.88	-1.39	-1.86	-1.25	-1.99	-2.42	-2.76	-3	-1.91	-2.61	-2.5	-2.95
2001	-3.3	-3.54	-2.72	-3.32	-3.48	-2.83	-2.87	-3.42	-3.21	-3.43	-4.26	-5.07
2002	-4.34	-4.69	-4.42	-5.02	-5.04	-5.61	-6.31	-6.08	-5.5	-4.28	-3.32	-2.2
2003	-2.5	-1.82	-0.74	3.45	4.7	5.27	5.83	6.11	5.8	5.21	4.21	3.68
2004	2.43	2.33	0.99	-2.43	-2.83	-2.06	-2	-1.31	2.61	1.88	1.73	1.23
2005	0.61	-1.27	-1.16	-1.16	-1.34	0.76	0.81	-0.28	-1.37	-0.95	-0.82	0.88
2006	-0.45	-1.17	-2.1	-2.33	-2.54	-0.84	-1.26	-0.93	-0.71	1.1	2.46	2.08
2007	1.65	1.17	-1.16	-1.03	-1.69	-1.69	-2.44	-3.88	-4.74	-4.54	-5.17	-5.09
2008	-5.31	-5.09	-4.82	-3.95	-3.84	-4.37	-4.25	-3	-2.1	-2.14	-1.84	-1.51
2009	-1.73	-2.19	-1.44	-1.67	-0.9	-1.03	-1.18	-1.77	-2.19	-2.15	-1.07	2.06
2010	2.25	2.23	1.68	0.58	-1.26	-1.46	-2.06	-2.43	-2.24	-2.81	-3.41	-4.09

Note: extreme drought conditions are highlighted in yellow.

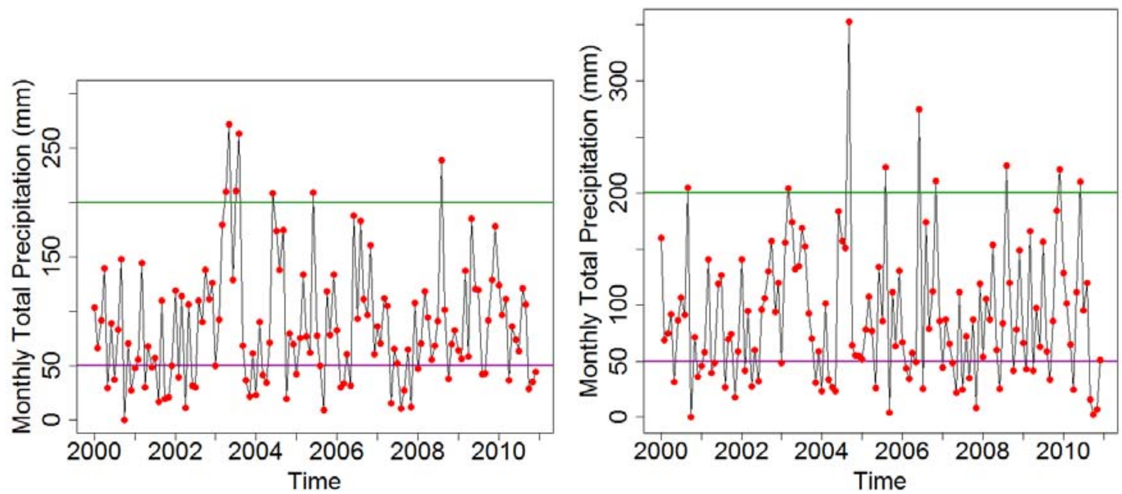
Both GHCN and USHCN temperature datasets show seasonal cycles over time (highest in summer and lowest in winter). The GHCN mean temperature data for the county by month (Figure 10a) indicate the highest temperature (28.7°C) in August 2007 and the lowest (1.5°C) in December 2010, while the monthly average temperature records from USHCN (Figure 10b) show a range from 29.32°C in August of 2005 to 2.25°C in December 2008.

There is less seasonality captured in the plot of monthly precipitation during 2000-2010 (Figure 11). The differences between the precipitation curves of GHCN and USHCN imply that Mecklenburg County has on average higher monthly rainfall than the region as a whole. There were two months (271.8mm in May 2003 and 263mm in August 2003) among the months of 2000-2010 having more than 250 millimeter rainfall. In October, 2000, the county had zero rain recorded for the entire month. The monthly precipitation from GHCN during the first drought period (mainly in 2002) is much higher than the one from USHCN. The differences in GHCN and USHCN precipitation records in 2007 (the early stage of the second drought) are less obvious.



(a) GHCN data for Mecklenburg County (b) USHCN data for the 5th climate division of NC

Figure 10: Monthly mean temperature, 2000-2010



(a) GHCN data for Mecklenburg County (b) USHCN data for the 5th climate division of NC

Figure 11: Monthly accumulative precipitation, 2000-2010

4.2.2 Single Family Residential Water Consumption in Charlotte

We use season-based annual use of SFR water consumption per household to understand its spatial and temporal variations across years from 2000 to 2010. Two ways were introduced in chapter 2 (on page 41) to conduct filtering and temporal aggregation of the monthly consumption dataset. Figure 12 shows the mean of average annual water consumption across block groups (the smallest geographical unit examined in this research) for each year within the time period 2000-2010, applying the two methods (month- and year-based filtering) described above. The two curves reveal similar trends in the change of annual consumption over time except between the years 2005 and 2006. We can recognize the first drought period from both curves (highest yearly consumption in 2002 and lowest in 2003). For the second drought period, the annual water use in 2007 is higher than the one in 2008, although during the early months of 2008 the drought condition had been persistently extreme. The decrease in yearly average consumption in 2008 is more likely to result from the water bans issued in Fall 2007, though it may also be attributed to the recovery from the regional drought condition in the second half of 2008.

When the annual consumptions based on the two methods are averaged across time for each block group and then mapped (Figure 13), we observe similar spatial patterns with higher usage concentrating in the south and the north of Mecklenburg County and lower usage surrounding the urban core.

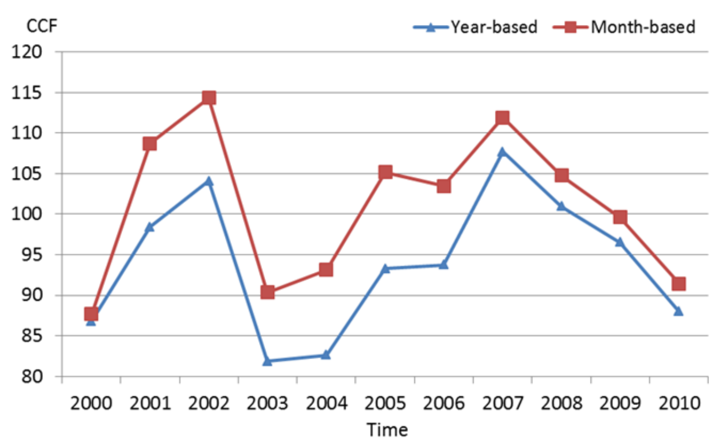
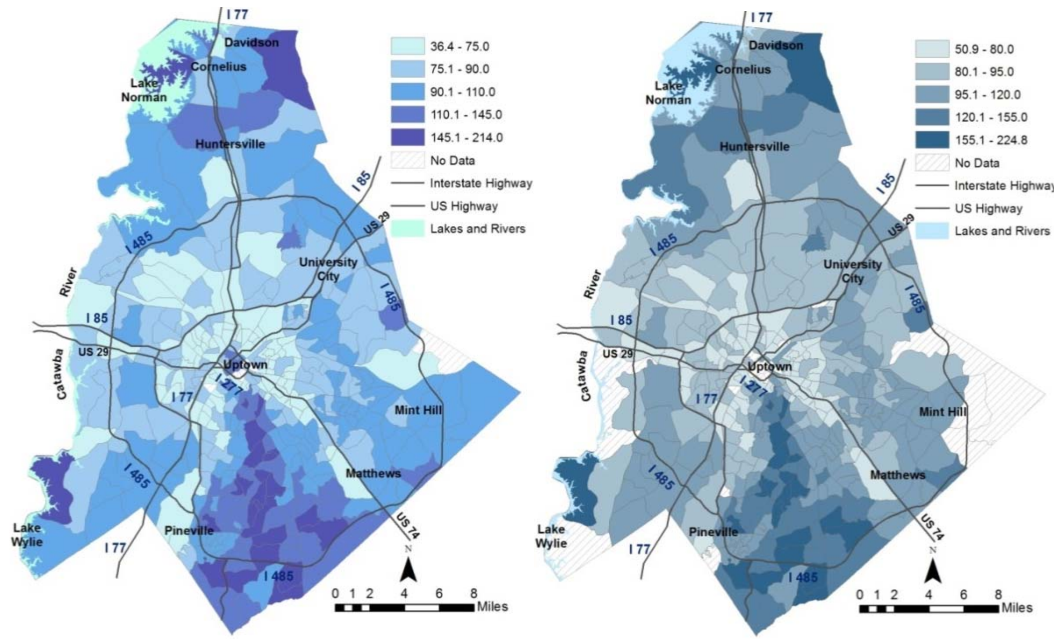


Figure 12: Average annual single family residential water consumption per household across block groups

(1 Centum Cubic Feet (CCF) = 748.05 Gallons = 2831.68 Liters)



(a) Year-based filtering (b) Month-based filtering

Figure 13: Average annual SFR water consumption per household by 2000 block group, 2000-2010

4.3 Spatial and Temporal Patterns of SFR Water Consumption per Household

4.3.1 Average Annual Water Consumption in Charlotte

We focus on investigating the average annual water consumption per household derived from year-based filtering and map its spatial patterns before, during and after the two drought periods in history (Figure 14 and 15). All the block-group-level annual values over the entire study period (2000-2010) are stacked together to define the classification scheme. The break points are chosen based on the standard deviation classification method. The seven classes in the legend (colored in blue-green-yellow-orange-red) represent the levels of less than -2.5, -2.5~-1.5, -1.5~-0.5, -0.5~0.5, 0.5~1.5, 1.5~2.5, and more than 2.5 standard deviations.

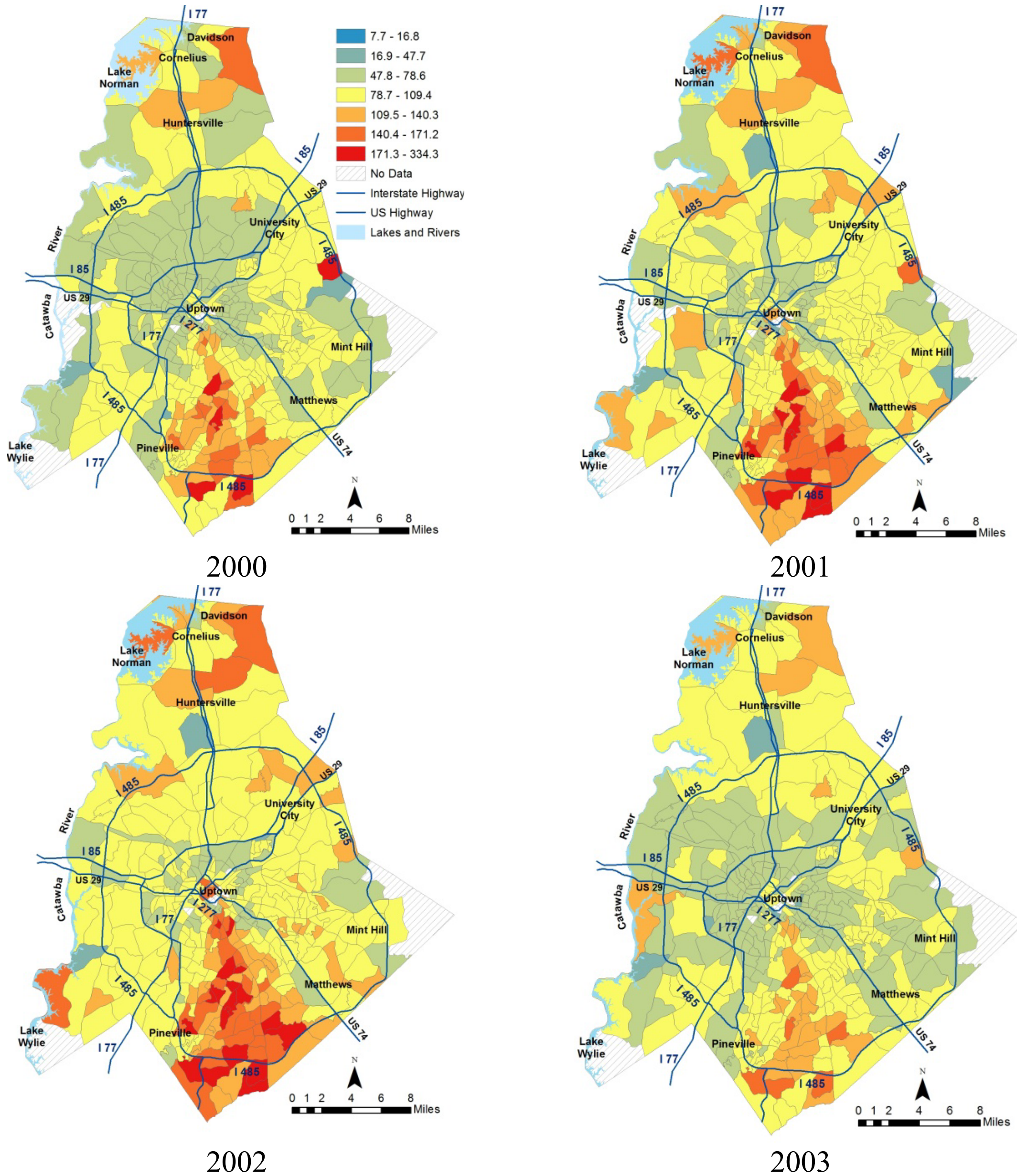


Figure 14: Average annual SFR water consumption per household by 2000 block group, 2000-2003

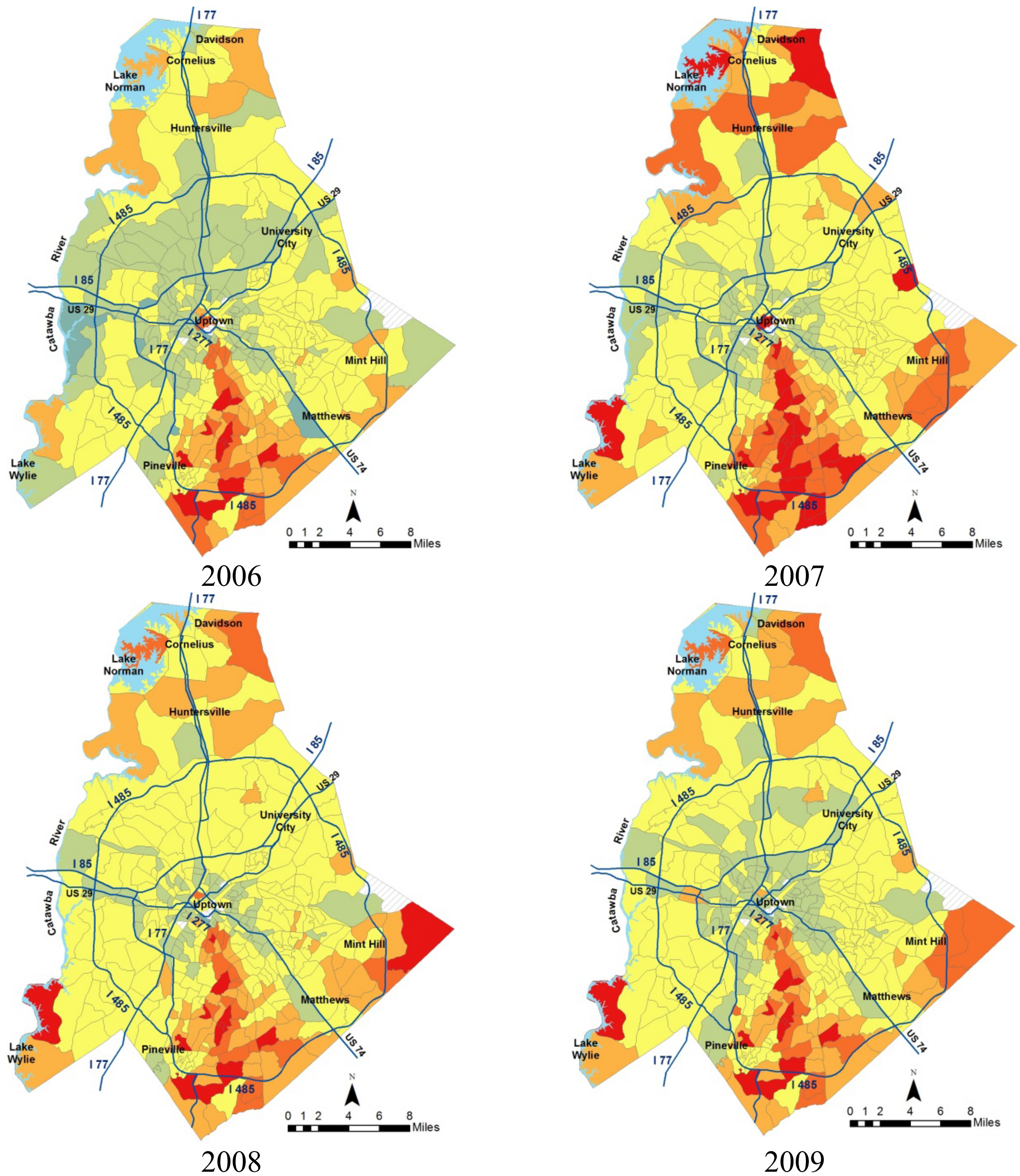


Figure 15: Average annual SFR water consumption per household by 2000 block group, 2006-2009

The general patterns across the county are consistent with the trend of average annual consumption per household shown in the curves of Figure 12 and with the pattern seen in Figure 13. From 2000 to 2002, more than half of the block groups experienced a gradual upward shift in the levels of water use; then the average annual water consumption levels quickly and dramatically decreased within a year in the majority of places within the Mecklenburg County. As for the second drought period, a number of block groups responded with an increase in water consumption, and the slowly improving drought conditions during the years 2008 and 2009 resulted in gentle down-grading in the levels of water use.

When we closely examine the geographical differences in the patterns seen in the maps for before, during and after the two droughts, it is noticeable that the responses of neighborhoods to droughts are not homogeneous. During the time period 2000-2003, the northern and the southern Charlotte neighborhoods were fairly responsive to the change in drought conditions. So were the municipalities of Davidson and Mint Hill. The towns of Cornelius and Pineville were more sensitive to extreme drought, while Matthews and the east and the west of Charlotte were more susceptible to moistness. The neighborhoods along Independence Boulevard (Highway 74) and along the I-77 South corridor behaved differently (following the weather conditions vs. responding more to the moist condition) when the first drought occurred. The areas west or south to the outer ring (I-485) showed the similar changes in pattern (responsive consistently), while most of the block groups east to the outer ring had a greater drop in water consumption when high rainfall occurred. The sensitivity of Huntersville neighborhoods to droughts showed a mixed pattern across the town.

Compared to the first drought, the second drought resulted in a simpler pattern of change. Most of the places responded correspondingly to the changes in the drought conditions (higher use in 2007) as well as the implementation of city-level water use restrictions in 2008 (lowered use). Only in a few neighborhoods such as southern Charlotte, along the I-77 South, and the southwestern/southeastern corners, average annual water usage per household kept increasing or remained same during 2007-2008, and the effect of the non-price policy was least observed. There were no obvious responses to the meteorological conditions observed (no change in the class of water consumption) within the center and west of Charlotte and the neighborhoods along Independence Boulevard.

In summary, the response of the average annual water consumption per household over time shows spatial heterogeneity across the county and this also transpires for the two drought periods with distinct characteristics. However, the annual aggregation masked the seasonal changes in water usage. To discover further evidence of geographic pattern of water consumption in weather sensitivity, it is better to examine the average usage during two distinct seasons – winter and summer.

4.3.2 Average Winter and Summer Water Consumption in Charlotte

Based on the average monthly data, we sum up the values in December, January and February as average winter usage and the values in June, July and August as average summer usage. Figure 16 presents the aggregated mean of winter and summer average across block groups over time. The red curve (for summer) bounces and drops throughout the study period (responding strongly to the trend of PHDI), while the blue curve (for winter) showed larger changes in value during 2000-2003, which was followed by slight variations even during the second drought period. The similar winter usage in 2007 and 2008 again implies the effectiveness of water use restriction policy.

Both the spatial distributions of the average winter and summer water consumptions across the 11 years highlight a concentration of higher water consumption neighborhoods in the southern part of the county (Figure 17) and of relatively smaller water consumption neighborhoods surrounding the urban core. The north of the county (including several towns) had moderate water usage level in winter and mixed water usage levels in summer. The water usage of more block groups within the eastern and western side of the county fell within the classes of higher level in winter, compared to the moderate level in summer.

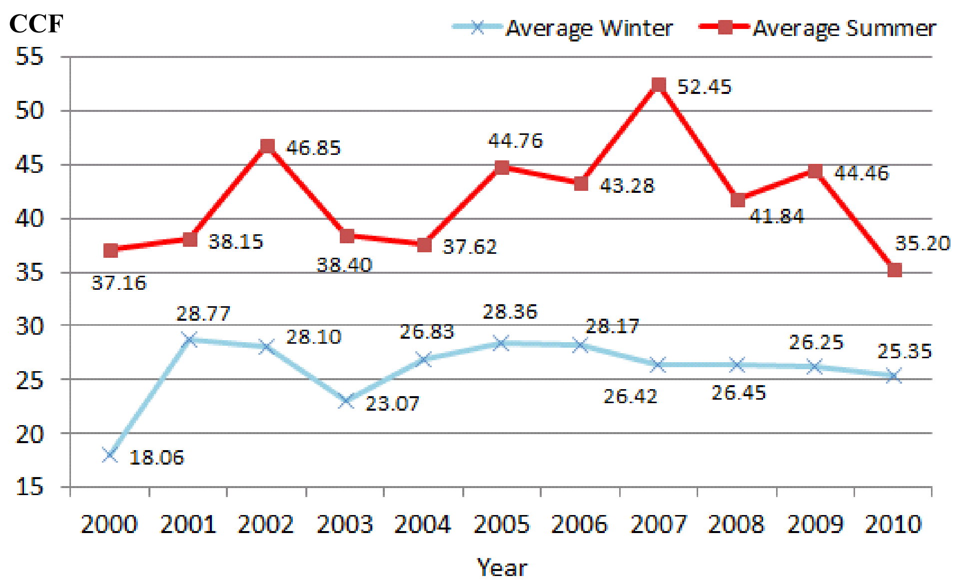
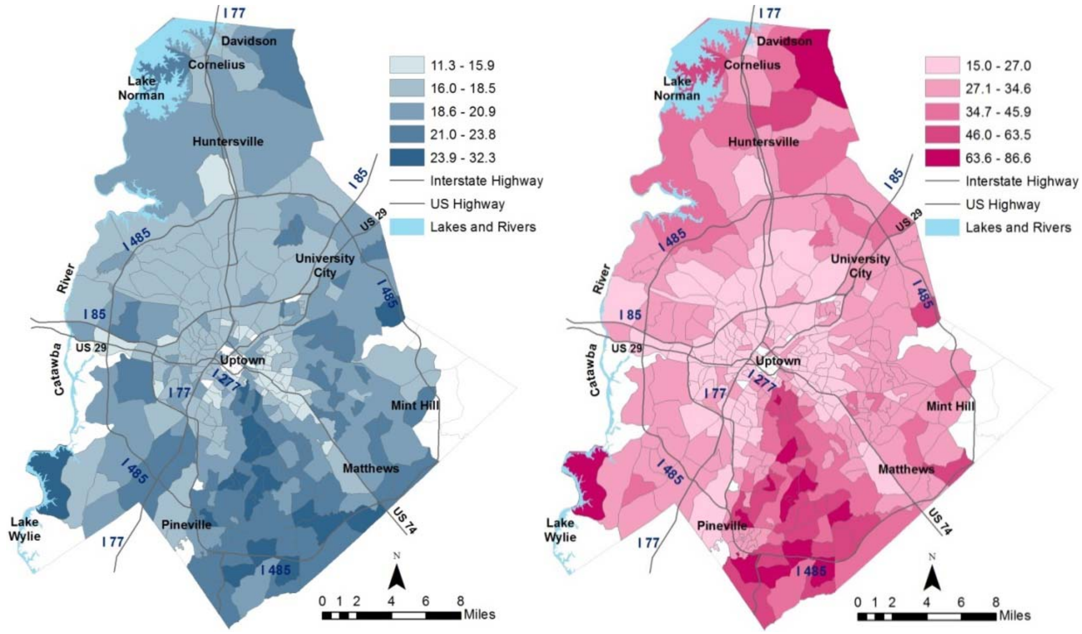


Figure 16: Average winter and summer SFR water consumption per household across block groups, 2000-2010

Comparing the value ranges of the classification of the water consumption for winter and summer, the lower user groups (first three classes) consumed twice as much water in summer as in winter, while the higher consumer groups tripled their average winter water use when summers came. Although we are not sure of the exact proportions of water usage for outdoor and indoor purposes, it is reasonable to assume that outdoor usage contributes a lot to the difference between summer and winter SFR water consumptions. Given that the neighborhoods in different areas have different combinations of water usage levels of winter and summer, we adopt the measure called "peak factor" to analyze geographic patterns in weather sensitivity.



