The Development of a New Model for Predictions of Urban Water Demand

Seyed Alireza Eslamian

A Thesis

in

The Department

of

Building, Civil and Environmental Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Applied Science (Civil Engineering) at

Concordia University

Montreal, Quebec, Canada

February 2016

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Seyed Alireza Eslamian

Entitled: The Development of a New Model for Predictions of Urban Water Demand

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Civil Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Catherine Mulligan Chair

Dr. Marius Paraschivoiu Examiner

Dr. Ali Nazemi Examiner

Dr. S. Samuel Li Co-Supervisor

Dr. Fariborz Haghighat Co-Supervisor

Approved by ____

Chair of Department or Graduate Program Director

____

Dean of Faculty

Date: ____

Abstract

The Development of a New Model for Predictions of

Urban Water Demand

Seyed Alireza Eslamian

Population growth, rapid urbanization and climate change have been producing an ever increasing stress on the limited resources of fresh water. This situation makes prediction models of water demand an important tool for decision making regarding urban water management and conservation. However, there is still the need for improvement of the predictions of the existing models, especially in their estimations of base water use. The purpose of this thesis is to develop a more reliable water demand prediction model which gives a better understanding of water use behaviour. The employed techniques represent a new approach to predictions of daily water use; its base use component is predicted using a function of