(a) Winter (December, January and February) (b) Summer (June, July, and August)

Figure 17: Average seasonal water consumption per household by block group, 2000-2010

The peak factor is calculated by dividing the summer usage of a block group by its winter usage. If we assume that indoor water usage is invariant throughout the year, outdoor water usage tends to spike in the summer months, thus the ratio of summer versus winter usage can capture the sensitivity to meteorological conditions to a large extent. Considering that sometimes there is a substantial amount of outdoor usage in fall/winter months (for example after overseeding lawn), the peak factor can be regarded as the variable measuring the extra water usage in hot seasons versus cold seasons.

Different aggregation procedures yield similar temporal trends in peak factors across the geographic space (Figure 18). One way is to get the summation of the average summer and winter usage for the entire city, then calculate the ratio (shown as the blue dashed line). This measure is a city-wide indicator. Another way is to get the ratio for each block group first, then calculate the mean of the block group level peak factor (shown as the red solid line). Such a peak factor represents the sample mean.

The first high spike of the peak factor appeared in 2002, following the first drought. The ratio of summer versus water usage reached the highest peak in 2007, and the peak factor in 2008 was lower than in 2007 and 2009. The latter observation seems surprising but is reasonable if the seasonal changes in drought condition are accounted for. The winter of 2008 is in the middle of the process that the region has been becoming more and more drought stricken, and the drought did not worsen in the summer period; instead, the PHDI indices in summer were lower than in winter, and even in August the value was around 3. The phenomenon that the 2009's peak factor was slightly higher than the 2008 peak is counterintuitive given the second drought was recovering from 2008 to 2009. The average summer use in 2009 increased, while the average winter usage decreased a little bit. The possible explanation is that the policy had a positive effect on winter use, and the previous years' drought made people desire more outdoor usage for watering lawn and engaging in water-related activities.

We observe a spatially clustered pattern on the map showing the peak factor averaged for each block group (Figure 19). In general, the neighborhoods in the eastern and western Charlotte and along the I-77 South have lowest values, and the south of the county contains block groups with highest peak factor as well as greatest level of average consumption. The peak factor values were moderately higher in northern Mecklenburg, in which areas a few block groups had relatively large water usage.

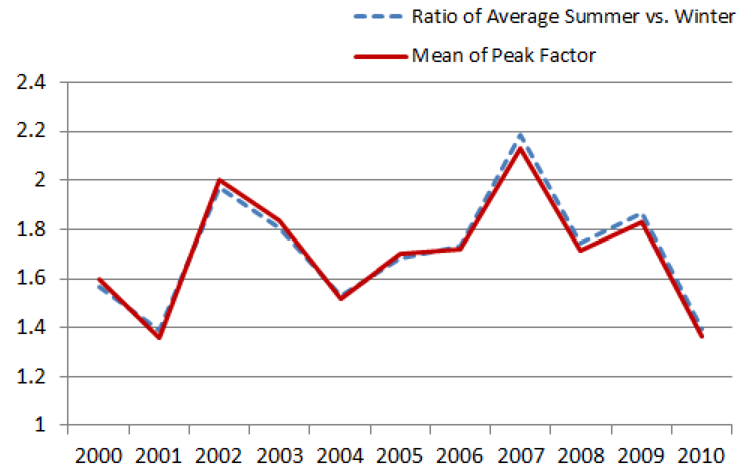
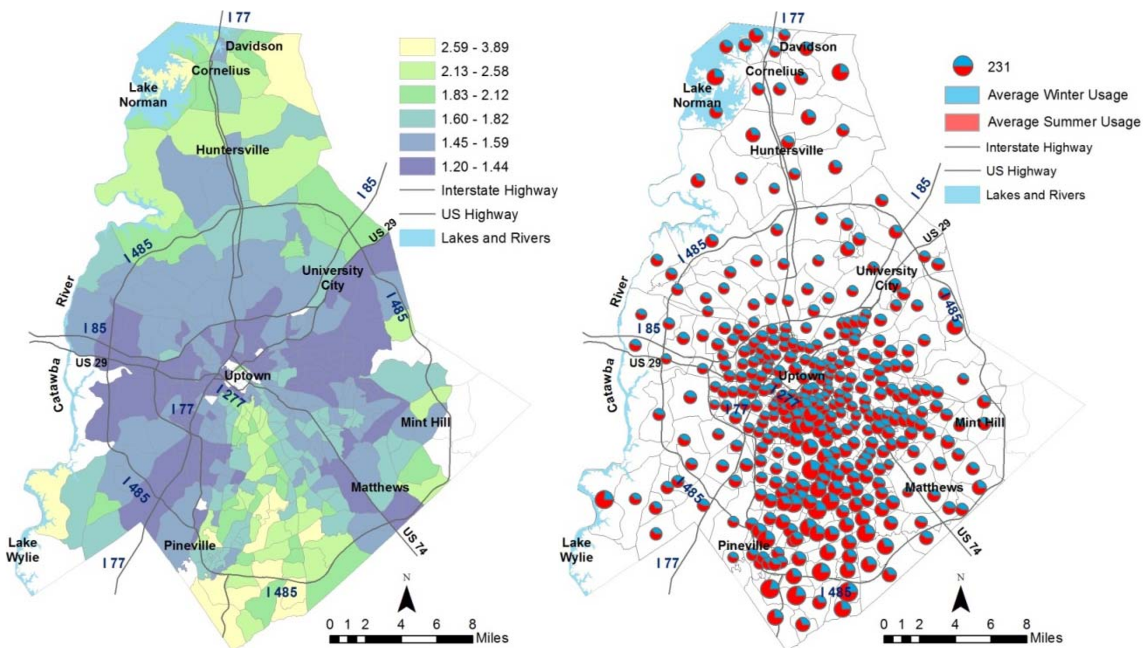


Figure 18: Peak factors across block groups, 2000-2010

(Blue: the ratio of average summer use versus average winter use across block groups; Red: the mean of block group level peak factor that is the ratio of block group's average summer use vs. average winter use)



(a) Ratio of Summer to Winter (b) Winter and Summer Combined

Figure 19: Average winter & summer SFR water consumption per household by block group, 2000-2010

The following series of maps demonstrates the spatial distributions of peak factors before, during and after the two drought periods (Figure 20 and 21). Before the first drought reached its peak, the areas within and bordering the outer ring had decreased peak factor, except a few blocks in the south of the city. In 2002, an increase in the ratio of summer versus winter was prevailing all over the study area. After rainfall returned to normal in 2003, the area with higher peak factor disappeared, instead, the neighborhoods to the east of Independence Boulevard and in the west and northwest of the urban core experienced a slight rise in peak factor.

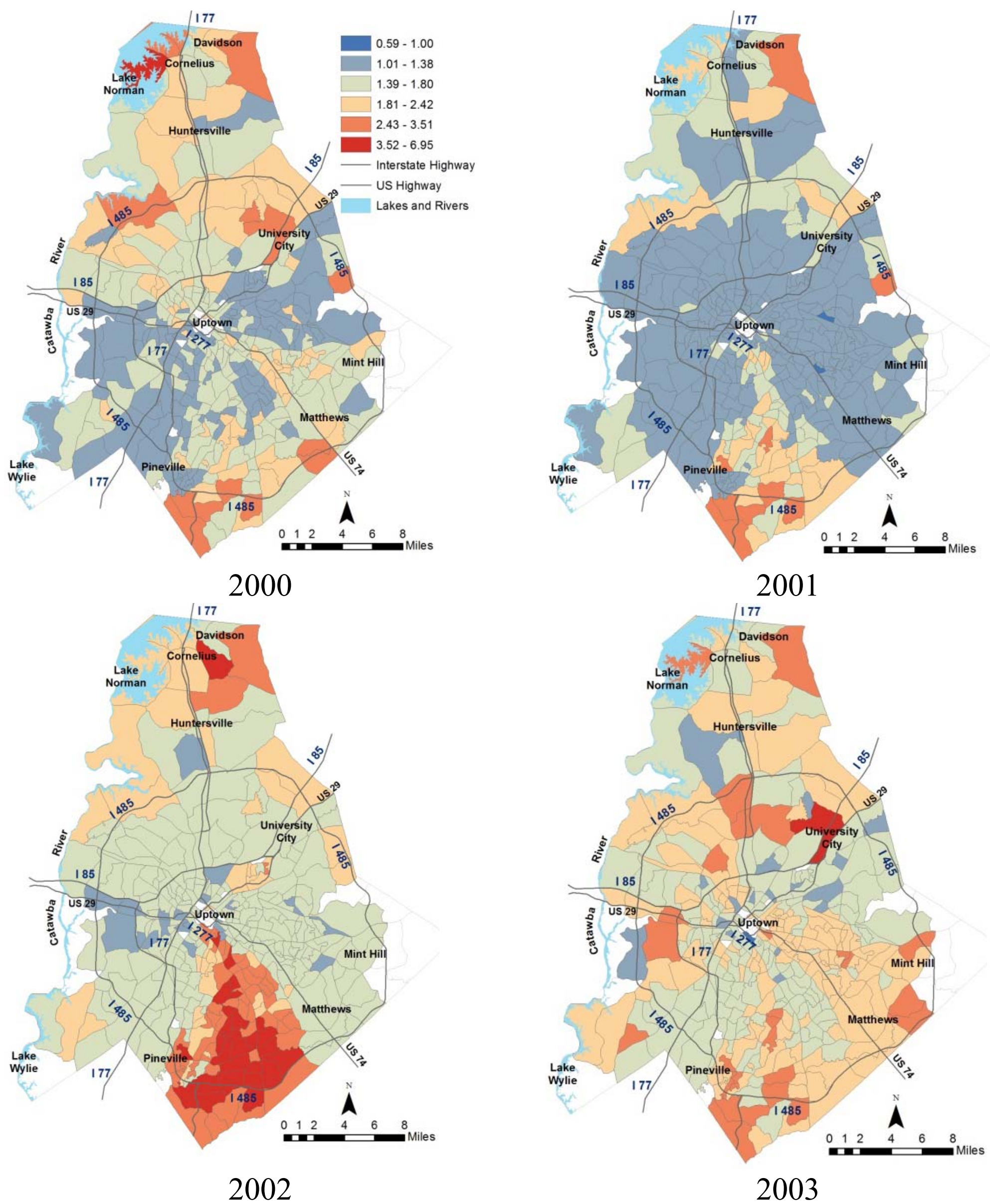


Figure 20: Peak factor by block group, 2000-2003

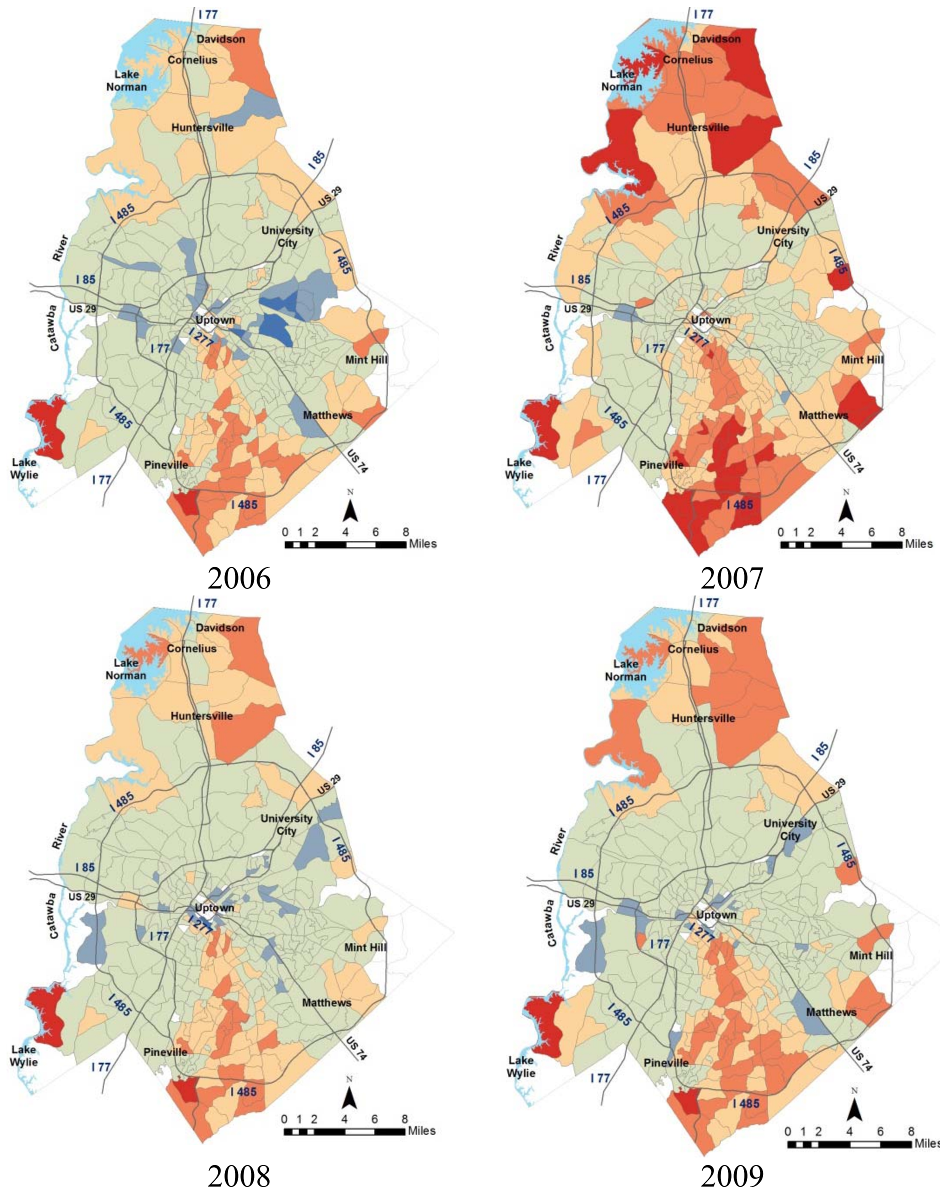


Figure 21: Peak factor by block group, 2006-2009

When the county was exposed to the second drought, interestingly, not all places involved a gain in peak factor. Only the north and the south of Mecklenburg County, and some neighborhoods distributed along the ring I-485 had increased value in 2007. After one year, the values in these areas dropped. A bounce back was observed for the northern and southern edges of the county in 2009. Most of the area within the I-485 ring kept the same level of the ratio before, during and after the second drought period. Relating such pattern changes with the counterintuitive variability of peak factor over time (Figure 18), we conclude that the areas with larger water consumption and more sensitive to drought dominated the overall temporal trend. The spatial and temporal patterns of the peak factor provide graphical evidence for meteorological sensitivity across block groups.

We also calculate the percent changes of winter, summer, and annual SFR water consumption per household from 2002 to 2000 and from 2003 to 2002, and created the cluster maps using queen contiguity-based spatial weight matrix. As shown in Figure 22, during the first drought, the neighborhoods to the north of the city center and along Independence Boulevard see an increase in household water usage in winter, while the residents in south Charlotte use much more water in summer 2002 than their regular usage before the drought. However, these clusters disappeared when the change in annual water consumption was considered. When the 2002 drought transitioned to the extremely moist condition in 2003, southern Charlotte exhibited a large decrease in water usage in summer, but a dramatic increase in winter water usage after the rainfall season past (Figure 23). The high-high clusters of percent change of water usage in summer are shown in the north and the west of the city center, while the low/low clusters of water usage percent change in winter appear along Independence Boulevard. The different clustering patterns described above indicate that households located in different neighborhoods have different preferences to winter/summer water use. Examining the yearly consumption only may mask such a preference difference. The remaining questions are why household preferences differ and what factors are influential.

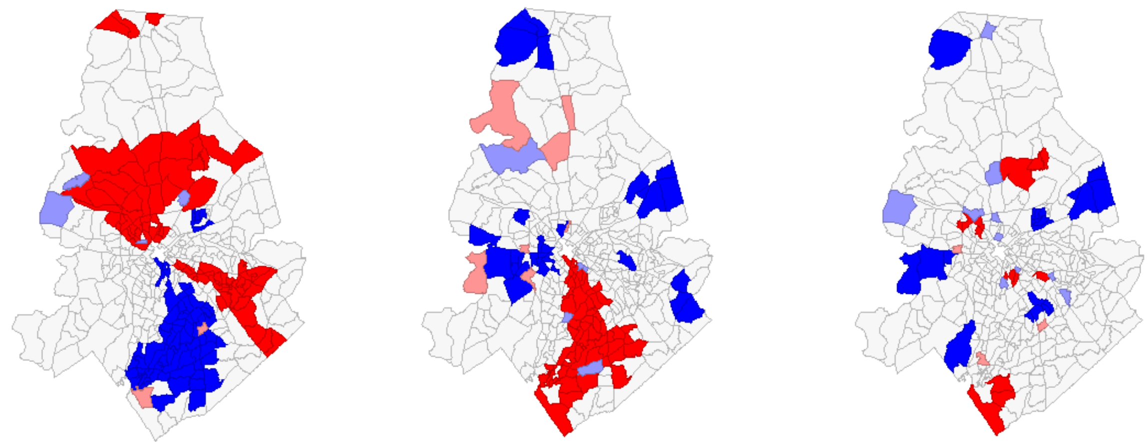


Figure 22: Clusters of the percent change of average winter, summer and annual water consumption per household, 2000-2002

(Red: High-high; Blue: Low-low; Pink: High-low; Purple: Low-high)

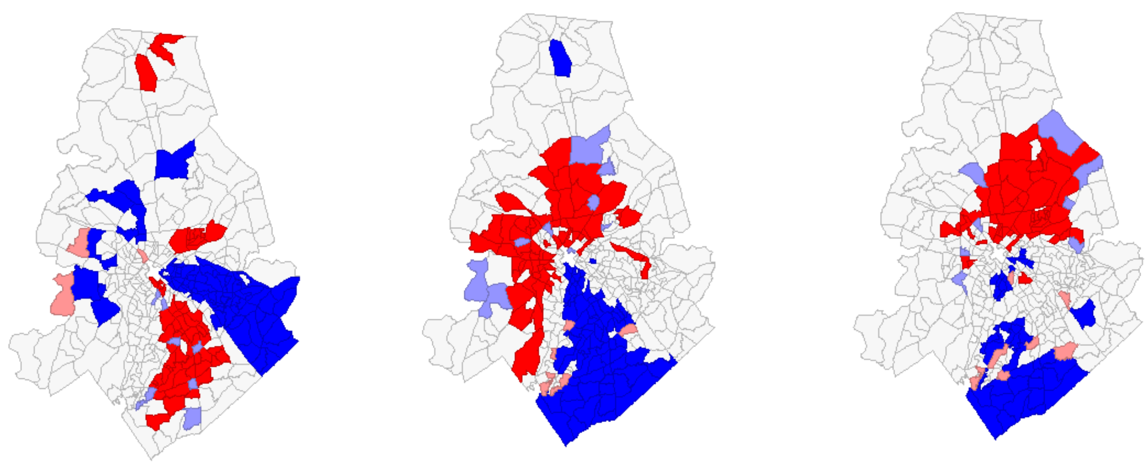


Figure 23: Clusters of the percent change of average winter, summer and annual water consumption per household, 2002-2003

(Red: High-high; Blue: Low-low; Pink: High-low; Purple: Low-high)

4.4 Temporal Analyses of Block Group Level Sensitivity of Water Use to Climatic Variations in Charlotte

Although we have found evidence supporting the hypothesis that neighborhoods in Charlotte responded differently during the drought periods via a mapping of the spatial patterns of average water usage per household living in a SFR home, it is important to test such an assumption in statistical ways. We follow the methodology of the case study of Phoenix (Balling et al. 2008) to test the hypothesis.

All monthly water consumption values in each block group are converted to the deviation from the mean values calculated for each month (mean monthly value) across all the block groups. The purpose of such a conversion is to eliminate the obvious annual cycle (Figure 24) in the monthly data. We do not want this cycle to dominate temporal variance in the water usage values. The new time series data has 132 rows, one for each month from January 2000 to December 2010, and 355 columns, one for each block group. The value is the deviation from normal monthly water consumption for each geographical unit.

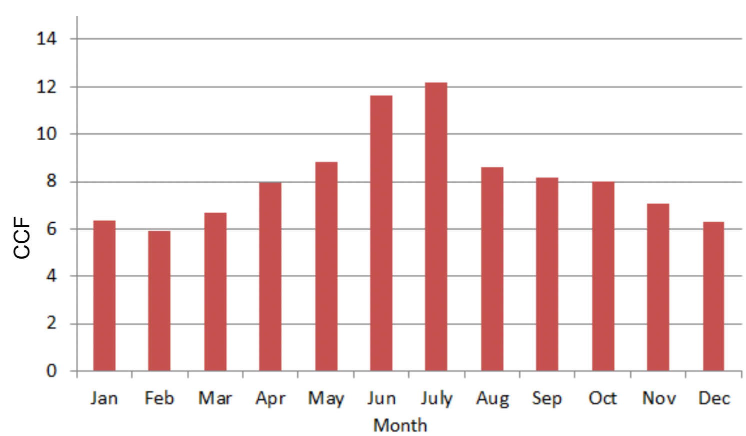


Figure 24: Mean monthly average SFR water consumption per household in Charlotte across the time period 2000-2010

The monthly mean and maximum temperature records show strong cyclic trend within a year (Figure 25). In meteorological studies, anomalies of such variables are usually derived from time series analysis. The anomalies of GHCN monthly temperature ranged from +3.618°C in December 2007 to -4.182°C in December 2010. This means Charlotte had a relatively warmer December in 2007 and a relatively colder December in 2010 when all the December temperatures during the time period 2000-2010 are compared. The range of the maximum temperature anomalies in Charlotte is from +4.218°C in August 2007 to -4.427°C in December 2000.

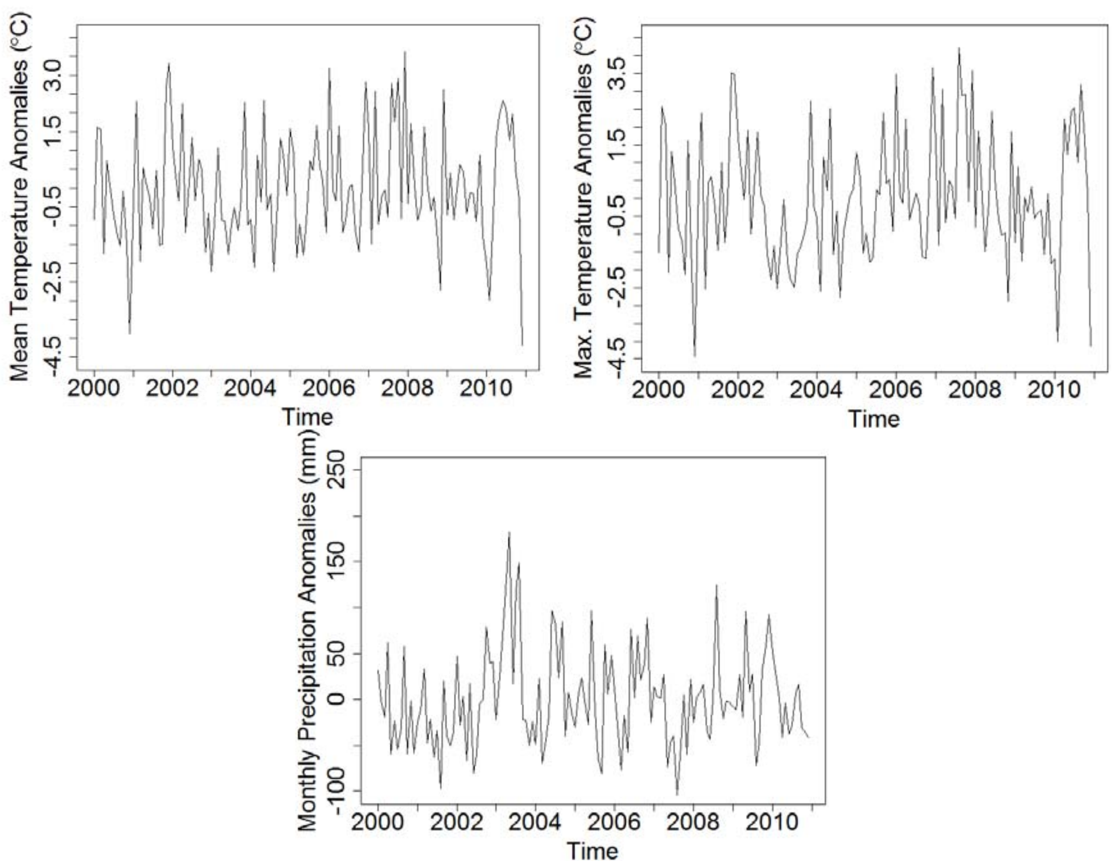


Figure 25: Monthly weather anomalies

(Upper left: mean temperature; Upper right: maximum temperature; Lower: precipitation)

Similarly, we transform the monthly precipitation values to departures from the mean (anomalies). The highest anomaly value (+182.7mm) in Charlotte data occurred in May 2003, when the 2002 drought transitioned to the most moist condition in history. The lowest anomaly value was -103.6mm in August 2007, corresponding to the driest summer time in Charlotte's history when the severest 2007-2008 droughts started. The PHDI data has no annual cycle, thus the original values are used.

For each block group, the monthly deviation values of monthly water consumption are compared with the temperature anomalies, precipitation anomalies, and PHDI values. The three climatic variables share little variance (<15%) over the 132 months and can thus be treated as independent variables. The mean and the range of the Pearson product-moment correlation coefficients (PCC) between the weather dataset of Charlotte and the monthly deviation are listed in Table 6. The results from the USHCN temperature and precipitation datasets are similar, thus not reported here.

Table 6: Pearson product-moment correlation coefficients between climatic variables and the deviation of monthly water consumption

Climatic variables	**Mean**	**Range**	**Num. of BG with the dominant sign**	**Data sources**
Maximum Temperature	0.112	-0.063 ~ 0.377	335 (+)	GHCN
Average Temperature	0.049	-0.123 ~ 0.265	280 (+)	GHCN
Total Precipitation	-0.16	-0.425 ~ 0.164	338 (-)	GHCN
PHDI	-0.175	-0.399 ~ 0.113	343 (-)	USHCN

Note: 355 block groups (BG) in total; the dominant sign in parentheses

The PCCs between the local average temperature anomalies and the monthly deviation from normal water usage range from -0.123 to 0.265, and the average is 0.049. The mean temperature anomalies on average explain only 0.24% of the variance in the monthly deviation of normal water usage. The majority (four fifth) of the block groups have a positive correlation coefficient, and the largest variance explained by temperature is 7%. We also calculate the coefficients between the monthly maximum temperature anomalies and the monthly deviation of water consumption and find that the maximum temperature variable explains water usage a little bit better than mean temperature, and the direction of the correlation for most of the block groups is positive.

Compared to temperature, monthly total precipitation better explains water consumption. Most of the block groups show the negative correlation, and the average PCCs is -0.16. The maximum variance accounted for by the local precipitation anomalies is approximately 18.1%.

The coefficients between the consumption values and PHDI range from -0.399 to 0.113, and the average is -0.175. Although a few block groups have positive PCCs, the water usage of most block groups is negatively correlated to drought condition as expected, meaning that more severe drought conditions stimulate higher water usage across space.

In summary, the PCCs between either of the climate anomalies and water consumption shows the expected sign, although the variances explained are not large. However, the mixed signs of the correlations between temperature (mean or maximum), precipitation or PHDI and water usage implies that different block groups did not respond to these variables in the same way. For some block groups with a sign different from the majority, there are other determinants that may better explain the monthly deviation from normal water usage than a single weather variable. For example, water use restrictions could be enacted during the persistent hot and dry time period, and the residents from some neighborhoods could choose to follow the policy and reduce their water consumption rather than responding to weather conditions, thus those neighborhoods may exhibit opposite associations with weather factors. Another possible reason is that different households may be sensitive to different weather variables (temperature vs. precipitation) since each weather factor is separately examined. Considering that the responses of block groups may be more heterogeneous than of census tracts, the same analyses were applied to monthly deviation of tract-level water consumption, and similar results were obtained.

Comparing to the results reported for the census tracts in Phoenix, the mean PCCs for all the climate variables for Charlotte are smaller (-0.05, -0.16 and -0.175) than the ones for Phoenix (0.32, -0.2, -0.26) as we anticipated. Unlike Charlotte, the water consumption in all the census tracts in Phoenix is correlated with each of the weather measures in the same way in terms of the sign of the PCCs. The range of the absolute values of PCCs is wider for Phoenix than for Charlotte, indicating the variation in the neighborhoods' climatic sensitivity in the arid-climate city is larger.

A multivariate regression analysis is employed to determine, for each block group, the portion of variance in monthly deviation to normal water consumption that can be explained by the temperature (mean or maximum), precipitation and PHDI.[3] The estimated model is:

$$Y = a + b*X1 + c*X2 + d*X3 \qquad \text{(Eq. 3.3)}$$

Where Y refers to the monthly deviation of normal water consumption, X1 is the mean or maximum temperature, X2 is precipitation, and X3 refers to PHDI. Each block

[3] Multivariate regression analysis is also applied to the USHCN data and the results are similar.

group has a different set of values representing its monthly water usage over the time period 2000-2010 (132 observations in total), while the values of the independent variables are the same for all the block groups. From each regression model calculated for each neighborhood, we obtained the coefficient of determination (R^2) as a proxy to measure climatic sensitivity of the corresponding neighborhood. We report here the results from the models using mean temperature, though the results based on the monthly maximum temperature anomalies are not much different.

Summarizing the 355 R-squared values, the average value is 0.06. These R^2 values range from near zero to 0.227. This indicates that some block groups have relatively substantial sensitivity to variations in weather conditions, while others are not sensitive to climatic dynamics at all. Given the 0.05 confidence level, only half of the block groups showed statistically significant association between water consumption and one of the weather variables. However, no block group has significant p-values for the temperature, precipitation and PHDI variables at the same time.

The LISA-based cluster map showing the R-squared values of all the block groups reveals a clustering pattern (Figure 26). The neighborhoods with higher sensitivity to weather conditions are mainly concentrated in the southern and northern portions of the county where higher water consumption is also observed. The block groups (in blue) that are not sensitive to the weather fluctuations are located just to the north of the Uptown area and along Independence Boulevard. Based on the Moran's *I* index (0.29), the global pattern is clustered, and the spatial autocorrelation of the block-group-level climate sensitivity is significant at $p < 0.001$ level of confidence. When only the block groups with R^2 more than 0.05 are considered, the neighborhoods in the furthest south of the county are still identified as the clusters with higher sensitivity, although there are less significant low-low clusters identified.

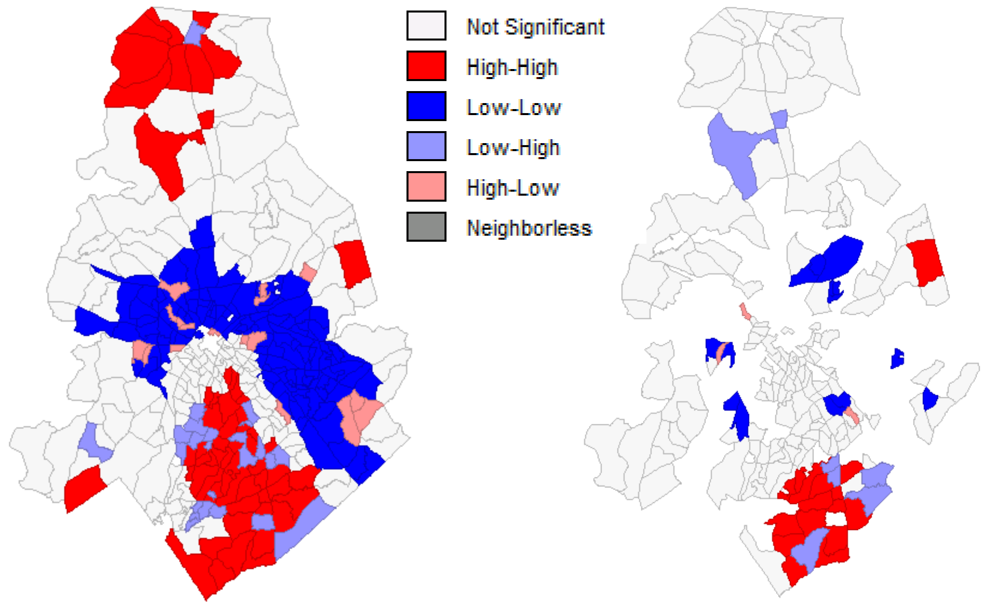


Figure 26: The cluster maps and Moran's *I* plots of the R-squared values as a measure of climate sensitivity from the regressions for block groups

(Left: based on all block groups; Right: the blocks groups with a R^2 larger than 0.05)

4.5 Spatial Analysis on Block Group Level Climate Sensitivity in Charlotte

The purpose of this analysis is to explain the variation in the R-squared values reported above (as the response variable) using sociodemographic and housing factors that could potentially determine water consumption (as the explanatory variables). Considering that some of the explanatory variables do not follow a normal distribution and there is no need to assume a linear association between the dependent and independent variables, spearman rank-order correlation coefficient (or called spearman rho) is employed to describe the relationship between the derived climatic sensitivity index and the selected factors using a monotonic function. We first test the spearman rank-order correlation on all the sociodemographic variables in 2000 or 2010 and the housing variables derived from 2008 parcel data. The correlation coefficients between the explanatory variables and the climatic sensitivity index (measured by R^2 from the last step) are listed in Table 7. The second and third columns show the spearman correlation coefficients and the p-values for the 2000 census variables and the housing variables in 2008, and the last two columns are for the 2010 census variables.

Table 7: Nonparametric Spearman rank-order correlation coefficients between water consumption and the explanatory variables representing sociodemographic and housing characteristics of neighborhoods

Category	Variables	Spearman correlation (2000/2008)	p-value (2000/2008)	Spearman correlation (2010)	p-value (2010)
Housing	Number of single-family houses built after 1992	0.334	0.000		
	Livable area	0.589	0.000		
	Area of lot	0.340	0.000		
	Assessed property value	0.613	0.000		
	Number of bathrooms and bedrooms	0.537	0.000		
	Average percent of households with a pool	0.351	0.000		
Socio-economic	Percent of population 19 years and under	-0.185	0.000	-0.132	0.013
	Percent of population 65 years old and over	0.108	0.042	0.126	0.017
	Household size	-0.115	0.030	-0.121	0.022
	Percent of population whose ethnicity is Hispanic	-0.365	0.000	-0.332	0.000
	Percent of owner-occupied housing units	0.304	0.000	0.343	0.000
	Median number of rooms	0.372	0.000	0.423	0.000
	Per capita income	0.567	0.000	0.524	0.000
	Median annual household income	0.504	0.000	0.507	0.000
	Percent of population 25 years old and over obtaining a college degree	0.499	0.000		
	Population density	-0.232	0.000		
	Ratio of own children under 18 years old over the number of households	-0.028	0.595		

All the housing factors have a significant correlation with the variation in the climatic sensitivity of water consumption across the county. A negative correlation for a variable implies that climatic sensitivity is higher for lower values of the variable. Thus, water consumption in block groups with larger and newer SFR houses, larger lot, higher house value, or more bathrooms and bedrooms is more sensitive to weather conditions. Most the sociodemographic variables from 2000 are significantly related to the dependent variable at the 0.05 level, except for the ratio of own children under 18 years old over the number of households. Weather sensitivity is higher for a lower percentage of population younger than 19, smaller family, a lower percentage of Hispanic population, and less dense neighborhood (in terms of population). When the percentage of owner-occupied housing units, the median number of rooms, the per capita income, or the median household income are larger, water consumption is more sensitive to weather conditions. With the added variable representing the percentage of parcels with swimming pool within a block group, weather sensitivity has a positive relationship. This means that block groups with a higher percentage of households with pool are susceptible to climate variation. The Spearman correlation coefficients calculated for 2010 sociodemographic variables are similar in terms of the magnitude, direction and significance as those obtained with 2000 variables.

Comparing with the correlation test results for Phoenix, we found that the correlation coefficients are generally higher for the case of Charlotte, and the household size variable is also a significant predictor to the index of sensitivity of water consumption to weather conditions. As we hypothesized, the household and housing factors that are closely linked to outdoor water usage such as area of lot, percent of households with a pool and household income are significantly correlated with the sensitivity being investigated, although the magnitude of their correlation coefficients are not highest.

As a nonparametric statistical method, Spearman rank-order correlation test cannot assess the associations between the dependent variable and multiple independent variables at the same time. One solution to this limitation is to generate components (or composite indices) from a group of factors and then evaluate the association between each component and the response variable of interest. We chose a few variables that are representative of water usage determinants (especially influential to outdoor usage), have a statistically significant Spearman correlation coefficient, and are comparable to the variable selection in the case study of Phoenix. These variables include area of lot, household size, percent of Hispanic population, median household income, and percent of parcels with a pool.

We apply principal components analysis[4] on these variables and obtained three factors with eigenvalues greater than 0.95. Together the three components explained 78.5% of the variation in the data (Table 8). The first component (mainly loading on income and percent of parcels with pool) represents the wealth of households, the second one is for household size (consisting of household size and percent of Hispanic population), and the variable area of lot alone is the third component. The results in Table 9 show that water consumption of affluent households (component 1) in 2000 or 2010 is more sensitive to weather conditions. The relatively strong relationship between climatic

[4] The principal components analysis method we applied does an eigen value decomposition, based on the correlation matrix of the variables involved. It is implemented using the "principal" package in R (Jolliffe, I. (2002) Principal Component Analysis (2nd ed,). Springer).

sensitivity and component 1 reinforces the associations individually observed with income and percent of parcels with pool. However, component 3 (area of lot) has a negative Spearman correlation coefficient, indicating the neighborhoods with smaller lot are more sensitive to weather conditions. This is opposite to the results when the single variable (area of lot) was investigated. The association of the household size component (component 2) with the R-squared values is negative but insignificant for the data in both years. Note the signs of the estimates of the component scores are different (positive in 2000 and negative in 2010), thus the interpretations are opposite as well - larger family in 2000 is less sensitive, while larger family in 2010 is more sensitive. This may be due to the change in household composition in terms of number of occupants within the last decade (households are becoming smaller).

Table 8: Unrotated principal component loadings and communalities

Variable	2000/2008				2010/2008			
	Component 1	**Component 2**	**Component 3**	**Communality**	**Component 1**	**Component 2**	**Component 3**	**Communality**
Area of lot	0.227	0.067	0.965	0.988	-0.188	-0.323	0.928	0.999
Household Size	-0.017	0.796	-0.118	0.648	0.248	-0.708	-0.197	0.601
Percent of Hispanic population	-0.38	0.554	0.134	0.462	0.491	-0.458	-0.06	0.455
Median annual household income	0.654	0.177	-0.148	0.481	-0.609	-0.246	-0.214	0.477
Average percent of households with a pool	0.613	0.153	-0.119	0.414	-0.540	-0.352	-0.227	0.467
Eigenvalue	1.786	1.178	0.96	3.924	1.784	1.219	0.941	3.944
Percent Variance	0.357	0.236	0.192	0.785	0.357	0.244	0.188	0.789

Table 9: Nonparametric Spearman rank-order correlation coefficients between water consumption and the principal components

Variables	Spearman correlation (2000/2008)	p-value (2000/2008)	Spearman correlation (2010)	p-value (2010)
component 1	0.523	0.000	-0.525	0.000
component 2	-0.067	0.210	-0.061	0.255
component 3	-0.225	0.000	-0.150	0.005

4.6 Conclusions and Future Work

In this chapter, we first analyzed the temporal trend and spatial patterns of water consumption during 2000-2010 and found that average annual SFR water consumption per household is not evenly distributed across the study area and the temporal variations of water consumption are consistent with the weather dynamics in terms of mean temperature, total precipitation and the drought index PHDI. A closer look at the average summer and water usage at the block group level reveals that certain neighborhoods (especially in southern Charlotte) consumed much more water in summer than winter, possibly due to their households' desire for lush lawn and/or outdoor activities. Although these neighborhoods would be more vulnerable to the drought conditions, they have greatest potentials in terms of water consumption reduction and could be targets for discovering the underlying behavioral processes, and for experimenting policy interventions and conservation strategies.

Furthermore, when the year-to-year changes in water usage are examined, we observed the heterogeneous responses at the neighborhood level to weather variations.

The analyses that followed specifically focus on the evaluation of the sensitivity of water consumption to atmospheric conditions. The major findings are: (1) despite the general low level of sensitivity compared to Phoenix, certain neighborhoods still exhibited mild but statistically significant sensitivity; (2) the temporal variations in water consumption explained by the three meteorological variables tend to cluster across geographical units; (3) high climatic sensitivity occurred in the neighborhoods with larger lots, more parcels with pools, larger and newer SFR houses, higher house value, or more number of bathrooms and bedrooms. Higher income neighborhoods or more owner-occupied areas are associated with larger sensitivity. Climatic sensitivity decreases with a high proportion of population younger than 19 years old, larger family, more percent of Hispanic population, or denser neighborhood.

This study has several limitations. First, correlation analysis (no matter Pearson or Spearman) cannot examine the relationship between dependent variable and multiple explanatory variables (no matter raw variables or principal components). Even if a multivariate statistical model is considered, we cannot avoid the issues such as distributional assumption and multicollinearity. Many machine learning techniques are largely free of these issues, and can be used to explore the complexity in the relationships between climatic sensitivity and household and housing characteristics in future. Second, the relatively low climatic sensitivity of the neighborhoods in Charlotte indicates that there could be other factors contributing the heterogeneous responses of neighborhoods to weather variations, such as water use restrictions or conservation campaigns taking effect during the droughts. We may need to capture the interaction of the weather variables and these factors when evaluating the sensitivity being investigated. Third, monthly accumulative precipitation has a sporadic nature - a rain event could occur late in the month and produce a positive monthly anomaly, when in reality, the bulk of the month was actually dry and a large amount of water had been consumed before the rainfall came (Balling et al. 2008). So we can substitute monthly precipitation with other rainfall-related variables to be derived from daily weather datasets. Other weather variables such as evapotranspiration can be tested as well. Fourth, few of the socioeconomic and housing factors we used to explain climatic sensitivity are relevant to indoor water consumption. If we can collect the household level data regarding water consumption habit and water-saving fixtures or appliances for a long time period, we could conduct sensitivity analysis at the household level thereby better understand the heterogeneity of climatic sensitivity.

CHAPTER 5: UNDERSTANDING SFR WATER USE IN CHARLOTTE FROM A SOCIO-ECOLOGICAL PERSPECTIVE

5.1 Introduction

Urban water usage or demand has been a subject of interest of scholars and professionals from various domains, across which the meaning of urban water usage or demand varies. It is a technical term for hydrological engineers, or an economic concept (demand), given that water is a commodity. It could be referred to as a quantity, an outcome from individuals or households' behavior, or a behavioral process per se. We acknowledge the systematic view on urban water usage/demand as a coupled human and natural system (also known as social-ecological system). In this system, individuals or households dominate the internal processes, including the actions of consuming and conserving water and making decisions for what purposes, when and how much water is used. Nevertheless, the external processes that are going on in the biophysical, social and economic environment can interfere with water usage behaviors, thereby influencing demand for water. Outdoor water usage can be greatly affected by the interactions between the household preferences for lush lawns and healthy vegetation and the natural processes related to water cycle (Figure 27), such as soil types, rates of evapotranspiration, precipitation, and types of vegetation. The social processes such as the implementation of price, non-price water policies and conservation campaigns at the municipal or regional level, and neighborhood-level rules and regulations (for example homeowner association's (HOA) mandatory lawn maintenance policies (Harlan et al. 2009) could constrain or facilitate water consumption. The complex interactions between human and natural system variables at multiple spatial and temporal scales make it complicated to predict and manage urban water demand (House-Peters and Chang 2011).

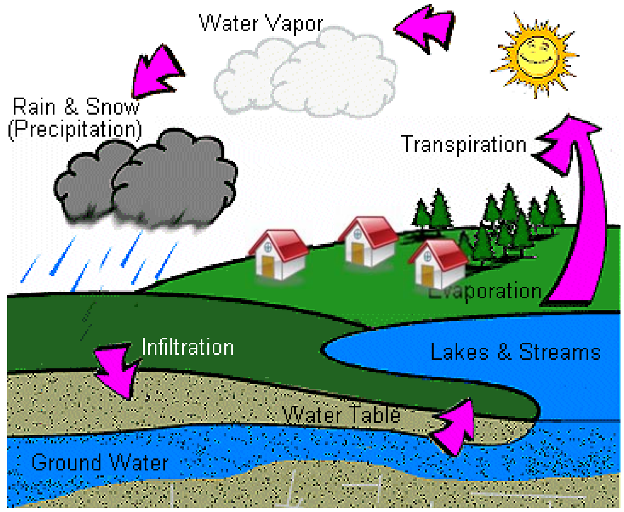


Figure 27: Water cycle

Source: Environmental Education for Kids by Dept. of Natural Resources, Wisconsin, US
dnr.wi.gov/org/caer/ce/eek/earth/groundwater/watercycle.htm

A pertinent framework has recently been suggested to examine the complexity involved in socio-ecological systems. Integrating different perspectives of system structure from hierarchy theory (Allen and Starr 1982), landscape ecology (Forman 1995; Pickett and Cadenasso 1995; Wiens 2000), and patch dynamics (Pickett and White 1985), Cadenasso et al. (2006) posited the existence of three dimensions of complexity, including spatial heterogeneity, organizational connectivity, and historical contingency. Spatial heterogeneity is often described as patches (discrete areas that differ in structure, composition, or function) in the ecology domain. In addition to land cover and land use patches, we can rely on various human and natural variables to construct patches, such as soil permeability, zoning and census geographic units (Cadenasso et al. 2006). From the regional science perspective, spatial heterogeneity refers to the uneven distribution of a trait, event, process, or relationship across a region (Anselin 2010). Organizational connectivity is about the processes, interactions and organizational structures related to the spatially arrayed components like patches. Historical contingency refers to relationships that extend beyond direct, contemporary ones. Indirect effects, lagged effects, and the impacts of past states of the system all are identified as historical contingency with higher complexity (Grove et al. 2012). As you can see, this complexity framework essentially emphasizes space, organization/structure, and time as necessary components for modeling and understanding socio-ecological systems (Cadenasso et al. 2006). These dimensions are also important for exploring urban water usage as a socio-ecological system.

Although being stored in water pipes lying underground, the water used by urban residents is essentially a kind of resource that is processed and ready-to-use. Consuming water is essentially the process of exploiting the processed natural resource, similar to the development of natural land into different uses (the difference is that water use does not physically exist as land use does). Since water consumption can be spatially attached to parcels, census geographies or other spatial units, we can conceptualize the distribution of urban water use as a virtual 'waterscape', and the basic components of waterscape would be water patches represented by the geometric shape of parcels or census geographies. In this way, we can easily apply the framework described above to study the complexity in urban water use along the three dimensions, that is spatial heterogeneity, organizational connectivity, and historical contingencies.

Research on water use started paying attention to spatial heterogeneity of residential water use only in the recent years. When the spatial distribution of residential water consumption was examined at either the census tract or block group scale, an evident clustering pattern was observed across all of the case studies (Guhathakurta and Gober 2007; Wentz and Gober 2007; Balling et al. 2008; Lee and Wentz 2008; Chang et al. 2010; House-Peters et al. 2010; Polebitski and Palmer 2010; Breyer et al. 2012; Ouyang et al. 2014). Advanced statistical models were employed in a few studies to reveal the role of spatial autocorrelation and spatial dependence (a spatial form of connectivity) in explaining the variation in water consumption. Using a geographically weighted regression (GWR) model, Wentz and Gober (2007) found there is spatial variation in the relationships between SFR water usage and four factors including household size, the presence of swimming pools, lot size, and the prevalence of landscaping; also neighboring census tracts in Phoenix respond similarly to changes in the significant independent variables. Chang et al. (2010) compared the results from the traditional OLS regression model for SFR water consumption per household and the spatial error model and concluded that the influence of building size and age was overestimated by the model without accounting for spatial dependence among residuals.

In terms of time dimension in the water use literature, three ways are commonly seen. One is to conduct analyses for multiple time points (usually yearly) separately then make comparison (of pattern or estimation) (House-Peters et al. 2010; Aggarwal et al. 2012) or calculate the mean estimates (Polebitski and Palmer 2010). Time series models are often developed at a coarser spatial scale (municipality or community) (Maidment and Parzen 1984; Gutzler and Nims 2005) and/or smaller temporal scale (daily and monthly) (Billings and Agthe 1980; Rhoades and Walski 1991). Moreover, several studies have used panel data approaches to incorporate temporal and subject-based variability for better estimates (Nieswiadomy and Molina 1989; Schneider and Whitlatch 1991; Höglund 1999; Kenney et al. 2008; Polebitski and Palmer 2010, to list a few). However, the exclusively time series analyses are limited by the (un)availability of longitudinal data for explanatory variables (House-Peters and Chang 2011), not to say the spatially explicit time series analyses of demand.

Most of the studies on water demand model have examined the effects of its determinants measured at the same time point as the dependent variable – water use. However, given the social processes involved in water consumption behavior (for example the desire of a community/neighborhood for a lush green landscape) and the slow changing process of land uses and housing development, the water usage phenomenon may exhibit some form of contingency, meaning that the current neighborhood-level water use is possibly dependent on the historical state of water usage (path dependence theory). Although adding a time lagged water consumption variable could directly capture such a historical contingency, this does not best serve the purpose of explaining the phenomenon and supporting policy-oriented decision-making. Instead, to examine separately the effects of the historical condition of associated factors and the effects of their changes on water use in a statistical model enables the question "whether or not the historical state of a factor is more important than its temporal change, or the two are equivalently important" to be answered, thus different policy implications could be provided accordingly.

This point is emphasized in Ouyang's research (2013). Using Phoenix as the case study, he studied the historically contingent effects of selected determinants on residential water consumption and found housing, household, and climate factors in the earlier year exhibit more important effects than their temporal change (between 2000 and 2009) on the SFR water usage. They argued that the studies quantifying spatial patterns did not consider how spatial dependence at one time manifests into future patterns of water usage. Thus they further quantified spatial heterogeneity and connections in the historically contingent effects they identified. This research represents an interesting application of the complexity framework of socio-ecological systems suggested by Cadenasso et al. (2006) in understanding the spatio-structural and temporal dimensions of urban water usage (Ouyang 2013), although the heterogeneity, connectivity, and contingency three dimensions are not investigated in parallel but in a nested way.

Following the same path, we propose to examine the complexity in SFR water usage in Charlotte, NC under this complexity framework. To approach this goal, we set three objectives: (1) Examine whether the socio-economic factors (household and housing) in history (the year 2000) and their changes between 2000 and 2008 are associated with SFR water usage, (2) Determine whether the associations between SFR water usage and these factors demonstrate spatial variability, (3) Explore the spatial dependence of the empirical relationship between SFR water usage and these factors.

This study focuses on yearly SFR water use at the block group scale. We apply OLS regression to quantify the relationship between annual SFR water usage per household and the previous status of household and housing factors and their changes in quantity between 2000 and 2008. The spatial heterogeneity and connectivity in the contingent effects of the determinants of water usage are then examined to help understand the location effects that influence the spatial variability of water consumption across the neighborhoods being studied. We use spatial statistical model and Geographically Weighted Regression (GWR) model to study these two aspects.

5.2 Data and Variables

Usually when the association between a dependent variable and its determinants is being studied, the contemporary states of the determinants are employed. However, considering the determinants may change over time, the effects of associated factors can be decomposed into two components; one is from their historical state, and another is their temporal change (the change between the current state and the historical state). A comparison of these two types of effects will result in three possible outcomes. Either the historical states of associated factors Xs are more (larger effects) or less (smaller effects) important than their temporal change (referred to as ΔXs), or the two are relatively of the same importance. In this way, the historically contingent effects of associated factors can be examined. Based on the three outcomes described above, Ouyang (2013) classified a system of residential water usage into three types (Types I, II, and III). In a Type I system, the historical states of associated factors Xs outweigh their temporal change) in influencing current residential water usage. The opposite condition characterizes a Type II system. In a Type III system, both historical state and temporal change affect water consumption in the similar magnitude. In order to assess what type of system the SFR water use in a study area belongs to, we need to compile datasets for two time points and calculate the temporal changes of each variable of interest.

Given that only parcel and building datasets in 2008 (used for deriving the factors related to housing characteristics) are available, we choose the annual SFR water consumption per household in 2008 as the dependent variable (Table 10). The earliest complete billing records we obtained are for the year 2000, which becomes the historical time point for examining the temporal contingency.

Housing variables included in this analysis are average house age, percent of units with a pool, area of irrigable land, and housing density. The housing density is calculated by dividing the total number of residential units within a block group by the total area of the block group. For the other three factors, values were first calculated for the premises that are located on the one-unit parcel and have non-zero water consumption in all of the twelve months of a certain year, and then averaged to obtain the measures for block groups. The specific definitions of these three variables are given in chapter 2 (in section 2.3.2). Note that the values for the two variables, percent of units with a pool and area of irrigable land, are estimated and errors may be non-trivial. The information about the impervious area of a parcel and the exact number of stories of a house is estimated, which are the important pieces for calculating irrigable land area. Although we are able to retrieve the parcel attribute about the pool from the table Special Features & Yard Items in the Mecklenburg County Assessor Database in 2006 and assume the existence of a swimming pool will not change once built, it is common in reality that the swimming pool of a house was removed or filled or built when or after the house was sold and re-occupied, and unfortunately such a change is not tracked by appraisers all the time. Thus the information on the presence of a swimming pool may not always be accurate, and so would the derived percentage variable.

The housing variables for 2000 are calculated based on buildings constructed before 2001. The reference year for the calculation of the house age variable is 2009, which means a housing unit built in 2008 is one-year-old.

For socioeconomic characteristics of households, we use decennial census 2000 and 2010 datasets. Although the 2008 household characteristics at the block group level may be different from the ones recorded in 2010, we could not find a dataset that would be closer to the 2008 reality and as reliable and accurate as 2010 decennial census data. We prepare a list of census variables that have been found to be empirically associated with water consumption, but only include median household income, the size of household, and the percentage of owner occupied housing units in the final analyses, each of which shows a strong and significant association (at the 0.05 level) with the dependent variable when being tested individually and does not cause multicollinearity when being added to the multivariate regression model.

To calculate the value change of census variables from 2000 and 2010, we have to resolve the issue of spatial mismatch since the boundary of some block groups may have changed. For this purpose, we modify the block relationship file provided by the Census Bureau, based on which the 2010 block group data (continuous variables) are adjusted to the 2000 block group boundaries.

Once the datasets for 2000 and 2008/2010 are created, the difference between the two values for each independent variable can be computed to represent the temporal change. We adjust the average age of buildings in 2000 to the year of 2008 by adding 8 years for each block group so that its temporal change reflects actual new housing development.

We cannot include climate factors because the entire county only has one observation station; thus temperature and precipitation for a certain year are constant across block groups. There are fourteen block groups that have less than 20 parcels with non-zero water consumption either in 2000 and 2008. We exclude them from the analyses to reduce the bias of the measures that may be introduced by small samples.

Table 10: Descriptive statistics of the dependent and independent variables

Variables	Units	Mean	Std. dev.	Min	Max
SFR Water Use per Household	CCF	104.442	26.554	64.452	237.92
Household Size in 2000	person	2.509	0.393	1.29	3.63
Change in Household Size		-0.027	0.242	-1.417	0.799
Median Household Income in 2000	1999 $	53950.44	28652.33	7717	200000
Change in Median Household Income		-19130.2	16182.95	-127849	20894
Housing Density in 2000	unit per square km^2	245.764	152.506	2.659	735.691
Change in Housing Density		27.708	35.595	0	242.953
Average Age of Buildings	year	32.78	17.391	2.141	82.3
Change in Average Age of Buildings		-3.733	5.866	-43.606	5.937
Percent of Households with Pools in 2000	%	2.454	3.134	0	21.74
Change in Percent of Households with Pools		-0.009	1.133	-10.462	4.731
Area of Irrigable Land in 2000	square feet	15689.25	8604.025	3095	92484.16
Change in Area of Irrigable Land		-630.94	4793.878	-67618.1	14867.16
Percent of Owner-Occupied Housing Units in 2000	%	62.985	25.205	4.829	99.327
Change in Percent of Owner Occupied Housing Units		-3.537	10.817	-64.3412	42.664

5.3 Methods

Ordinary least squares (OLS) regression, spatial regression, and geographically weighted regression (GWR) are used to determine the relationship between SFR water consumption per household in 2008 and the explanatory variables (see Table 10). Both OLS and spatial regression models assume a global relationship across spatial units (spatial stationarity), while GWR assumes spatial non-stationarity in the relationship.

An OLS regression model is defined as:

$$Y_i = \beta_0 + \beta X_i + \varepsilon_i \qquad \text{(Eq. 4.1)}$$

where Y_i is the SFR water consumption per household in block group i ($i = 1, \ldots, N$), X_i is a vector of observations on the explanatory variables in block group i, β_0 is the intercept, β is a vector of coefficients for the explanatory variables, and ε_i is the error term. ε_i is assumed independently and identically distributed for all i ($\varepsilon_i \sim$ N(0, σ_ε^2)). However, spatial autocorrelation that possibly exist in the data will violate the independence assumption of OLS regression. Spatial regression models can explicitly account for spatial autocorrelation in the form of lag or error dependence (Ward and Gleditsch 2008).

There are two basic types of spatial regression models, spatial lag and spatial error models. In spatial lag models the dependent variable is affected by the dependent variables in adjacent places, while in spatial error models, the error terms across different spatial units are correlated. The spatial lag model is expressed as:

$$Y_i = \beta_0 + \rho W Y_j + \beta X_i + \varepsilon_i \qquad \text{(Eq. 4.2)}$$

where ρ is the spatial autoregressive coefficient, W refers to a spatial weight matrix, and the other notation is as before. The spatial error model is specified as:

$$\begin{cases} Y_i = \beta_0 + \beta X_i + u_i \\ u_i = \lambda W u_j + \varepsilon_i \end{cases} \qquad \text{(Eq. 4.3)}$$

where u_i is a vector of the spatially correlated error terms, λ is the spatial autocorrelation coefficient.

Two test statistics—Lagrange Multiplier (LM) statistic and Robust LM statistic can be used to test whether the spatial lag or spatial error model should be used instead of

OLS. The LM statistic from our analysis suggests that the spatial lag model is a better option to incorporate spatial dependence. A queen contiguity spatial weight matrix is constructed for the spatial lag model.

In both spatial lag and error models, the parameter estimates of the independent variables (β) and the intercept β_0 are the same for all spatial units (block group). GWR differs from spatial regression models in that it calculates a unique set of parameters for each spatial unit, defined by geographic coordinates (u, v). A GWR model for spatial unit i is defined as:

$$Y_i = \beta_0(u_i, v_i) + \beta(u_i, v_i)X_i + \varepsilon_i \qquad \text{(Eq. 4.4)}$$

As seen from the Equation 4.4, the coefficients $\beta_0(u_i, v_i)$ and $\beta(u_i, v_i)$ of block group i are dependent on its location. GWR runs a local regression model for each spatial unit from the values of the dependent and independent variables at that location and weighted values (typically weighted by an inverse distance) of neighboring units. Nearby values either can be sampled from a fixed distance from the observation (called a fixed kernel) or a varying distance (called an adaptive kernel). The optimal distance (for a fixed kernel) or optimal number of neighbors (for an adaptive kernel) is determined using an (minimize) Akaike Information Criterion (AIC) or a (leave-one-out) cross-validation approach.

GWR generates coefficients, standard errors, t-scores, and r^2 values for each spatial unit. We can map these results to show the heterogeneous effect of each independent variable on water use in 2008 varying by location. Thus GWR is a useful tool in depicting spatial heterogeneity in the relationships between block-level SFR water usage per household and the determinants of water consumption.

5.4 Regression Analysis

The average water usage per household in 2008 ranges from 64.45 to 237.92 CCF (1 CCF is equivalent to 748.05 gallons, thus 48,211.8 to 177,976.1 gallons), but 90% of the values fall below 139 CCF (103,979 gallons). The global Moran's *I* scores (in GeoDa) for 2000 and 2008 water usage are 0.617 and 0.584 respectively ($p < 0.001$), which indicates that water consumption is spatially autocorrelated across the block groups and this spatial dependence persists over time. The block groups with higher water usage in 2008 are mostly located in southern Mecklenburg where the neighborhoods are relatively more affluent (Figure 28). In contrast, the low or lower water consumption neighborhoods concentrate in the eastern periphery of Uptown as well as the areas near and surrounding the urban core. These areas are also associated with lower household income.

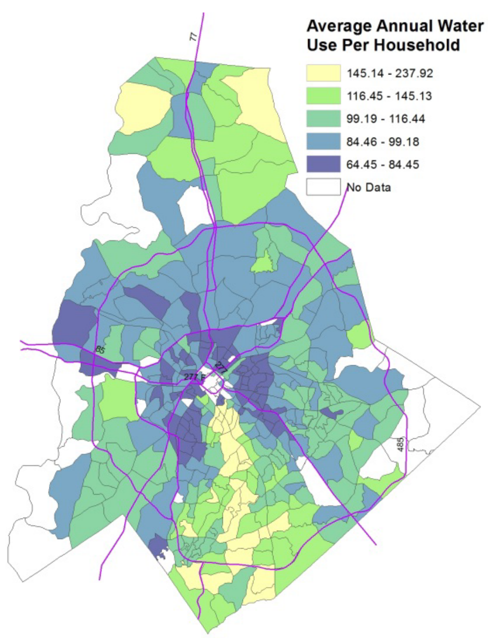


Figure 28: Spatial distribution of SFR water usage per household in 2008

We start with OLS regression. The basic assumptions of OLS models (such as multicollinearity, homoscedasticity and spatial independence of errors) are tested in GeoDa. A spatial autocorrelation is found in the residuals of the OLS model (the Moran's *I* score is 0.185, $p \leq 0.0001$). Therefore, we tested for the spatial dependence using the queen-contiguity-based spatial weight matrix. The diagnostics for spatial dependence showed that both simple tests (Lagrange Multiplier (LM)) of the lag and error models are significant, indicating the presence of spatial dependence. The robust measure for the lagged dependent variable is also significant, while the robust LM test for the lagged error is not. Comparing the Akaike Info Criterion (AIC), Schwarz Criterion (SC) (also called Bayesian Information Criterion) and Log Likelihood (LL) values of the spatial lag and spatial error models, the spatial lag model (AIC=2822, SC = 2876, LL=-1397) has a slightly better fit than the spatial error model (AIC=2833, SC = 2883, LL=-1404). Therefore, only the results of the spatial lag model are reported. The OLS and SLM results are summarized in Table 11. The variables of age of housing units in 2000 and its change did not have significant associations with water usage in 2008 and are thus not reported.

All the explanatory variables together explain 67.6% of the variance in the 2008 average water usage in the OLS regression model. Most of the variables in the model have statistically significant effects at the 95% significance level, except for the area of irrigable land in 2000 and the temporal change in housing density. The variables median household income, household size, percent of households with pools and percent of owner-occupied housing units in 2000 have higher standardized coefficients than their temporal changes between 2000 and 2008/2010. This indicates that the historical states of the variables included have more important influences on water consumption in 2008.

Compared to the OLS regression model, the spatial lag model yields a better fit since it has higher log-likelihood and lower AIC and BIC (Table 11). The housing density in 2000 is not significant any more at the 95% confidence level in the spatial lag model. The changes in the magnitudes of the significant variables present no systematic trend. The coefficient of the lagged water consumption in 2008 (0.371) is highly significant ($p < 0.001$), and the likelihood ratio test also shows significant result (16.926, p ~= 0.000). In addition, the Moran's *I* score (0.0046, p = 0.413) indicates no significant spatial autocorrelation in the residuals. These results confirm the effectiveness of the spatial lag. However, the Heteroscedasticity Test remains significant. So the addition of the lagged dependent variable to incorporate spatial autocorrelation in the model improves the model fit, but it does not make the spatial effects go away. The GWR model on average explains about 69.1% of the variance in water use in 2008. The AIC score of the GWR model (2865) is slightly lower than those of the OLS regression and slightly higher than of spatial lag model (2869 and 2822, respectively), indicating that the GWR model is a better fit to the data than the OLS regression, but not as good as the spatial lag model. The Moran's *I* score (0.192, p ~= 0.000) indicates there is still significant spatial autocorrelation in the residuals of the GWR model. It seems that the GWR model does not account for the spatial dependence effects.

Table 11: Results of ordinary least squares model and spatial lag model

Variables	**Coefficient (Standard Error)**		**t/z-value**		**Standardized coefficient**
	OLS	SLM	OLS	SLM	OLS
Median household income in 2000	0.00082*** (0.000006)	0.00065*** (0.000006)	13.827	11.180	0.885
Change in Median household income	0.00054*** (0.000008)	0.00044*** (0.000008)	6.434	5.805	0.330
Household Size in 2000	8.263** (2.547)	9.257*** (2.304)	3.245	4.018	0.122
Change in Household Size	10.682** (4.052)	10.161** (3.652)	2.636	2.783	0.097
Percent of Households with Pools in 2000	2.565*** (0.399)	1.756*** (0.374)	6.437	4.697	0.303
Change in Percent of Households with Pools	2.572** (0.810)	2.121** (0.732)	3.175	2.897	0.110
Area of Irrigable Land in 2000	0.00026 (0.00016)	0.00027 (0.00015)	1.639	1.843	0.086
Change in Area of Irrigable Land	0.00066** (0.00024)	0.00056** (0.00022)	2.767	2.597	0.120
Housing Density in 2000	-0.0163* (0.0064)	-0.0111 (0.0058)	2.536	-1.912	-0.0936
Change in Housing Density	0.00045 (0.028)	-0.0116 (0.026)	0.016	-0.455	0.0006
Percent of Owner Occupied Housing Units in 2000	-0.268*** (0.052)	-0.260*** (0.047)	-2.465	-5.526	-0.254
Change in Percent of Owner Occupied Housing Units	-0.233* (0.094)	-0.226** (0.085)	-5.122	-2.654	-0.0947
Lagged Water Usage in 2008	--	0.371*** (0.051)	--	7.228	--

*: p<0.05; **: p<0.01; ***: p<0.001.

R^2: R^2(OLS) = 0.676; R^2(SLM) = 0.727 (not directly comparable)
Log-likelihood (LL): LL(OLS) = -1422; LL(SLM) = -1397
Akaike information criterion (AIC): AIC(OLS) = 2869; AIC(SLM) = 2822
Schwarz criterion (SC): SC(OLS) = 2919; SC(SLM) = 2876

Figure 29 shows the spatial distribution of the local R^2 for each block group, with values ranging from 0.672 to 0.703. Overall, the local models do not explain much better than the OLS regression model the variance in water consumption in 2008, but the local R^2 shows a clear trend of southern low values and northern high values.

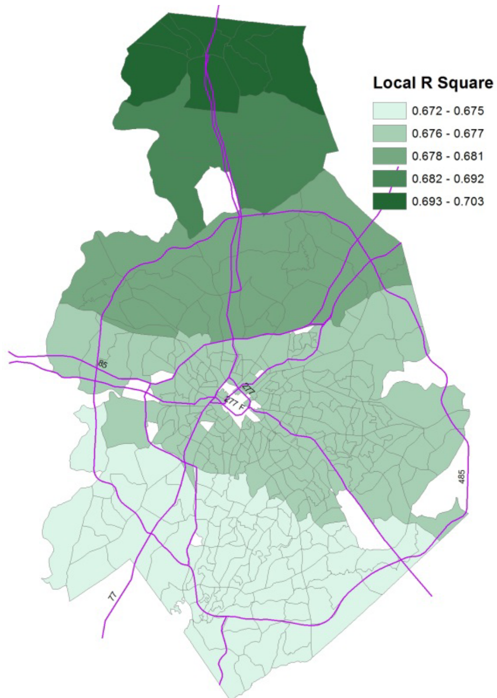


Figure 29: Spatial distribution of the local R-squared values from the GWR model

The essence of GWR model is to construct local relationship of dependent variable and independent variables for each spatial feature to understand the individual underlying process under a local context. Thus regression coefficients for each geographic unit (block group in this case) are calculated, and examining the spatial variations in the local coefficients will help reveal spatial non-stationarity in the relationship (e.g. SFR water consumption in 2008 and the household and housing variables). The spatial patterns of all the estimated local coefficients (no matter significant or not) for all the explanatory variables present gradual changes (decreasing or increasing coefficients from south to north or from east to west).

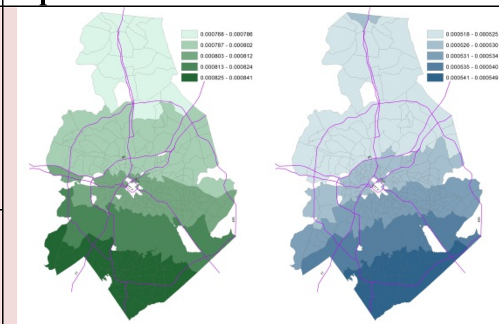
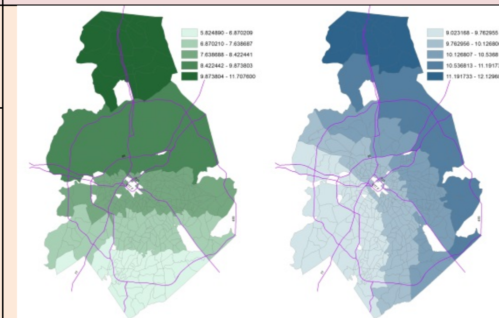
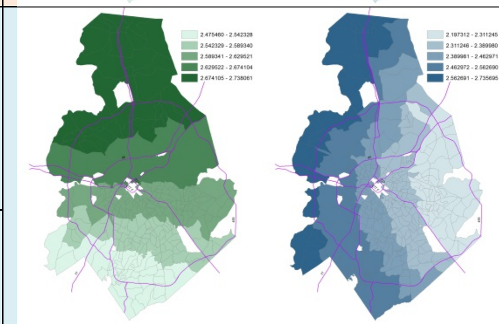
In terms of statistical significance, most of the variables have significant local coefficients in all or part of the block groups, except the temporal change in housing density. For the variable percent of owner occupied housing units, less than 10% of the block groups have insignificant coefficients. Only a quarter of the local coefficients of the area of irrigable land in 2000 are significant, while less than 5% of the block groups have insignificant coefficients for the temporal change in this variable. Next we describe the spatial variations in the significant local coefficients for all the variables (Table 12).

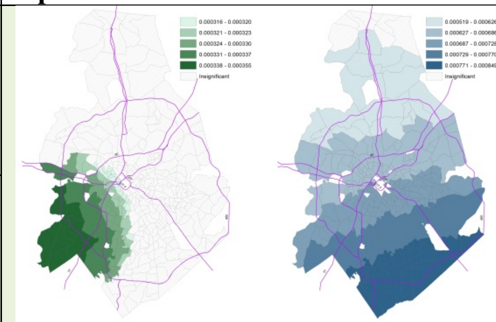
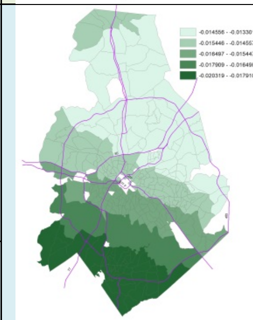
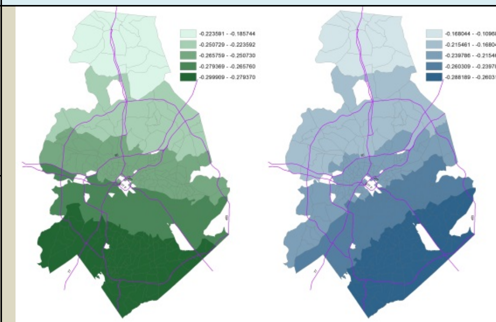
5.5 Discussion

5.5.1 Historical Contingency

The results of OLS regression and spatial lag models suggest that socioeconomic and housing variables in 2000, except for the area of irrigable land, have more important impacts than their temporal change during 2000-2008 in predicting SFR water usage per household in 2008. Although Charlotte and Phoenix are located within different climate zones, SFR water consumption in Charlotte belongs to the same Type I system (defined by Ouyang (2013)) as the one in Phoenix. This implies the persistent lifestyles of households largely determine water consumption assuming no behavioral change and water conservation policy take effects (these factors are not included in the analyses). Moreover, the time span of this study is relatively short (less than a decade), and the built environments and the household structure and characteristics changed relatively slowly. Thus to observe a dramatic response of life-style-determined water consumption to temporal changes in housing and household factors given a relatively conservative political environment in Charlotte is not easy or expected.

Table 12: Spatial patterns of local effects of explanatory variables on SFR water consumption per household in 2008

Variables	Spatial pattern of local coefficients from the GWR model	Spatial distribution of local coefficients
Median household income in 2000	It has a significant effect across all the block groups. The local coefficients range from 0.00077 to 0.00084. The gradient pattern (the map on the left) indicates that the 2008 water uses of the southern neighborhoods have highest/higher associations with the historical value of household income (in 2000), and the ones of the northern neighborhoods are least/less influenced by historical income levels.	
Change in Median household income	Similar to the historical household income, all the block groups have significant coefficients. The range is relatively smaller (between 0.00052 and 0.00055). The map (on the right) of the local coefficients showed a similar gradient pattern with a slightly different orientation.	
Household Size in 2000	All the block groups have significant local coefficients. The magnitude of its association with 2008 water usage varies from 5.825 to 11.708. The effects decrease from the north to the south gradually (the map on the left).	
Change in Household Size	It has a significant effect across all the block groups. The local coefficients range from 9.628 to 12.13. The curvy and gradient pattern (the map on the right) indicates that the 2008 water uses of the northern and eastern neighborhoods have highest/higher associations with the temporal change in household size (in 2000), and the ones of the western neighborhoods are least/less influenced.	
Percent of Households with Pools in 2000	Significant effects are observer in space, although the range is narrow (2.475~2.738). Their spatial pattern shows the gradual changes of values increasing from the south to the north. The spatial extent of the highest level of the associations expands from the north to the northern part of the outer ring (I-485), covering all the municipalities within the north of the county.	
Change in Percent of Households with Pools	All the block groups have significant local coefficients. The values range from 2.197 to 2.736. The map (on the right) of the local coefficients shows radiating gradients – higher values in the north, west and southwest, and lower values in the east.	

Variables	Spatial pattern of local coefficients from the GWR model	Spatial distribution of local coefficients
Area of Irrigable Land in 2000	It does not have a significant effect on water use in 2008 at large. Only the neighborhoods in the southwestern area have significant associations between historical irrigable lot size and water consumption. The spatial pattern still shows radiating gradients with the concentration of higher effects at the furthest southwestern corner.	
Change in Area of Irrigable Land	All of the block groups have significant local coefficients, except a few neighborhoods at the edge of the northern county. The values range from 0.000519 to 0.000849. The map (on the right) of the coefficients showed a gradient pattern increasing from the south to the north.	
Housing Density in 2000	All of the block groups have significant negative associations between the 2008 water use and the housing density in 2000. The local coefficients range from -0.0179 to -0.0146. The gradient pattern in the map indicates that the 2008 water uses of the southern neighborhoods have highest/higher effects (smaller negative values), and the ones of the northern and northeastern neighborhoods are least/less influenced by historical housing density levels.	
Change in Housing Density	It does not show significant local coefficients in any of the block groups.	
Percent of Owner Occupied Housing Units in 2000	It has a significant negative effect across all the block groups. The values range from -0.279 to -0.224. The gradient pattern shown in the map (on the left) has higher values in the southern neighborhoods and lower values in the north of the county.	
Change in Percent of Owner Occupied Housing Units	Similar to the historical percent of owner occupied housing units, all the block groups have significant coefficients. The range is relatively smaller (between -0.26 and -0.168). The map (on the right) of the local coefficients showed a similar gradient pattern with a slightly different orientation.	

After incorporating spatial autocorrelation, the spatial lag model shows that the temporal changes in household size and pool ownership have slightly larger effects than their historical status. The historical contingency of water use on the area of irrigable land is not observed; instead, the increase in the size of irrigable areas exerts stronger and significant effect on water usage in 2008.

Among the housing and socioeconomic factors, the historical state of and temporal change in median household income have the largest positive associations with water usage. The effects of pool ownership rank second. The effects of temporal change in household size and area of irrigable land indicate that the decreasing household size and the development of single-family homes with smaller yards would decrease water usage. Similar to the Phoenix case, neither of the factors of historical housing density or its temporal change is statistically significant in the spatial lag model. Be aware that the density is derived based on the total area of block groups, not the areas of land uses that were developed with or zoned for SFR houses.

5.5.2 Spatial Connectivity and Spatial Heterogeneity

We have the following observations in terms of spatial connectivity in SFR water usage and its relationship with associated factors. First, SFR water consumption shows a significant spatial dependence between neighboring block groups in Charlotte. Higher water usage neighborhoods cluster in the south of the county, while some blocks groups with lower water consumption form significant clusters in the eastern areas close to urban core. Second, for most of the household and housing factors, the local coefficients in the GWR model are significant, but their magnitudes vary in space and demonstrate gradients patterns. Third, beyond the effects of those factors and spatial autocorrelation of the dependent variable we have considered in the spatial lag model, the significant Heteroscedasticity Test suggests that SFR water consumption in 2008 is also affected by some other omitted factors locally and between neighboring block groups.

Spatial heterogeneity is also evident in terms of the relationship between SFR water usage and the explanatory factors. The spatial distribution of the local R^2 from the GWR model indicates that the household and housing factors better explain the variability of SFR water usage in the northern neighborhoods than in the southern neighborhoods. All the historical and temporal variables exhibit significant spatial variations if all the block groups with significant or insignificant local coefficients are included. The strong clustering patterns of statistically significant local coefficients also provide evidence of the spatial heterogeneity in the relationship between SFR water use and most of the factors except area of irrigable land in 2000 and the temporal change in housing density.

5.6 Conclusions

We applied a framework originally proposed for socio-ecological systems to study the complexity in SFR water use in Charlotte along the three dimensions - spatial heterogeneity, spatial connectivity, and historical contingency. The spatial pattern of the "waterscape", no matter it is for a single year (annual water use in 2008 in this chapter) or for multiple-year average (average annual water use as in chapter 4), reveals evident spatial heterogeneity and significant spatial dependence. This is consistent with the existing literature with a focus on the spatial dimension of water usage. To explore the role of historical contingency in modeling residential water usage, we separately estimate the effects of a factor at a historical time point and its temporal change from the past to the present on SFR water usage per household in Charlotte. Similar to the findings from the case study of Phoenix, most of the historical household and housing variables (in 2000) have larger significant influences than their temporal change during 2000-2008 when an OLS regression model is estimated. As Ouyang indicated (2013), although average water usage generally declined in the past decade, the behaviors that tend to use more water (more likely for outdoor activity) may not experience a radical change probably due to a long-established lifestyle. For those factors, it would be more useful to promote behavioral change than the reconfiguration of physical or social structure (the latter one is less controllable or changeable through governance).

The effect of the area of irrigable land in 2000 is not significant. In contrast, the temporal change in area of irrigable land between 2008 and 2000 is significantly positively associated with the SFR water usage in 2008. The possible implication is that the new houses with larger or smaller than historically average irrigable lot size in a neighborhood may induce a larger increase or decrease in water use, assuming the household preference to watering lawn or other vegetation does not change over time. The temporal change in housing density has a positive sign, though it is not significant. Considering the assumption that increased housing density would lead to less water usage, the positive sign may reflect the nonlinearity in the relationship between housing density and SFR water consumption, meaning that when housing density become higher and higher, its effects on inducing water usage reduction may become smaller or disappear.

The incorporation of the spatial dependence in the dependent variable to the model improves the model fit and manifests a decrease in the coefficient estimates in majority of the explanatory factors.

The historical contingency in terms of housing density became insignificant. The results from the GWR model show obvious spatial heterogeneity and connections in the relationship between SFR water usage per household and the factors being examined. The spatial variabilities exhibited in the local associations in terms of the historical variables and the change variables differ for most factors except median household income. The possible policy implications from the GWR analyses are, for the factors with which history contingency dominates over temporal change, it is better to establish different policies or programs for the neighborhoods in the south vs the north of the Charlotte region, while for the factors whose temporal change exerted better effects, it would be useful to differentiate the intervention strategies for the east vs. the west of the Charlotte region.

There are a number of limitations preventing us drawing definite conclusions from this study. First, the spatial lag model did not eliminate the spatial autocorrelation in the residuals, indicating the possibility of model misspecification. To some extent, the spatial heterogeneity observed in the GWR results itself also implies this problem, and GWR models can be vulnerable to misspecification issue. Although we have tested all the possible determinants we have data for, we did not resolve the issue. We may try to add spatial lag variable to GWR model, or include the lag of the explanatory factors in OLS and spatial statistical models. Second, the idea of decomposing a variable into historical and temporal change component would introduce multicollinearity. Although multicollinearity issue is not manifested in the OLS and SLM models we estimated, the decomposition may potentially increase the likelihood of inducing local collinearity in GWR applications (Wheeler and Tiefelsdorf 2005). The imperfection existing in the dataset (due to temporal mismatch of parcel/building and census data, spatial mismatch of census boundary, and the incomplete or un-updated information about lawn and pool) would add concerns to the coefficient estimates. Thus, this study is more exploratory in nature, and we are more cautious with offering policy implications, although the general conclusions such as the presence of the complexity in SFR water use in terms of spatial heterogeneity, spatial connectivity, and historical contingency should hold. The study serving the first but important step will motivate us to make greatest endeavor to improve the model, including identifying more valuable factors, seeking or generating data for them, resolving spatial/temporal mismatch issues, getting better estimates for the current variables, exploring seasonal water use instead of yearly with different sets of factors, experimenting water use data from a different year, etc.

CHAPTER 6: A SPATIAL ECONOMETRIC ANALYSIS OF URBAN WATER DEMAND: EVIDENCE FROM CHARLOTTE, NORTH CAROLINA

6.1 Introduction

Research on urban water usage has mainly focused on identifying determinants and quantifying their effects in order to understand the processes underlying the phenomenon (of water usage). For this purpose, panel data is preferred (if available) because of their advantages over cross-sectional and time-series data. In a panel dataset, there are multiple subjects with repeated observations over multiple time periods. Using panel data, we can incorporate both temporal- and subject-based variability into the model specification and control for the effects of omitted or unobserved variables, and thus obtain more efficient and consistent parameter estimates (Arbués et al. 2003; Polebitski and Palmer 2010; Ouyang 2013). A number of empirical studies exist that use panel data to estimate and predict water usage (for a list of literature, refer to Arbués et al. 2003, Worthington and Hoffman 2008, Polebitski and Palmer 2010, Ouyang 2013). A majority of them employed the household or city/town as the unit of subject, and the temporal scale of observations in the panel data varies from daily, monthly to seasonal and yearly. Few researches have applied panel data models to investigate water usage at the multihouse, or neighborhood (typically census tract or block group) level, which could be appropriate spatial scales for planning and policy-making purposes (Polebitski and Palmer 2010).

In contrast, there is no lack of water demand models based on cross-sectional data at the census tract or block group level. Some of them have paid attention to the spatial pattern of residential water consumption. The research groups who have been continuously focusing on residential water consumption in Phoenix, Arizona (Guhathakurta and Gober 2007; Wentz and Gober 2007; Balling et al. 2008; Lee and Wentz 2008; Ouyang et al. 2014) and Oregon, Portland (Chang et al. 2010; House-Peters et al. 2010; Breyer et al. 2012) have reported clustering patterns of low and high water users at the census tract or block group level. To address the impacts of spatial autocorrelation or dependence on model estimation, various spatially explicit methods have been adopted for better modeling practices. This includes Spatial Lag Model (SLM) (House-Peters and Chang 2011), Spatial Error Model (SEM) (Chang et al. 2010), and Bayesian Maximum Entropy (BME) mapping method (Lee and Wentz 2008).

The presence of spatial autocorrelation and spatial dependence in water consumption needs to be accounted for when using neighborhood-level panel data since it may invalidate the independence assumption of panel data models. Failure to do so may produce inconsistent and biased parameter estimates (LeSage and Pace 2009). The recent development of spatial panel data models (Anselin et al. 2008; Elhorst 2010) offers a great solution to this issue.

Scant literature exists on panel data modeling to explain the associations between water consumption and its determinants, accounting for spatial effects in panel data model, except for a recent study by Ouyang and his collaborators (2013). In their empirical study of Phoenix, several traditional (also referred to non-spatial) and spatial panel data models were applied to examine which model is more appropriate in the context of residential water usage research using tract-level panel data.

Our previous analysis (in chapter 4) on the spatial pattern of SFR water consumption in Charlotte has provided evidence on spatial heterogeneity and spatial dependence and we have monthly and yearly panel data to model the relationship of SFR water usage and various factors. Therefore, we are interested in quantifying the effects of determinants of water consumption by means of spatial panel models. The benefits of using spatial panel models are that we not only could get more reliable estimates, but also understand how spatial effects contribute to modeling SFR water usage in the fast-growing southeastern city – Charlotte – and furthermore gain some insights on demand-management policy from such understanding. Additionally, this will add a new empirical application of spatial panel models in the context of water research.

In this research, we propose two objectives: (1) Examine the empirical needs for incorporating spatial dependence and heterogeneity into panel data models of water consumption; (2) Identify the observable factors and quantify their associations with SFR water consumption in Charlotte by the integration of panel data and the spatial econometric modeling framework.

At variance with the case study in Phoenix, we look at a finer geographical scale (block group) and water consumption in the recent years. The price and water usage restriction variables are included in our empirical models.

Based on the findings in the literature (summarized in chapter 2) and preliminary analyses (the associations between water usage and weather factors in chapter 4 and the analysis about the historical contingency in chapter 5), we have the following hypotheses about the effects of determinants: (1) weather factors would make a great contribution to local water usage on a monthly basis in Charlotte, (2) some household and housing characteristics such as household size, income, and lot size better explain monthly water consumption than the others (for example age factors and housing density), (3) price effect would be inelastic, and become smaller with the intervention of water usage restrictions which supposedly show negative effects.

6.2 Data and Variables

The observations for the dependent variable are monthly SFR water usage per household at the block group level. We choose this temporal scale for two reasons. First, the effects of weather factors on monthly water usage are expected to be more significant than on annual water use. Second, we attempt to evaluate the effects of the price and water usage restrictions variables during the 2007-2008 droughts. The month-based temporal scale is more appropriate than the temporal scale based on years for measuring water usage restrictions since they were enacted during the time period between October 2007 and September 2008. Each block group has 36 observations (for the months from 2007 to 2009), which were derived from the water billing records. Although we tested all of the variables described in the data collection and processing section of chapter 2, only a selection of them are included in the final models due to multicollinearity concerns. The variables in the final list (Table 13) represent the factors from different categories including price, non-price policy, weather, sociodemographic characteristics of household, and housing factors.

Table 13: Descriptive statistics of the dependent and independent variables

Variable	Unit	Mean	Std. dev.	Min	Max
Logged Monthly Water Usage Per Household	CCF	2.062	0.379	1.059	4.102
Logged Average Price	1999 $	1.465	0.116	0.710	1.680
Monthly Cumulative Precipitation	tenth of millimeter	839.86	484.52	105	2385
Monthly Average Maximum Temperature	tenths of degrees Celsius	221.68	76.54	97.74	358.29
Household Size	person	2.47	0.388	1.27	3.5
Logged Median Household Income	1999 $	10.744	0.512	8.966	12.217
Percentage of houses built after 1992	%	0.259	0.302	0	1
Area Of Lot	square feet	16656	7478.722	5312.552	63764.42
Percentage Of Irrigable Land	%	79.164	2.64	67.554	85.935
Housing Units Density	housing units per square kilometers	273.82	148.914	17.042	738.52
Water Usage Restrictions Dummy	NA	0.361	0.480	0	1
Logged Average Price * Water Usage Restrictions Dummy	1999 $	0.518	0.691	0	1.622

6.3 Spatial Effects and Underlying Processes

Before stepping into the specification and estimation details of spatial panel data approach, we will first elucidate the connection between spatial patterns and spatial effects.

Spatial patterns could be explained by spatial effects, which may take two general forms, spatial dependence and spatial heterogeneity (Anselin 1988). Spatial dependence is regarded as the most commonly conceptualized form of spatial effects; in the context of water consumption, it implies that water usage in one location is affected by or functionally related to the water usage in neighboring locations (Anselin and Griffith 1988). The inclusion of spatial dependence in applied econometric models is typically driven by theoretical economic models of interacting agents and social interaction, which are labelled differently in various subfields such as social norms, neighborhood effects, peer group effects, social capital, etc. (Anselin 2002). Subjective norm and perceived behavior controls as direct/indirect factors of water consumption behaviors reviewed in chapter 2 are in line with these theoretical grounds. For example, the neighborhoods with a consensus on keeping lush lawns to maintain their property values may consume more water (Askew and McGuirk 2004; Domene et al. 2005); even without consensus, households may mimic their neighbors' behaviors they perceived regarding when and how (much) to water lawn. Similarly, households may exhibit similar responses to water usage restrictions due to neighborhood effects or social interaction (Aitken et al. 1991; Ramachandran and Johnston 2011; Breyer 2014).

Moreover, the spatial pattern emerging from interactions between adjacent neighborhoods can be attributed to the spillover effect of factors in the neighboring environment. Take the outdoor water usage as an example again. The rate of evapotranspiration or temperature in a neighborhood can be lowered by the existence of a park or trees in the surrounding areas, resulting in less demand of water usage outdoor. Households who live close to a public swimming pool or water body may not build their own pools or frequently fill the pools.

Linking the spatial dependence to spatial econometric models, the spatial lag operators of various types are introduced. A spatially lagged dependent variable (lagged Y) such as water consumption can capture the effect of spatial interaction of neighboring units (e.g. household). The spillover effects of local factors could be expressed by their lags (lagged Xs). If model misspecification (due to missing variables) is assumed related to spatial dependence, we can utilize a spatial autoregressive process for the error term to address such implicit spatial dependence. No matter how spatial interactions/spillovers are formulated, a structure for spatial correlation is needed to define the interacting process; in empirical practice we construct a spatial weight matrix to quantify the range and strength of the relations between the neighboring units (locations).

As another form of spatial effect, spatial heterogeneity could explain spatial patterns in water consumption. Spatial heterogeneity reflects structural instability. In other words, the stability of relationship varies from one location to another (Fornango 2010). The capability and context factors (e.g. weather, household income and size, housing density, to name a few) can be viewed as the potential source of spatial heterogeneity of water consumption as well as the behavioral factors, since they collectively contribute to the spatial process or structure in an observational unit. According to Anselin (Anselin 2003), structural instability can be expressed in the form of variable regression coefficients or in the form of non-constant error variances in a regression model (heteroscedasticity). The incorporation of fixed or random effects into the model will help generate variable coefficients, thus a heterogeneous structure, while spatial heteroscedasticity following from missing spatially-correlated variables or other forms of misspecification related to the variation to location, area, or differential spatial structure (Anselin 1988) could make use of spatially lagged error term.

The incorporation of spatial effects into a model can be complicated by several situations. First, the two forms of spatial effects are not mutually exclusive, thus the joint presence of heterogeneity and spatial dependence necessitates specialized tests and estimation techniques (Anselin 1990). Second, spatial dependence is likely to be present in the error terms when aggregated cross-sectional data is being used and the scale and location of the data does not match the one of the process under study (Anselin 1990, 2002).

Since both spatial dependence and spatial heterogeneity may violate the assumptions of OLS regression, it is important to account for them in order to get accurate estimated models. This applies to panel data model as well.

6.4 Methods - Spatial Panel Data Models

In this research, we first conduct a spatial clustering analysis to assess the spatial patterns of the panel data. The spatial clustering analysis approach has been described in the methods section of chapter 4. Therefore, we mainly focus on the introduction of spatial panel data models here.

Various model specifications for spatial processes have been proposed in the spatial econometrics literature (e.g. (LeSage and Pace 2009; Elhorst 2010). Similar to spatial statistical models, we can construct the spatial lag and error versions of panel data model (Anselin et al. 2008), and they are referred to as Spatial Autoregressive Model (SAR) and Spatial Error Model (SEM). A spatial autoregressive model with auto regressive disturbances model (SAC) is a combination of spatial lag and error models. When a spatial lag of an explanatory variable matrix is included into SAR, the model is called Spatial Durbin Model (SDM) (LeSage and Pace 2009). Consider the following general specification for static spatial panel data models:

$$\begin{cases} Y_{it} = \alpha + \rho W y Y_{jt} + \beta X_{it} + \theta D Z_{jt} + u_i + \gamma_t + v_{it} \\ v_{it} = \lambda E v_{jt} + \varepsilon_{it} \end{cases} \qquad \text{(Eq. 5.1)}$$

where Y_{it} is the dependent variable for observational unit i at time t (i = 1, …, N; t = 1, …, T);

X_{it} is a 1*k row vector of independent variables for observational unit i at time t *(k is the number of the independent variables)*;

Z_{jt} is a 1*m row vector of spatially lagged independent variables for observational unit i at time *t (m is the number of the spatially lagged independent variables, j is the index of neighboring units); sometimes* Z_{jt} and X_{it} can refer to the same set of independent variables*;*

v_{it} is the spatially correlated error terms for all i and t;

ε_{it} is the disturbance term that is independently and identically distributed for all i and t *(*$\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$*)*;

W is the $N \times N$spatial weight matrix for the autoregressive component (or spatially lagged dependent variable), and for each observational unit i, $W = \sum_{j=1}^{N} w_{ij}$;

D is the $N \times N$ spatial weight matrix for the spatially lagged independent variables, for each observational unit i, $D = \sum_{j=1}^{N} d_{ij}$;

E is the $N \times N$spatial weight matrix for the idiosyncratic error component, for each observational unit i, $E = \sum_{j=1}^{N} e_{ij}$;

u_i is the individual fixed or random effect, for random effect $u_i \sim N(0, \sigma_\mu^2)$;

γ_t is the time effect (fixed or random);

α is the intercept;

β is the vector of coefficients;

ρ is the spatial autoregressive coefficient;

θ is the coefficients for the spatially lagged independent variables;

λ is the spatial error autocorrelation coefficient.

When ρ, θ, and λ are zero, the model specification becomes the specification for traditional (or non-spatial) panel data models with both individual and time effects (Table 14). If the optional terms (u_i and γ_t) are removed, we get a (pooled) OLS model. If either u_i or γ_t terms are specified in a traditional model, they correspond to individual effect model or time model. The effect could be fixed or random depending on the purpose and the assumption on the underlying processes (or the relations between the dependent variable and the observed and unobservable independent variables). In this study, we estimated all the panel data models listed in Table 14 except the one with only individual effects.

Table 14: Various types of panel data models

Panel data models	Parameters	Specification
Pooled OLS model	$\rho = \theta = \lambda = u_i = v_t = 0$	$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it}$
Panel data model with time effects	$\rho = \theta = \lambda = u_i = 0$	$Y_{it} = \alpha + \beta X_{it} + \gamma_t + \varepsilon_{it}$
Panel data model with individual effects	$\rho = \theta = \lambda = v_t = 0$	$Y_{it} = \alpha + \beta X_{it} + u_i + \varepsilon_{it}$
Panel data model with both individual and time effects	$\rho = \theta = \lambda = 0, u_i = v_t \neq 0$	$Y_{it} = \alpha + \beta X_{it} + u_i + \gamma_t + \varepsilon_{it}$

If the parameters ρ, θ, or λ in the generalized specification are not zero, different combinations of these three parameters yield a variety of static spatial panel data models as listed in Table 15.

Table 15: Various types of spatial panel data models

Spatial panel data models	Parameters	Specification
Spatial Autoregressive Model (SAR)	$\theta = \lambda = 0$	$Y_{it} = \alpha + \rho W Y_{jt} + \beta X_{it} + u_i + \gamma_t + \varepsilon_{it}$
Spatial Error Model (SEM)	$\theta = \rho = 0$	$Y_{it} = \alpha + \beta X_{it} + u_i + \gamma_t + v_{it}$ $v_{it} = \lambda E v_{jt} + \varepsilon_{it}$
Spatial Durbin Model (SDM)	$\lambda = 0$	$Y_{it} = \alpha + \rho W Y_{jt} + \beta X_{it} + \theta D Z_{jt} + u_i + \gamma_t + \varepsilon_{it}$
Spatial Autoregressive Model with Autoregressive Disturbances Model (SAC)	$\theta = 0$	$Y_{it} = \alpha + \rho W Y_{jt} + \beta X_{it} + u_i + \gamma_t + v_{it}$ $v_{it} = \lambda E v_{jt} + \varepsilon_{it}$

Note: The spatial Durbin model is simplified to the spatial lag model if $\theta = 0$, and to the spatial error model if $\theta + \lambda\beta = 0$.

Linking the mathematical specification back to the underlying processes discussed above, spatial weight matrices (W, D and E) define the structure of spatial correlation among observational units. WY_{jt} denotes the spatial effect that the dependent variable of neighboring units have on Y_{it}; DZ_{jt} captures the spillover effects of the explanatory variables of neighboring units; similarly, Ev_{jt} represents the spatial correlations of errors among observational units, which may result from unobservable factors shared by units. u_i represents one way of constructing the variable regression coefficients of the structural relationship for each unit, and γ_t denotes time heterogeneity. Both individual effects and time effects are also the important components in traditional panel model.

In this study, we estimate three specifications of spatial panel data models: spatial autoregressive model (or spatial lag model), spatial error model and spatial Durbin model. The Maximum likelihood (ML) method is commonly used to estimate the parameters of spatial panel data models specified above, although there are other estimation methods such as instrumental variables and generalized method of moments. The software package (xsmle, a user-written STATA program (Belotti et al. 2013)) which takes the ML approach is used here.

In all these models, we include the fixed or random effect for the spatial unit (u_i) and the fixed time effect (γ_t). The random effect estimator is appropriate only if u_i are uncorrelated with the included explanatory variables. If this assumption is violated, the model will produce biased and inconsistent parameter estimates (Elhorst 2010). In contrast, the fixed effect estimator allows for the correlation between time-invariant heterogeneity at the spatial unit level (block group) and included predictors. The disadvantage of the fixed effect in a panel data model specification is that no time-invariant variables can be included in the specification and there is concern that fixed effects may bias the estimates if variables are changing slowly over time (Plümper and Troeger 2004). A Hausman test can assist in deciding between fixed or random effects, which basically tests whether the unique errors (u_i) are correlated with the regressors. However, in empirical practices, the choice may be built upon the needs for evaluating time-invariant variables and the purpose of model generalization. We restrict the time effect (γ_t) to be fixed since all the months during the three year time period are included in the model.

Different tests can be applied to the traditional panel data model to decide whether fixed or random effects should be included in the model. An F-test or likelihood ratio (LR) test can be employed to investigate the null hypothesis that the fixed effects (for either spatial units or time) are jointly insignificant. A Lagrange multiplier (LM) test developed by Breusch and Pagan (Breusch and Pagan 1980) is useful for examining the null hypothesis of no random effects.

To determine the most appropriate spatial panel data model, we first test whether the spatial version of panel data model (SAR and SEM) is better than traditional (non-spatial) panel data model using LM tests and robust LM tests. Elhorst (2010) generalized for spatial panel data the LM test and robust LM test that were proposed for cross-sectional spatial data by Burridge (1980) and Anselin et al. (1996) respectively. This is based on a panel data model without spatial effects (when $\rho = \theta = \lambda = 0$). If the LM and robust LM tests support either spatial a lag model or a spatial error model, we estimate the spatial Durbin model, and test whether it can be simplified to the spatial lag or spatial error models using the Wald tests.

For the spatial lag model or spatial Durbin model, which includes a spatial lag of the dependent/independent variable(s), caution is needed when interpreting their parameters in these models (LeSage and Pace 2009). The looped feedback effects among neighboring geographical units need to be accounted for. To explain the effect of an explanatory variable on the dependent variable in a correct way, one should look at the average direct, indirect and total effects. In the context of water demand, the direct effect measures the average change in average water usage in the spatial unit i (y_i) caused by a unit change in this unit's explanatory variable (e.g. area of lot). It includes the feedback effect that passing through neighboring spatial units and back to the original spatial unit (Ouyang 2013). The indirect effect (e.g. spatial spillover effect) refers to the average change in average water usage in the spatial unit i (y_i) caused by a unit change in the explanatory variable of its neighboring spatial units. The total effect is a summary measure of both direct and indirect effects associated with a unit change in an independent variable. The three types of effects were proposed by LeSage and Pace (2009) for a cross-section panel data model, and Elhorst (2012a) extended the idea to spatial panel models. All of these effects are reported in the results section. However, when discussing our results, we mainly focus on the total effects.

Spatial weight matrices for the dependent variable, independent variables and idiosyncratic error (W, D and E) can be the same, or constructed differently, depending on the spatial processes that are hypothesized or conceptualized. In this study, we employ two types of spatial weight matrix. One is to define neighbors of a spatial unit based on first-order queen contiguity of spatial relationship. This conceptual spatial arrangement assumes, for a block group, the water consumption of its immediate neighbors (sharing a common border and corner with it) affects or is functionally related to its own water consumption. Another way to create a spatial weight matrix is based on the k-nearest neighbor criterion. We assume that the six nearest neighbors of a block group contribute to its water usage. For our spatial panel data models, the same spatial weight matrix is employed because we assume that no matter its water consumption, its determinants or error, the spatial effects at the aggregated geographical level take place in a similar way.

Although we know this assumption may be too simple, this serves our purpose as an initial step of evaluating spatial effects in a panel data setting.

6.5 Results and Discussion

Before conducting multivariate statistical modeling, we performed a test of spatial autocorrelation using Moran's *I* statistic to establish the extent of spatial effects in our monthly water use dataset. A positive test would provide initial justification for incorporating spatial effects into spatial panel data models.

For all of the 36 monthly data series, the Moran's *I* statistics have a value larger than 0.4, which is statistically significant at the 0.01 level (Table 16). This means that high water users cluster together, and there is less than 1% likelihood that the observed clustered pattern of residential water consumption in Charlotte area could have occurred by chance. This spatial dependence persists over time.

Table 16: Global Moran's *I* statistics of monthly average SFR water usage per household at block group level from January in 2007 to December 2008

month	2007		2008		2009	
	Moran's *I*	**Z score**	**Moran's *I***	**Z score**	**Moran's *I***	**Z score**
Jan	0.574	18.071	0.497	15.658	0.551	17.345
Feb	0.546	17.191	0.552	17.394	0.576	18.149
Mar	0.630	19.909	0.558	17.607	0.573	18.032
Apr	0.596	18.874	0.684	21.606	0.596	18.883
May	0.618	19.537	0.533	16.816	0.573	18.188
Jun	0.608	19.290	0.639	20.249	0.632	20.000
Jul	0.526	16.633	0.467	14.778	0.563	17.859
Aug	0.597	18.895	0.553	17.548	0.590	18.701
Sep	0.607	19.247	0.550	17.441	0.599	19.017
Oct	0.619	19.591	0.575	18.236	0.592	18.783
Nov	0.519	16.424	0.550	17.433	0.603	19.074
Dec	0.543	17.163	0.499	15.805	0.585	18.431

Note: the p-value for all the months is below the 0.0001 level.

Next, we start with the panel data models without spatial interactions (Table 17). The time-fixed effects (actually composed of month and year fixed effects) and the fixed and random effects of spatial units are first tested. The F-test (315.25, df = 13, $p < 0.001$) indicates that the time-fixed effects are jointly statistically significant, and should be included. The LM tests also indicate the need for adding spatial fixed or random effects. The result of the Hausman test rejects the null hypothesis and suggests that the fixed effects model is more appropriate. The results of four non-spatial models are reported as follows.

Only the coefficients for logged average price and monthly average maximum temperature have a consistent and significant sign across all the models. The price variable has a negative effect on the monthly SFR water consumption per household. The elasticity is relatively higher than the ones seen in the literature. One possible reason is that the calculation of the average price in this study is based on the total charges on water bills instead of the charges for the water portion only. The temperature variable has a positive and significant effect, meaning that more water is consumed when the temperature is higher. Monthly cumulative precipitation shows a negative association with water demand in the models with time fixed effects, but not in the OLS model. The dummy variable for the policy of restricting water usage during October 2007 and September 2008 has a positive sign in the coefficient estimate for models 1, 3 and 4, while the interaction term of price and non-price policy variable always shows an opposite sign. The total effect of the water usage restrictions variable is negative while holding price constant. This also indicates the price effect is greater when the water usage restriction policy is practiced.

Table 17: Results of non-spatial models, block group level

Independent Variables	**Model 1: Pooled OLS Regression Model**	**Model 2: Panel data Model with Time-Fixed Effects**	**Model 3: Panel Data Model with Spatial Fixed Effects and Time-Fixed effects**	**Model 4: Panel Data Model with Spatial Random Effects and Time-Fixed effects**
Logged Average Price	-1.18355*** (-0.0196)	-2.43925*** (-0.0338)	-1.88988*** (-0.0260)	-1.90730*** (-0.0242)
Monthly Precipitation	0.00006*** (0.0000)	-0.00002*** (0.0000)	-0.00003*** (0.0000)	-0.00003*** (0.0000)
Monthly Avg Max. Temperature	0.00231*** (0.0000)	0.00148*** (-0.0001)	0.00100*** (0.0000)	0.00101*** (-0.0001)
Household Size	0.02639*** (-0.0051)	0.04006*** (-0.0042)	0.00406 (-0.0049)	0.00336 (-0.0065)
Logged Median HH Income	0.24661*** (-0.0045)	0.23904*** (-0.0038)	0.00063 (-0.0077)	0.03415*** (-0.0079)
Percentage of houses built after 1992	-0.08058*** (-0.0083)	-0.10291*** (-0.0068)	1.06153 (-0.9148)	0.04118 (-0.0344)
Area Of Lot	0.00001*** (0.0000)	0.00001*** (0.0000)	0.00004 (0.0000)	0.00002*** (0.0000)
Percentage Of Irrigable Land	-0.00847*** (-0.0011)	-0.00856*** (-0.0009)	0.05449 (-0.0766)	-0.01327** (-0.0047)
Housing Density	-0.00018*** (0.0000)	-0.00019*** (0.0000)	0 (.)	0.00007 (-0.0001)
Water Usage Restrictions	1.32023*** (-0.0581)	-0.12809 (-0.0657)	0.0426 (-0.0346)	0.03746 (-0.0456)
Logged Price * Restrictions	-0.99437*** (-0.0402)	0.03667 (-0.0455)	-0.07085** (-0.0234)	-0.06763* (-0.0315)
2.month		-0.12718*** (-0.0079)	-0.11596*** (-0.0014)	-0.11633*** (-0.0054)
3.month		-0.08961*** (-0.0128)	-0.04583*** (-0.0032)	-0.04727*** (-0.0089)
4.month		-0.10975*** (-0.0167)	(-0.0068)	-0.02755* (-0.0116)
5.month		0.03438 (-0.0220)	0.12914*** (-0.0104)	0.12606*** (-0.0152)
6.month		0.03766 (-0.0295)	0.21319*** (-0.0125)	0.20755*** (-0.0205)
7.month		0.11005*** (-0.0289)	0.26788*** (-0.0123)	0.26274*** (-0.0201)
8.month		0.12433*** (-0.0298)	0.20269*** (-0.0131)	0.20012*** (-0.0207)
9.month		0.12163*** (-0.0234)	0.16779*** (-0.0119)	0.16626*** (-0.0162)
10.month		0.12060*** (-0.0166)	0.13765*** (-0.0075)	0.13708*** (-0.0115)
11.month		0.10814*** (-0.0104)	0.09625*** (-0.0036)	0.09659*** (-0.0072)

Independent Variables	Model 1: Pooled OLS Regression Model	Model 2: Panel data Model with Time-Fixed Effects	Model 3: Panel Data Model with Spatial Fixed Effects and Time-Fixed effects	Model 4: Panel Data Model with Spatial Random Effects and Time-Fixed effects
12.month		0.17704*** (-0.0084)	0.15133*** (-0.0027)	0.15210*** (-0.0058)
2008.year		0.18524*** (-0.0060)	0.13301*** (-0.0029)	0.13452*** (-0.0042)
2009.year		0.36734*** (-0.0077)	0.22118*** (-0.0046)	0.23073*** (-0.0058)
intercept	1.18010*** (-0.1040)	3.11771*** (-0.0961)	-0.81694 (-5.8540)	4.84199*** (-0.3723)
Log-likelihood	2279.06889	4753.83044	9515.57515	8607.27657
AIC	-4534.13777	-9457.66089	-18985.1503	-17160.55314
BIC	-4444.97828	-9271.91195	-18814.26127	-16959.94428

1. standard errors in parentheses;
2. Significance level: *0.05 **0.01 ***0.001;
3. January is the reference month for the month fixed effects, and 2007 is the reference year for the year fixed effects.

Among the household and housing factors, household size, median household income, and lot area show a consistent and positive association with the dependent variable across all four models. In the OLS model and the time fixed effect panel data model, the variables percentage of houses built after 1992, percentage of irrigable land and housing density have negative and significant signs. However, when the fixed/random effects of a spatial unit are accounted for, the signs of these three variables show an opposite direction.

For the fixed effects model (Model 3), all the factors related to household and housing characteristics fail to be significant. This implies there are unobserved factors having more influence than the ones being tested, and their effects are captured by the fixed effects component of the observational units in the model. Another possible reason is the change in the explanatory variables being estimated is small within the 36-month periods.

For the random effects model (Model 4), median household income, lot area and percentage of irrigable land show significant associations with water consumption. The sign of the variable average percentage of irrigable land at block group level is negative, although we expected the areas with higher percentage of irrigable land would demand more outdoor water use.

In Models 3 and 4, the coefficient estimates for months reveal that water usage became less in February and were highest in the June and July months in summer. The fixed effect for the year 2009 is larger than the one in 2008 and 2007. This is counterintuitive since the overall trend of the SFR water consumption decreased.

Based on the log-likelihood and other goodness-of-fit measures such as AIC and BIC, the panel data model with spatial fixed effects and time-fixed effects (Model 4) performs better than the other three. However, given that Models 1-4 do not consider the effect of spatial dependence, the parameter estimates of non-spatial panel data models may be biased.

Next, we test whether the spatial lag model or the spatial error model is more appropriate than a model without spatial dependence. The results of the LM test (1209.54, $p<0.001$) and of the robust LM test (3.86, $p<0.001$) support the inclusion of spatially autocorrelated error terms. The LM test (1205.68, $p<0.001$) and the robust LM test (3.07, $p<0.001$) for spatial lag model are significant, suggesting that the spatial lag model is a better option than its corresponding non-spatial panel data model. Thus, the preliminary tests suggest that either SEM or SAR with both spatial fixed effects and month/year fixed effects is an appropriate model.

As for the choice of the specification of spatial panel data models, we start with the spatial Durbin models. We do not include the spatial lags of all of the explanatory variables because some of them have the same value for the entire city given a specific time point such as weather factors and non-price policy, which would create an identification problem. Thus they may not exhibit spillover effects over through neighbor effects or spatial interaction. Only the lag variables of the household and housing factors are included in SDM.

For the SDM with fixed effects (Model 6a), the Wald tests show that the null hypothesis that the spatial Durbin model can be simplified to the spatial lag model or spatial error model are not rejected. Thus either SAR with fixed effects (Model 5a) or SEM with fixed effects (Model 7) is a better option than the spatial Durbin model. When random effects are considered instead of fixed effects, the test finds that the SAR model (Model 5b) is the most appropriate model and SDM (Model 6b) is preferred when compared to SEM. For comparison purposes, we report the results from SAR and SDM with either fixed or random effects and from SEM with fixed effects (Table 19 and 20). The results are based on the spatial weight matrix from the first-order queen contiguity specification. Another specification of the spatial weight matrix is also tested and yields similar results (generally getting higher parameter estimates).

The significant and positive coefficient (ρ) of the spatially lagged dependent variable in SAR and SDM models (Model 5 and 6) indicate that the monthly SFR water usage per household in one block group increases (decreases) in response to the increase (decrease) in the monthly SFR water usage per household in neighboring block groups.

This is consistent with our empirical finding from Moran's *I* test and the findings in a cross-sectional setting in the literature reviewed earlier.

Examining the results for SAR models (both fixed and random effects), the price, non-price policy variables, their interaction term, and two weather factors are all significant at the 0.01 level in terms of the estimated coefficients, their direct, indirect and total effects. Compared to the non-spatial version of panel data models (Models 3 and 4), the direction of the effect of the water usage restriction variable changes from positive to negative, indicating that the presence of non-price policy induced a decrease in water consumption. The positive sign of the interaction term of the price and water usage restriction variable implies that the price effect would become less influential when the non-price policy is implemented. This conclusion is consistent with the literature that our third hypothesis is based on (Kenney et al. 2008). The opposite results obtained from the traditional panel data models may reflect the biased specifications of those models.

Regarding the household and housing factors, the estimated coefficients have different sign and magnitude in SAR models, and all of them are non-significant no matter with fixed or random effects, except that the variable lot area is significant and positive in the SAR with random effects (Model 5b). The total effects of all the explanatory variables at least doubled the direct effect (see Table 19) or the coefficient (see Table 18), after the looped feedback effects from spatial dependence are counted in. Comparing the total effects in SAR and the estimated coefficients in the panel data models, the parameters for all the variables are underestimated (in terms of absolute values) in the non-spatial models, although they are overestimated compared to the direct effects of corresponding variables in SAR models. Interestingly, the two SAR models (fixed effects Model 5a and random effects Model 5b) generate similar parameter estimates.

The results of the spatial Durbin models show that the estimated coefficients of the factors (Xs) (Table 18) are similar to their corresponding direct effects (Table 20) in terms of signs, sizes, and statistical significance. However, the parameter estimates of the spatially lagged independent variables (W*Xs) are slightly different from their corresponding indirect effects. None of the independent variables has statistically significant lagged effects at the 0.05 level. Neither are their indirect effects significant, except that area of irrigable land has a significant positive indirect effect in the SDM with random effect (Model 6b). Comparing SAR and SDM, the results are very similar. This is possible since the difference between the two model specifications is the addition of lagged independent variables, and those lagged variables are not significant in our case. The effects of the observational units (block group) probably exerted greater influences on water consumption than the selected factors do.

The spatial error panel data model with fixed effects (Model 7) has similar results in terms of the direction and significance of coefficient estimates as the ones in the SAR model with fixed effects (Model 5a). The price, precipitation, and water usage restriction variables have significant and negative effects on water consumption, while the temperature and interaction term have significant and positive effects. The household and housing variables are not significant. Comparing SEM and SAR with fixed effects, the magnitude of the coefficient estimates in SEM is much higher than their counterpart of SAR, but slightly lower than the total effects derived in SAR. The statistically significant value of λ in the SEM suggests that a spatial process generate the errors.

In summary, all of the models except the pooled OLS model confirm our first hypothesis about the effects of the weather variables. The results from the spatial panel models support the third hypothesis regarding the interaction effects of price and non-price variables. Although the housing and household factors being investigated are not significant in the spatial panel settings, the income, lot size and percentage of irrigable land variables exhibit statistically significant effects compared to the other factors related to age of housing and density.

6.6 Conclusions

This study explores the role of the spatial effects in understanding of the relationship between monthly SFR water usage and its determinants in Charlotte. Using a set of panel data obtained from a variety of sources at block group level, we estimated and compared several non-spatial and spatial panel data models. Diagnostic tests suggested that the non-spatial panel data model with block-group fixed effects and time fixed effects is most appropriate. When incorporating the spatial lags of the dependent variable (water consumption) or of the errors, the spatial autoregressive coefficient or spatial error autocorrelation coefficient is significant, indicating the existence of spatial dependence and/or spatial heterogeneity. After accounting for the possible spatial effects, the price, temperature, precipitation and water usage restriction variables have significant and expected association with monthly SFR water consumption per household. For these variables, a large difference between the direct effects from each of the spatial panel models and the effects found in traditional panel data models is observed. The price effect becomes smaller with the intervention of water usage restrictions policy in a spatial panel setting, while the non-spatial panel data model suggested the opposite conclusion.

Although we estimated the effects of household and housing factors, the results from all of the spatial panel models did not find a significant association between each factor and water consumption. The spatial Durbin model shows little evidence of spillover effects of the socioeconomic and housing variables, since the coefficient estimates of all the lagged explanatory variables are insignificant and the indirect effects are not statistically insignificant as well. This may be due to the slow change in these variables within a relatively short time period (three years), or because there would be no spillover effects from the selected variables. In conclusion, the monthly SFR water consumption at the block group level is better explained by price, non-price policy and weather variations rather than by household and housing factors, when the block-group level structural variations are accounted for by the fixed effects of block groups.

There are a number of limitations in this study. First, the model specification may still need to be improved by including or seeking different sociodemographic and housing variables, and we have not explored the vulnerability of these empirical spatial panel data models to the modifiable areal unit problem by comparing modeling results obtained through analyses at different geographical scales (e.g. census tracts). The sensitivity of the model results to the conceptualization of spatial weight matrices requires further study. Second, the temporal span or scale of the panel dataset may not be very helpful in uncovering the effects of the socioeconomic and housing variables on water consumption. We could either add more monthly data from other years or focus on yearly consumption. Although we discussed the possible mechanism for explaining the spatial dependence we observed, the results from this study did not provide any evidence. Third, we need other data sources to explore the influences of social connectivity, social networking, and shared attitudes on water consumption behavior, based on which to develop a theoretical framework and then test the mechanism explicitly. Once we have better capabilities in interpreting the directions of the direct and spillover effects, we could pursue more complicated modeling frameworks such as dynamic spatial panel data models that include temporally lagged effects (Elhorst 2012b) and would be useful in predicting water consumption.

Table 18: Coefficient estimates of spatial panel models, block group level

Independent Variables	Model 5: Spatial Lag Panel Model		Model 6: Spatial Durbin Model		Model 7: Spatial Error Panel Model
	a. Fixed effect	b. Random effect	a. Fixed effect	b. Random effect	Fixed effect
Logged Average Price	-0.91536*** (-0.0423)	-0.92817*** (-0.0427)	-0.91630*** (-0.0422)	-0.92819*** (-0.0426)	-2.43604*** (-0.0761)
Monthly Precipitation	-0.00001*** (0.0000)	-0.00001*** (0.0000)	-0.00001*** (0.0000)	-0.00001*** (0.0000)	-0.00003*** (0.0000)
Monthly Mean Maximum Temperature	0.00057*** (0.0000)	0.00058*** (0.0000)	0.00058*** (0.0000)	0.00058*** (0.0000)	0.00143*** (-0.0001)
Household Size	-0.00073 (-0.0046)	0.0009 (-0.0047)	-0.00289 (-0.0052)	-0.00029 (-0.0050)	-0.00211 (-0.0051)
Logged Median HH Income	-0.00081 (-0.0078)	0.01219 (-0.0074)	-0.00139 (-0.0080)	0.01111 (-0.0079)	-0.00246 (-0.0069)
Percent of houses built after 1992	1.53351 (-0.9699)	0.02187 (-0.0305)	1.47409 (-1.0196)	0.07921 (-0.0464)	1.15162 (-0.9081)
Area Of Lot	0.00004 (0.0000)	0.00001*** (0.0000)	0.00003 (0.0000)	0.00001*** (0.0000)	0 (0.0000)
Percentage Of Irrigable Land	0.00982 (-0.0739)	-0.0074 (-0.0044)	0.01473 (-0.0753)	-0.00105 (-0.0058)	0.07078 (-0.0602)
Housing Density	0 (.)	0.00003 (0.0000)	0 (.)	0.00001 (-0.0001)	0 (.)
Water Usage Restrictions	-0.11559*** (-0.0302)	-0.11702*** (-0.0303)	-0.11458*** (-0.0302)	-0.11572*** (-0.0303)	-0.41965*** (-0.0769)
Logged Price * Restrictions	0.06132** (-0.0206)	0.06204** (-0.0206)	0.06055** (-0.0206)	0.06114** (-0.0206)	0.24340*** (-0.0515)
2.month	-0.04397*** (-0.0027)	-0.04440*** (-0.0028)	-0.04397*** (-0.0027)	-0.04439*** (-0.0028)	-0.12673*** (-0.0070)
3.month	-0.03721*** (-0.0029)	-0.03796*** (-0.0029)	-0.03730*** (-0.0029)	-0.03800*** (-0.0029)	-0.08330*** (-0.0134)
4.month	-0.05006*** (-0.0051)	-0.05136*** (-0.0051)	-0.05027*** (-0.0051)	-0.05149*** (-0.0050)	-0.09146*** (-0.0230)

Independent Variables	Model 5: Spatial Lag Panel Model		Model 6: Spatial Durbin Model		Model 7: Spatial Error Panel Model
	a. Fixed effect	b. Random effect	a. Fixed effect	b. Random effect	Fixed effect
5.month	-0.00617 (-0.0063)	-0.00719 (-0.0062)	-0.00634 (-0.0063)	-0.00732 (-0.0062)	0.04309 (-0.0354)
6.month	-0.01961* (-0.0095)	-0.02155* (-0.0095)	-0.01990* (-0.0096)	-0.02176* (-0.0095)	0.05172 (-0.0401)
7.month	0.00688 (-0.0096)	0.00528 (-0.0095)	0.00653 (-0.0096)	0.00508 (-0.0095)	0.11967** (-0.0444)
8.month	0.02714** (-0.0084)	0.02657** (-0.0084)	0.02685** (-0.0084)	0.02650** (-0.0084)	0.12954** (-0.0473)
9.month	0.03175*** (-0.0072)	0.03157*** (-0.0072)	0.03170*** (-0.0072)	0.03154*** (-0.0072)	0.11949** (-0.0417)
10.month	0.03748*** (-0.0048)	0.03764*** (-0.0048)	0.03739*** (-0.0048)	0.03761*** (-0.0048)	0.12321*** (-0.0277)
11.month	0.03867*** (-0.0030)	0.03911*** (-0.0031)	0.03871*** (-0.0031)	0.03911*** (-0.0031)	0.10898*** (-0.0157)
12.month	0.06467*** (-0.0035)	0.06544*** (-0.0035)	0.06482*** (-0.0035)	0.06541*** (-0.0035)	0.18261*** (-0.0106)
2008.year	0.07160*** (-0.0034)	0.07298*** (-0.0034)	0.07200*** (-0.0036)	0.07292*** (-0.0034)	0.18686*** (-0.0110)
2009.year	0.12143*** (-0.0060)	0.12660*** (-0.0061)	0.12477*** (-0.0064)	0.12957*** (-0.0061)	0.31408*** (-0.0145)
Spatial autoregressive coefficient (ρ)	0.66997*** -0.0218	0.66730*** -0.0221	0.66979*** -0.0219	0.66754*** -0.022	
Spatial error autocorrelation coefficient (λ)					0.80452*** -0.0192
Spatially lagged Household Size			0.01226 (-0.0108)	0.00904 (-0.0107)	
Spatially lagged Logged Median HH Income			0.00788 (-0.0121)	0.01082 (-0.0117)	
Spatially lagged Percpost1992			-2.19961 (-2.6827)	-0.07999 (-0.0629)	
Spatially lagged Area Of Lot			0.00008 (-0.0001)	0 (0.0000)	
Spatially lagged Percentage Of Irrigable Land			-0.23245 (-0.1684)	-0.01006 (-0.0092)	
Spatially lagged Housing Density			0 (.)	0.00003 (-0.0001)	
Log-likelihood	12901.20093	11971.51647	12904.46011	11975.29489	14359.10395
AIC	-25614.40186	-23743.03294	-25604.92023	-23738.58978	-28668.20789
BIC	-24915.98584	-23000.03717	-24847.06454	-22951.01426	-28482.45895

1. standard errors in parentheses; 2. significance level: *0.05 **0.01 ***0.001;
3. January is the reference month for the month fixed effects, and 2007 is the reference year for the year fixed effects.

Table 19: Direct, indirect and total effects of spatial Durbin model, block group level

Independent Variables	Spatial Lag Panel Model					
	Fixed Effect			Random Effect		
	Direct	Indirect	Total	Direct	Indirect	Total
Logged Average Price	-1.02864***	-1.74898***	-2.77762***	-1.04177***	-1.75393***	-2.79570***
	(-0.0393)	(-0.1131)	(-0.0994)	(-0.0395)	(-0.1157)	(-0.1011)
Monthly Precipitation	-0.00001***	-0.00001***	-0.00002***	-0.00001***	-0.00001***	-0.00002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Monthly Mean Maximum Temperature	0.00065***	0.00110***	0.00175***	0.00066***	0.00111***	0.00177***
	(0.0000)	(-0.0001)	(-0.0001)	(0.0000)	(-0.0001)	(-0.0001)
Household Size	-0.0008	-0.00145	-0.00225	0.00101	0.00162	0.00263
	(-0.0051)	(-0.0089)	(-0.0140)	(-0.0052)	(-0.0089)	(-0.0141)
Logged Median HH Income	-0.00094	-0.00157	-0.00251	0.01375	0.02328	0.03704
	(-0.0084)	(-0.0145)	(-0.0229)	(-0.0080)	(-0.0139)	(-0.0218)
Percentage of houses built after 1992	1.78087	3.04298	4.82385	0.02632	0.04419	0.07051
	(-1.0742)	(-1.8597)	(-2.9242)	(-0.0340)	(-0.0574)	(-0.0913)
Area Of Lot	0.00004	0.00007	0.00011	0.00001***	0.00002***	0.00003***
	(0.0000)	(-0.0001)	(-0.0001)	(0.0000)	(0.0000)	(0.0000)
Percentage Of Irrigable Land	0.01174	0.0193	0.03104	-0.00835	-0.01427	-0.02262
	(-0.0821)	(-0.1409)	(-0.2228)	(-0.0049)	(-0.0087)	(-0.0136)
Housing Density	0	0	0	0.00004	0.00007	0.00011
	(0.0000)	(0.0000)	(0.0000)	(-0.0001)	(-0.0001)	(-0.0001)
Water Usage Restrictions	-0.12996***	-0.22386**	-0.35382***	-0.13129***	-0.22390**	-0.35518***
	(-0.0346)	(-0.0702)	(-0.1039)	(-0.0342)	(-0.0691)	(-0.1023)
Logged Price * Restrictions	0.06884**	0.11897**	0.18781**	0.06949**	0.11890**	0.18840**
	(-0.0236)	(-0.0461)	(-0.0693)	(-0.0234)	(-0.0453)	(-0.0682)

1. standard errors in parentheses; 2. significance level: *0.05 **0.01 ***0.001;

Table 20: Direct, indirect and total effects of spatial Durbin model and spatial error model, block group level

Independent variables	Spatial Durbin Model						Spatial Error Model
	Fixed Effect			Random Effect			Fixed Effect
	Direct	Indirect	Total	Direct	Indirect	Total	Coefficient
Logged Average Price	-1.02960***	-1.74937***	-2.77897***	-1.04185***	-1.75477***	-2.79662***	-2.43604***
	(-0.0392)	(-0.1140)	(-0.0998)	(-0.0395)	(-0.1162)	(-0.1024)	(-0.0761)
Monthly Precipitation	-0.00001***	-0.00001***	-0.00002***	-0.00001***	-0.00001***	-0.00002***	-0.00003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Monthly Mean Maximum Temperature	0.00065***	0.00110***	0.00175***	0.00066***	0.00111***	0.00177***	0.00143***
	(0.0000)	(-0.0001)	(-0.0001)	(0.0000)	(-0.0001)	(-0.0001)	(-0.0001)
Household Size	-0.00102	0.02789	0.02687	0.00133	0.02412	0.02545	-0.00211
	(-0.0054)	(-0.0287)	(-0.0308)	(-0.0053)	(-0.0289)	(-0.0312)	(-0.0051)
Logged Median HH Income	-0.00016	0.01875	0.01859	0.01451	0.05114	0.06565	-0.00246
	(-0.0084)	(-0.0324)	(-0.0368)	(-0.0082)	(-0.0318)	(-0.0359)	(-0.0069)
Percentage of houses built after 1992	1.32334	-3.05584	-1.73249	0.07781	-0.06894	0.00887	1.15162
	(-1.1867)	(-7.6091)	(-8.3241)	(-0.0457)	(-0.1387)	(-0.1468)	(-0.9081)
Area Of Lot	0.00005	0.00031	0.00036	0.00001***	0.00001*	0.00002***	0
	(0.0000)	(-0.0002)	(-0.0003)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Percentage Of Irrigable Land	-0.02575	-0.62694	-0.65269	-0.00289	-0.02971	-0.0326	0.07078
	(-0.0919)	(-0.5138)	(-0.5784)	(-0.0056)	(-0.0231)	(-0.0243)	(-0.0602)
Housing Density	0	0.00001	0.00001	0.00002	0.00009	0.00011	0
	(0.0000)	(-0.0001)	(-0.0001)	(-0.0001)	(-0.0003)	(-0.0003)	(.)
Water Usage Restrictions	-0.12879***	-0.22162**	-0.35041***	-0.12991***	-0.22169**	-0.35159***	-0.41965***
	(-0.0346)	(-0.0698)	(-0.1034)	(-0.0344)	(-0.0693)	(-0.1028)	(-0.0769)
Logged Price * Restrictions	0.06795**	0.11733*	0.18528**	0.06852**	0.11733**	0.18585**	0.24340***
	(-0.0236)	(-0.0458)	(-0.0690)	(-0.0235)	(-0.0454)	(-0.0685)	(-0.0515)

1. standard errors in parentheses; 2. significance level: *0.05 **0.01 ***0.001;

CHAPTER 7: CONCLUSIONS

7.1 Research Findings

This dissertation consists of three pieces of empirical work that aim to collectively contribute to the understanding of the dynamics of single family residential (SFR) water usage in Charlotte at a fine temporal and spatial resolution. The studies cover various issues from weather sensitivity, historical contingency, and relational assessment on water demand's determinants within a spatio-temporal framework. Using different datasets and analytical methods, a relatively complete picture of SFR water consumption as a coupled human and natural system is depicted in terms of its state, pattern and process.

We start with the overall state of "urban waterscape". After calculating and mapping the average annual and seasonal SFR water consumption per household during 2000-2010 at the block group level, we found that SFR water consumption is not evenly distributed across the study area and its overall patterns in winter and summer are consistent with its all-year-round pattern. The temporal trend of water consumption is analyzed by aggregating the average annual and seasonal SFR water consumption per household across all block groups. Not surprisingly, the temporal variations of annual water consumption are consistent with the weather dynamics in terms of mean temperature, cumulative precipitation and the Palmer Hydrological Drought Index (PHDI). However, the trends of summer and winter SFR water consumption differ.

The summer water usage follows closely the trend of PHDI and shows large variations during the entire time period, while the temporal variation of the winter water consumption is relatively large during 2002-2003 (Charlotte's first severe drought in the 21st century) and becomes small during 2007-2008 (the second severe drought period). The similar winter usage in 2007 and 2008 implies to some extent the effectiveness of the water usage restriction policy implemented during that time. Furthermore, when the year-to-year changes in water usage are examined, we observed the heterogeneous responses at the neighborhood level to weather variations.

In terms of weather sensitivity, we first use the peak factor measure to analyze its geographic patterns. As a ratio of summer usage over winter usage, the peak factor can be regarded as the variable measuring the extra water usage in hot seasons versus cold seasons. From the spatial distribution of the peak factor, we observe that the neighborhoods in the eastern and western Charlotte and along the I-77 South corridor have the lowest values, and the south of the county contains block groups with the highest peak factor as well as the greatest level of average consumption. The peak factor values were moderately higher in northern Mecklenburg, in which areas a few block groups had relatively large water usage. These findings provide evidence supporting the hypothesis that neighborhoods in Charlotte responded differently to atmospheric conditions.

Next, applying correlation and regression analysis on the monthly SFR water consumption data and weather datasets, we evaluated the sensitivity of water consumption in a statistical way. The major findings are: (1) despite the general low level of sensitivity compared to Phoenix, certain neighborhoods still exhibited mild but statistically significant sensitivity; (2) the temporal variations in water consumption explained by the three meteorological variables tend to cluster across geographical units; (3) high climatic sensitivity occurred in the neighborhoods with larger lots, more parcels with pools, larger and newer SFR houses, higher house value, or more bathrooms and bedrooms. Neighborhoods with higher income population and more owner-occupied housing units are associated with larger sensitivity. Climatic sensitivity decreases where a high proportion of the population is under 19 years old, and in neighborhoods with larger family, a higher percentage of Hispanics, or higher density.

The spatial pattern of the "urban waterscape" reveals evident spatial heterogeneity and significant spatial dependence. To uncover the processes underlying the patterns being observed, we adopt the statistical modeling approach.

Based on path dependence theory, we explore the role of historical contingency in modeling annual SFR water usage in 2008. The OLS model results show that most of the historical household and housing variables (in 2000) have larger significant influences than their temporal change during 2000-2008. When the spatial dependence in water consumption is incorporated via a spatial lag model, the model fit is improved and the majority of coefficient estimates decreases. The results from the GWR model show obvious spatial heterogeneity in the relationship between SFR water usage per household and the associated factors. The spatial variabilities exhibited in the local associations in terms of the historical variables and the change variables differ for most factors except median household income.

The concepts of spatial dependence and spatial heterogeneity are emphasized again in quantifying the relationship between monthly SFR water usage per household and its determinants in Charlotte under a spatial panel data modeling framework. After accounting for the possible spatial effects, the price, temperature, precipitation and water use restriction variables have significant and expected association with water consumption. For these variables, a large difference between the direct effects from each of the spatial panel models and the effects found in traditional panel data models is observed. The price effect becomes smaller with the intervention of water use restrictions policy in a spatial panel setting, while the non-spatial panel data model suggested the opposite conclusion.

Although we estimated the effects of household and housing factors, the results from all of the spatial panel models did not find a significant association between each factor and water consumption. The spatial Durbin model shows little evidence of spillover effects of the socioeconomic and housing variables.

It is the ultimate goal of this study to derive policy implications for water demand management from the analyses on the state, pattern and process of SFR water consumption. We summarize the specific findings and their policy implications here.

When comparing the spatial patterns of average summer water usage at the block group level, it is found that certain neighborhoods (especially in South Charlotte) consumed much more water in summer than winter, possibly due to their households' desire for lush lawns and/or outdoor activities. These neighborhoods have the greatest potential in terms of water consumption reduction and could be targets for discovering the underlying behavioral processes, and for experimenting policy interventions and conservation strategies.

The findings that the effects of some factors at a historical time point are larger than the counterparts of their temporal changes imply that the behaviors that tend to use more water (more likely for outdoor activity) may not experience a radical change probably due to a long-established lifestyle. We suggest it would be more useful to focus on the change in water consumption behavior itself rather than on the reconfiguration of physical or social structure measured by those factors.

The temporal change in area of irrigable land between 2008 and 2000 is significantly positively associated with the SFR water usage in 2008. The possible implication is that the new houses with larger or smaller than historically average irrigable lot size in a neighborhood may induce a larger increase or decrease in water usage, assuming the household preference to watering lawn or other vegetation does not change over time. The positive sign of the temporal change in housing density may reflect the diminishing effects of housing density (when increasing) on SFR water consumption.

The possible policy implications from the GWR analyses are that, for the factors with which history contingency dominates over temporal change, it is better to establish different policies or programs for the neighborhoods in the south versus the north of the Charlotte region, while for the factors whose temporal change exerted better effects, it would be useful to differentiate the intervention strategies for the east versus the west of the Charlotte region.

Although we have not been able to explain spatial dependence using an explicit mechanism or any theory that exists regarding water consumption behaviors, the general policy implication is to take a regional approach when directing water conservation efforts and resources to groups of spatially clustered neighborhoods with large water usage.

In summary, this study has strong social value and can help enhance the knowledge of local community on the history and present of SFR water consumption. The research findings highlight the impacts of price and water usage restrictions during and after the 2007-2008 droughts. It also advances our understanding of spatial and temporal dependence in modeling the relationship between SFR water use and some of its determinants.

7.2 Limitations and Future Work

We identified a number of limitations of this study from four aspects, namely data, methods, theoretical framework, and policy support.

To obtain a reliable longitudinal dataset is crucial to any study on the spatio-temporal dynamics of a phenomenon. Although we compiled the data from 2000 to 2010 for the socio-demographic variables, the datasets are from different sources and a majority of them are estimates derived from Census results, on which we have not given a comprehensive evaluation in terms of quality and measurement errors. The decennial census in 2010 are reported using the 2010 census geographies, and we have to calculate the variables for the 2000 census geographies based on the relations of the census boundaries in 2000 and 2010. The process itself is bound to introduce (random or systematic) errors to the data. We only have the datasets of parcels and buildings for a single year for calculating housing and urban structure related variables, and we make the assumption that housing characteristics are time-invariant, which is not true. The lack of variations in these variables (especially the urban structure variable) will prevent us from accurately evaluating their impacts on water consumption. A few variables such as the area of irrigable land and the area of pool are estimated based on incomplete information and simple assumptions. The weather variables from the only one local weather station in Mecklenburg County or from the regional climate division cannot capture the nuances in the micro-climate of neighborhoods. Being aware of all the data limitations mentioned above will make us more cautious and less frustrated when interpreting the model results. Some of the limitations may be overcome by seeking more data (e.g. use the historical parcel/building datasets, use high-quality remote sensing data to estimate the area of impervious areas, lawn and gardens and even derive the vegetation types, recent weather data collected from the multiple stations located in the county). We may consider census tract as the analytical unit since there are more and (presumably) better data source for census variables.

From the methodological perspective, there are several drawbacks in this dissertation. First, model misspecification issues are still relevant in our models in chapters 4 and 5. We can substitute the variables that have been used with different ones or add new variables to the models to see whether the models could be improved. Due to multicollinearity concerns, we dropped a number of variables that are identified as potential determinants of water consumption, implying that there may be omitted variables. One solution could be the use of principal components or factors in the models (although this will bring difficulties in interpreting the results and making suggestions on policy implications). Another alternative is to change the analytical framework from statistical models to machine learning approaches and structural equation models, and the latter is free of normality and multicollinearity concerns, and is particularly useful when the purpose is to predict water consumption. Second, our analyses are based on the aggregated data at block group level (despite its relatively small geography). The choice of this scale involves several methodological problems including the modifiable areal unit problem (MAUP) (Openshaw and Taylor 1979), the ecology fallacies (Openshaw 1984) and the uncertain geographic context problem (Kwan 2012). Under the current modeling framework, we can test the same models (such as spatial panel data models) at different geographical levels (e.g. census tract) to address MAUP. Our further direction is to employ multilevel modeling approach to understand the relationships between water consumption and its determinants at household level (premise or parcel as the proxies). Third, there exist other more complex methods to help examine the spatio-temporal dynamics of water consumption, for example, spatial Markov models and self-organizing maps. We could gain deeper insights on the patterning changes of water consumption by applying these methods. Fourth, as for spatial panel data models, it can be improved by adding temporally lagged effects that are commonly observed in panel data. Dynamic spatial panel data models will be useful for incorporating temporal dependence and spatial dependence simultaneously.

We have acknowledged the importance of behavioral and psychological determinants of water consumption in the conceptual model proposed in chapter 2, but we cannot explore those dimensions due to data availability. Although multilevel modeling using household level data seems promising, without the variables that directly measure or are associated with water consumption behaviors, the significance of such a study will be limited. We definitely need a more interdisciplinary empirical research dedicated to household survey and qualitative analysis to discover the underlying behavioral mechanisms of Charlotteans. We also call for a research agenda to explore the influences of social connectivity, social networking, and shared attitudes on water consumption behavior, and furthermore frame an explicit mechanism or even theory to explain spatial dependence we observed.

Future work may also be directed to other research themes related to policy making. This includes special studies on evaluating the effectiveness of pricing implemented in Charlotte and the equity issue that prices and their changes in time have involved, and on exploring the spatial heterogeneity in household or neighborhood reactions to water usage restrictions in Charlotte. Further efforts are required to fully understand the impacts of urban structure and land-use planning on water consumption. More importantly, we need an analytical framework that can integrate the mutual impacts between land development and water infrastructure in order to evaluate water- or land use-related policy and moreover achieve the sustainability in water use, land use, and urban development. The development of a useful water demand model to predict households' water consumption is part of the framework.

In summary, while this dissertation has answered a number of critical questions in urban water research, it also sets an open-ended research agenda, and we foresee a lot of research potential extending this study. We hope this research will raise the attention of local scholars and water authorities to participate and collaborate in future research opportunities for disentangling the problems pertaining to water management in a comprehensive manner.

REFERENCES

Aggarwal RM, Guhathakurta S, Grossman-Clarke S, Lathey V (2012) How do variations in Urban Heat Islands in space and time influence household water use? The case of Phoenix, Arizona. Water Resour Res 48:W06518. doi: 10.1029/

Agthe DE, Billings RB (1997) Equity and conservation pricing policy for a government-run water utility. Aqua - J Water Supply Res Technol 46:252–260.

Agthe DE, Billings RB (1987) Equity, Price Elasticity, and Household Income Under Increasing Block Rates for Water. Am J Econ Sociol 46:273–285.

Aitken CaK, Mcmahon TA, Wearing Brian L Finlayson AJ, Aitken C (1994) Residential water use: Predicting and reducing consumption. J Appl Soc Psychol 24:136–158.

Aitken CK, Duncan H, McMahon TA (1991) A cross-sectional regression analysis of residential water demand in Melbourne, Australia. Appl Geogr 11:157–165. doi: 10.1016/0143-6228(91)90041-7

Ajzen I (1999) The Theory of Planned Behavior. Organ Behav Hum Decis Process 50:179–211.

Ajzen I, Fishbein M (2000) Attitudes and the attitude-behavior relation: Reasoned and automatic processes. Eur Rev Soc Psychol 11:1–33. doi: 10.1080/14792779943000116

Allen TFH, Starr TB (1982) Hierarchy: Perspectives for Ecological Complexity. University of Chicago Press, Chicago

Anselin L (1995) Local indicators of spatial association - LISA. Geogr Anal 27:93–115. doi: 10.1111/j.1538-4632.1995.tb00338.x

Anselin L (2010) Thirty years of spatial econometrics. Pap Reg Sci 89:3–25.

Anselin L (1988) Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, Dordrecht, The Netherlands

Anselin L (2002) Under the hood: Issues in the specification and interpretation of spatial regression models. Agric Econ 27:247–267.

Anselin L (2003) Spatial Econometrics. In: Baltagi BH (ed) A Companion to Theoretical Econometrics. Blackwell, Malden, MA, p 728

Anselin L (1990) Spatial dependence and spatial structural instability in applied regression analysis. J Reg Sci 30:185–207. doi: 10.1111/j.1467-9787.1990.tb00092.x

Anselin L, Bera AK, Florax R, Yoon MJ (1996) Simple diagnostic tests for spatial dependence. Reg Sci Urban Econ 26:77–104. doi: 10.1016/0166-0462(95)02111-6

Anselin L, Gallo J, Jayet H (2008) Spatial panel econometrics. In: Mátyás L., Sevestre P (eds) The Econometrics of Panel Data. Springer-Verlag Berlin Heidelberg, pp 625–660

Anselin L, Griffith DA (1988) Do spatial effecfs really matter in regression analysis? Pap Reg Sci 65:11–34. doi: 10.1111/j.1435-5597.1988.tb01155.x

Arbués F, Barberán R, Villanúa I (2004) Price impact on urban residential water demand: A dynamic panel data approach. Water Resour Res 40:W11402. doi: 10.1029/2004WR003092

Arbués F, García-Valiñas MÁ, Martínez-Espiñeira R (2003) Estimation of residential water demand: A state-of-the-art review. J Socio Econ 32:81–102. doi: 10.1016/S1053-5357(03)00005-2

Arbués F, Villanúa I, Barberán R (2010) Household size and residential water demand: An empirical approach. Aust J Agric Resour Econ 54:61–80. doi: 10.1111/j.1467-8489.2009.00479.x

Arnold CA (Tony) (2005) Introduction: Integrating Water Controls and Land Use Controls: New Ideas and Old Obstacles. In: Arnold CA (Tony) (ed) Wet Growth: Should Water Law Control Land Use? Environmental Law Institute, Washington, D.C., pp 1–489

Askew LE, McGuirk PM (2004) Watering the suburbs: distinction, conformity and the suburban garden. Aust Geogr 35:17–37. doi: 10.1080/0004918024000193702

Baerenklau KA, Schwabe KA, Dinar A (2014) The residential water demand effect of increasing block rate water budgets. Land Econ 90:683–699.

Balling RC, Gober P (2007) Climate variability and residential water use in the city of Phoenix, Arizona. J Appl Meteorol Climatol 46:1130–1137. doi: 10.1175/JAM2518.1

Balling RC, Gober P, Jones N (2008) Sensitivity of residential water consumption to variations in climate: An intraurban analysis of Phoenix, Arizona. Water Resour Res 44:W10401. doi: 10.1029/2007WR006722

Bates BC, Kundzewicz ZW, Wu S, Palutikof JP (2008) Climate Change and Water. IPCC Secretariat,

Belotti F, Hughes G, Mortari AP (2013) XSMLE - A Command to Estimate Spatial Panel Models in Stata. In: 2013 German Stata Users Group Meeting. Potsdam, German, Billings RB, Agthe DE (1980) Price elasticities for water: A case of increasing block rates. Land Econ 56:73–84.

Breusch TS, Pagan AR (1980) The Lagrange multiplier test and its applications to model specification in econometrics. Rev Econ Stud 47:239–253.

Breyer B, Chang H, Parandvash GH (2012) Land-use, temperature, and single-family residential water use patterns in Portland, Oregon and Phoenix, Arizona. Appl Geogr 35:142–151. doi: 10.1016/j.apgeog.2012.06.012

Breyer EY (2014) Household Water Demand and Land Use Context: A Multilevel Approach. Portland State University

Burridge P (1980) On the Cliff-Ord test for spatial autocorrelation. J R Stat Soc B 42:107–108.

Cadenasso ML, Pickett STA, Grove JM (2006) Dimensions of ecosystem complexity: Heterogeneity, connectivity, and history. Ecol Complex 3:1–12. doi: 10.1016/j.ecocom.2005.07.002

Carter DW, Milon J. Walter (2005) Price knowledge in household demand for utility services. Land Econ 81:265–283.

Chang H, Parandvash GH, Shandas V (2010) Spatial variations of single-family residential water consumption in Portland, Oregon. Urban Geogr 31:953–972. doi: 10.2747/0272-3638.31.7.953

Charlotte Observer (2015) Charlotte population tops 800,000, new census estimates show. http://www.charlotteobserver.com/news/local/article21523356.html.

Charlotte Water (2015) About Charlotte Water. http://charmeck.org/city/charlotte/Utilities/AboutUs/Pages/About.aspx.

Charlotte Water (2016a) Charlotte Water Facts And Figures. http://charmeck.org/city/charlotte/Utilities/AboutUs/Pages/About.aspx. Accessed 1 May 2016

Charlotte Water (2016b) Utilities News (archived). http://cmudata.com/newsstoryarchive.cfm. Accessed 1 May 2016

Clark WA, Finley JC (2007) Determinants of water conservation intention in Blagoevgrad Bulgaria. Soc Nat Resour 20:613–627.

Cole MA (2004) Economic growth and water use. Appl Econ Lett 11:1–4. doi: 10.1080/1350485042000187435

Corbella HM, Pujol DS (2009) What lies behind domestic water use? A review essay on the drivers of domestic water consumption. Boletín la Asoc Geógrafos Españoles

50:297–314.

Corral-Verdugo V, Bechtel RB, Fraijo-Sing Blanca (2003) Environmental beliefs and water conservation: An empirical study. J Environ Psychol 23:247–257.

Cox W (2014) The Evolving Urban Form: Charlotte.

Coyne W (2003) The Fiscal Cost of Sprawl: how sprawl contributes to local governments' budget woes, Colorado.

Dalhuisen JM, Florax RJGM, Groot JhLF de, Nijkamp P (2003) Price and income elasticities of residential water demand: A meta-analysis. Land Econ 79:292–308.

Day D, Howe C (2003) Forecasting peak demand - What do we need to know? Water Sci Technol Water Supply 3:177–184.

Division of Water Resources NCD of E and NR (2014) Local Water Supply Plans - Charlotte Water. http://www.ncwater.org/Water_Supply_Planning/Local_Water_Supply_Plan/report.php?pwsid=01-60-010.

Domene E, Saurí D (2006) Urbanisation and water consumption: Influencing factors in the metropolitan region of Barcelona. Urban Stud 43:1605–1623.

Domene E, Saurí D, Parés M (2005) Urbanization and sustainable resource use: The case of garden watering in the metropolitan region of Barcelona. Urban Geogr 26:520–535. doi: 10.2747/0272-3638.26.6.520

Dziegielewski B (1999) Management of Water Demand: Unresolved issues. J Contemp Water Res Educ 114:1–7.

Eagly AH, Chaiken S (1993) The Psychology of Attitudes. Harcourt Brace Jovanovich College Publishers, Orlando, FL

Edwards M, Ferrand N, Goreaud F, Huet S (2005) The relevance of aggregating a water consumption model cannot be disconnected from the choice of information available on the resource. Simul Model Pract Theory 13:287–307. doi: 10.1016/j.simpat.2004.11.008

Elhorst JP (2010) Spatial panel data models. In: Fischer MM, Getis A (eds) Handbook of Applied Spatial Analysis. Springer-Verleg Berlin Heidelberg, pp 377–407

Elhorst JP (2012a) Matlab software for spatial panels. Int Reg Sci Rev 37:389–405. doi: 10.1177/0160017612452429

Elhorst JP (2012b) Dynamic spatial panels: models, methods, and inferences. J Geogr Syst 14:5–28. doi: 10.1007/s10109-011-0158-4

Forman RTT (1995) Land Mosaics: The Ecology of Landscapes and Regions. Cambridge University Press, New York

Fornango RJ (2010) When space matters: Spatial dependence, diagnostics, and regression models. J Crim Justice Educ 21:117–135. doi: 10.1080/10511251003693652

Fox C, McIntosh BS, Jeffrey P (2009) Classifying households for water demand forecasting using physical property characteristics. Land use policy 26:558–568. doi: 10.1016/j.landusepol.2008.08.004

Franczyk J, Chang H (2008) Spatial analysis of water use in Oregon, USA, 1985-2005. Water Resour Manag 23:755–774. doi: 10.1007/s11269-008-9298-9

Gage E, Cooper DJ (2015) The influence of land cover, vertical structure, and socioeconomic factors on outdoor water use in a Western US city. Water Resour Manag 29:3877–3890. doi: 10.1007/s11269-015-1034-7

Gato S, Jayasuriya N, Roberts P (2007) Temperature and rainfall thresholds for base use urban water demand modelling. J Hydrol 337:364–376. doi: 10.1016/j.jhydrol.2007.02.014

Gaudin S, Griffin RC, Sickles RC (2001) Demand specification for municipal water management: Evaluation of the Stone-Geary form. Land Econ 77:399. doi: 10.2307/3147133

Geller ES, Erickson JB, Buttram BA (1983) Attempts to promote residential water conservation with educational, behavioral and engineering strategies. Popul Environ 6:96–112. doi: 10.1007/BF01362290

Geolytics Inc. (2016) Annual Estimates for 2001 through 2009 - Methodology. http://www.geolytics.com/USCensus,Annual-Estimates-2001-2005,Data,Methodology,Products.asp.

Gibbs KC (1978) Price variable in residential water demand models. WATER Resour Res 14:15–18.

Gilg A, Barr S (2006) Behavioural attitudes towards water saving? Evidence from a study of environmental actions. Ecol Econ 57:400–414. doi: 10.1016/j.ecolecon.2005.04.010

Gleick P (1996) Basic water requirements for human activities: Meeting basic needs. Water Int 21:83–92. doi: 10.1080/02508069608686494

Gregory GD, Di Leo M (2003) Repeated behavior and environmental psychology: The role of personal involvement and habit formation in explaining water consumption. J Appl Soc Psychol 33:1261–1296.

Griffin RC, Chang C (1990) Pretest analyses of water demand in thirty communities. Water Resour Res 26:2251–2255. doi: 10.1029/WR026i010p02251

Grove JM, Pickett STA, Whitmer A, Cadenasso ML (2012) Long Term Socio-Ecological Research: Studies in Society-Nature Interactions Across Spatial and Temporal Scales. Springer, Dordrecht, Netherlands

Guhathakurta S, Gober P (2007) The impact of the Phoenix Urban Heat Island on residential water use. J Am Plan Assoc ©American Plan Assoc Chicagti 73:317–329.

Gutzler DS, Nims JS (2005) Interannual variability of water demand and summer climate in Albuquerque, New Mexico. J Appl Meteorol 44:1777–1787. doi: 10.1175/JAM2298.1

Harlan SL, Yabiku ST, Larsen L, Brazel AJ (2009) Household water consumption in an arid city: Affluence, affordance, and attitudes. Soc Nat Resour 22:691–709. doi: 10.1080/08941920802064679

Henderson B (2015) Protecting our water: We know what needs to be done. Charlotte Obs.

Hoffmann M, Worthington A, Higgs H (2006) Urban water demand with fixed volumetric charging in a large municipality: The case of Brisbane, Australia. Aust J Agric Resour Econ 50:347–359. doi: 10.1111/j.1467-8489.2006.00339.x

Höglund L (1999) Household demand for water in sweden with implications of a potential tax on water use. Water Resour Res 35:3853–3863. doi: 10.1029/1999WR900219

House-Peters L, Pratt B, Chang H (2010) Effects of urban spatial structure, sociodemographics, and climate on residential water consumption in Hillsboro, Oregon. J Am Water Resour Assoc 46:461–472. doi: 10.1111

House-Peters LA, Chang H (2011) Urban water demand modeling: Review of concepts, methods, and organizing principles. Water Resour Res 47:W05401. doi: 10.1029/2010WR009624

Hurd BH (2006) Water conservation and residential landscapes: Household preferences, household choices. J Agric Resour Econ 31:173–192.

Ingram H, Whiteley JM, Perry R (2008) The Importance of Equity and the Limits of Effi ciency in Water Resources. In: Whiteley JM, Ingram H, Perry R (eds) Water, Place, and Equity. The MIT Press, pp 1–32

Jorgensen B, Graymore M, O'Toole K (2009) Household water use behavior: An integrated model. J Environ Manage 91:227–236. doi:

10.1016/j.jenvman.2009.08.009

Kenney DS, Goemans C, Klein R, et al (2008) Residential water demand management: Lessons from Aurora, Colorado. J Am Water Resour Assoc 44:192–207. doi: 10.1111/j.1752-1688.2007.00147.x

Kwan M-P (2012) The Uncertain Geographic Context Problem. Ann Assoc Am Geogr 102:958–968. doi: 10.1080/00045608.2012.687349

Lackstrom K, Dow K, Ferguson S, Kettle N (2011) Understanding Needs for a Drought Early Warning System: Urban Water Systems in North Carolina. Columbia, SC

Lam SP (2006) Predicting intention to save water: Theory of planned behavior, response efficacy, vulnerability, and perceived efficiency of alternative solutions. J Appl Soc Psychol 36:2803–2824. doi: 10.1111/j.0021-9029.2006.00129.x

Lee SJ, Wentz EA (2008) Applying Bayesian Maximum Entropy to extrapolating local-scale water consumption in Maricopa County, Arizona. Water Resour Res 44:W01401. doi: 10.1029/2007WR006101

LeSage J, Pace RK (2009) Introduction to Spatial Econometrics. CRC Press, Boca Raton, FL

Lyman RA (1992) Peak and off-peak residential water demand. Water Resour Res 28:2159–2167. doi: 10.1029/92WR01082

Maidment DR, Miaou S-P (1986) Daily water use in nine cities. Water Resour Res 22:845–851. doi: 10.1029/WR022i006p00845

Maidment DR, Parzen E (1984) Time patterns of water use in six Texas cities. J Water Resour Plan Manag 110:90–106. doi: 10.1061/(ASCE)0733-9496(1984)110:1(90)

Martinez-Espiñeira R (2002) Residential water demand in the northwest of Spain. Environ Resour Econ 21:161–187.

Messner SF, Anselin L (2004) Spatial analyses of homicide with areal data. In: Goodchild M, Janelle D (eds) Spatially Integrated Social Science. Oxford University Press, New York, NY, pp 127–144

Mizzell Hope, Carbone G, Dow K, Rhee J (2010) Addressing monitoring needs for drought management. In: 18th Conference on Applied Climatology. Atlanta, GA,

Moran PAP (1950) A test for the serial independence of residuals. Biometrika 37:178–181.

Mylopoulos YA, Mentes AK, Theodossiou I (2004) Modeling residential water demand using household data: A cubic approach. Water Int 29:105–113. doi: 10.1080/02508060408691753

National Oceanic and Atmospheric Administration (2016) Monitoring Global and U.S. Temperatures at NOAA's National Centers for Environmental Information. https://www.ncdc.noaa.gov/monitoring-references/faq/temperature-monitoring.php.

Nauges C, Thomas A (2000) Privately Operated Water Utilities, Municipal Price Negotiation, and Estimation of Residential Water Demand: The Case of France. Land Econ 76:68–85.

Nieswiadomy M, Cobb S (1993) Impact of pricing structure selectivity on urban water demand. Contemp Policy Issues 11:101–113.

Nieswiadomy ML (1992) Estimating urban residential water demand: Effects of price structure, conservation, and education. Water Resour Res 28:609–615.

Nieswiadomy ML, Molina DJ (1989) Comparing Residential Water Demand Estimates under Decreasing and Increasing Block Rates Using Household Data. Land Econ 65:280–289.

Nordin JA (1976) A proposed modification on Taylor's demand–supply analysis comment. Bell J Econ 7:719–721.

North Carolina Division of Water Resources, Department of Environment and Natural Resources (2005) North Carolina Drought Management Advisory Council Activities Report - 2005.

Olmstead SM, Hanemann WM, Stavins RN, Kennedy JF (2003) Does Price Structure Matter? Household Water Demand Under Increasing-Block and Uniform Prices. Cambridge, Massachusetts

Olmstead SM, Michael Hanemann W, Stavins RN (2007) Water demand under alternative price structures. J Environ Econ Manage 54:181–198. doi: 10.1016/j.jeem.2007.03.002

Openshaw S (1984) Ecological Fallacies and the Analysis of Areal Census Data. Environ Plan A 16:17–31. doi: 10.1068/a160017

Openshaw S, Taylor P (1979) A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In: Wrigley N (ed) Statistical Application in the Spatial Sciences. Pion, London, pp 127–144

Ouyang Y (2013) Relationship between Single-family Residential Water Use and its Determinants - A Spatio-Temporal Study of Phoenix, Arizona. Arizona State University

Ouyang Y, Wentz EA, Ruddell BL, Harlan SL (2014) A multi-scale analysis of single-family residential water use in the phoenix metropolitan area. J Am Water Resour Assoc 50:448–467. doi: 10.1111/jawr.12133

Pfeffer MJ, Stycos JM (2002) Immigrant Environmental Behavior in New York City. Soc Sci Q 83:64–81.

Pickett STA, Cadenasso ML (1995) Landscape ecology: spatial heterogeneity in ecological systems. Science (80-) 269:331–334.

Pickett STA, White PS (1985) The Ecology of Natural Disturbance and Patch Dynamics. Academic Press, Orlando, FL

Pint EM (1999) Household Responses to Increased Water Rates During the California Drought. Land Econ 75:246–266.

Plümper T, Troeger VE (2004) The estimation of time-invariant variables in panel analyses with unit fixed effects. SSRN Electron J 15:124–139. doi: 10.2139/ssrn.565904

Polebitski AS, Palmer RN (2010) Seasonal residential water demand forecasting for census tracts. J Water Resour Plan Manag 136:27–36. doi: 10.1061/ʹASCÊWR.1943-5452.0000003

Polebitski AS, Palmer RN, Waddell P (2011) Evaluating water demands under climate change and transitions in the urban environment. J Water Resour Plan Manag 137:249–257. doi: 10.1061/(ASCE)WR.1943-5452.0000112

Porcher S (2013) Efficiency and equity in two-part tariffs: the case of residential water rates. Appl Econ 46:539–555.

Ramachandran M, Johnston RJ (2011) Quantitative Restrictions and Residential Water Demand: A Spatial Analysis of Neighborhood Effects. Worcester, MA

Randolph B, Troy P (2008) Attitudes to conservation and water consumption. Environ Sci Policy 11:441–455. doi: 10.1016/j.envsci.2008.03.003

Renwick ME, Archibald SO (1998) Demand side management policies for residential water use: Who bears the conservation burden? Land Econ 74:343–359.

Renwick ME, Green RD (2000) Do residential water demand side management policies measure up? An analysis of eight California water agencies. J Environ Econ Manage 40:37–55. doi: 10.1006/jeem.1999.1102

Rhoades SD, Walski TM (1991) Using regression analysis to project pumpage. J Am Water Work Assoc 83:45–50.

Russell S, Fielding K (2010) Water demand management research: A psychological perspective. Water Resour Res 46:W05302. doi: 10.1029/2009WR008408

Sanchez A, Vojinovic Z, Waly M (2011) Towards an approach to the evolution of urban drainage networks using agent-based models. In: 12th International Conference on

Urban Drainage. Porto Alegre, Brazil,

Schleich J, Hillenbrand T (2009) Determinants of residential water demand in Germany. Ecol Econ 68:1756–1769. doi: 10.1016/j.ecolecon.2008.11.012

Schneider ML, Whitlatch EE (1991) User-Specific Water Demand Elasticities. J Water Resour Plan Manag 117:52–73.

Sewell, W. R. D., Roueche L (1974) Peak load pricing and urban water management-Victoria BC a Case study. Nat Resour J 14:383–400.

Shandas V, Lehman R, Larson KL, et al (2015) Stressors and Strategies for Managing Urban Water Scarcity: Perspectives from the Field. Water 7:6775–6787. doi: 10.3390/w7126659

Shandas V, Parandvash GH (2010) Integrating urban form and demographics in water-demand management: An empirical case study of Portland, Oregon. Environ Plan B Plan Des 37:112–128. doi: 10.1068/b35036

Shin J-S (1985) Perception of price when price information is costly: Evidence from residential electricity demand. Rev Econ Stat 67:591–98.

Shove E (2003) Converging conventions of comfort, cleanliness and convenience. J Consum Policy 26:395–418. doi: 10.1023/A:1026362829781

Smith A, Ali M (2006) Understanding the impact of cultural and religious water use. Water Environ J 20:203–209. doi: 10.1111/j.1747-6593.2006.00037.x

Steg L, Vlek C (2009) Encouraging pro-environmental behaviour: An integrative review and research agenda. J Environ Psychol. doi: 10.1016/j.jenvp.2008.10.004

Stern PC (2000) Toward a coherent theory of environmentally significant behavior. J Soc Issues 56:407–424. doi: 10.1111/0022-4537.00175

Stevens TH, Miller J, Willis C (1992) Effect of price structure on residential water demand. J Am Water Resour Assoc 28:681–685. doi: 10.1111/j.1752-1688.1992.tb01489.x

Syme GJ, Shao Q, Po M, Campbell E (2004) Predicting and understanding home garden water use. Landsc Urban Plan 68:121–128. doi: 10.1016/j.landurbplan.2003.08.002

Taylor LD (1975) The demand for electricity: A survey. Bell J Econ 6:74–110.

Timmins C (2003) Demand-side technology standards under inefficient pricing regimes: Are they effective water conservation tools in the long-run. Environ Resour Econ 26:107–124. doi: 10.1023/A:1025689706396

Trumbo CW, O'Keefe GJ (2005) Intention to conserve water: Environmental values, reasoned action, and information effects across time. Soc Nat Resour 18:573–585. doi: 10.1080/08941920590948002

Urich C, Bach P, Hellbach C (2011) Dynamics of cities and water infrastructure in the DAnCE4Water model. Proc 12th Int Conf Urban Drain Porto AlegreBrazil 1015 Sept 2011 59:1–4. doi: 10.2307/1938769

Utilities C-M (2011) Strategic Operating Plan FY2012 – FY2013. Charlotte, NC

Ward MD, Gleditsch KS (2008) Spatial Regression Models. Sage, Thousand Oaks, CA

Weaver JC (2005) The Drought of 1998-2002 in North Carolina - Precipitation and Hydrologic Conditions.

Wentz EA, Gober P (2007) Determinants of small-area water consumption for the city of Phoenix, Arizona. Water Resour Manag 21:1849–1863. doi: 10.1007/s11269-006-9133-0

Wheeler D, Tiefelsdorf M (2005) Multicollinearity and correlation among local regression coefficients in geographically weighted regression. J Geogr Syst 7:161–187. doi: 10.1007/s10109-005-0155-6

Wiens JA (2000) Ecological Heterogeneity: An ontogeny of concepts and approaches. In: Hutchings MJ, John EA, Stewart AJA (eds) The Ecological Consequences of Environmental Heterogeneity. Blackwell, Malden, MA, pp 9–31

Worthington AC, Hoffman M (2008) An empirical survey of residential water demand modelling. J Econ Surv 22:842–871. doi: 10.1111/j.1467-6419.2008.00551.x

Zhou SL, McMahon TA, Walton A, Lewis J (2000) Forecasting daily urban water demand: A case study of Melbourne. J Hydrol 236:153–164. doi: 10.1016/S0022-1694(00)00287-0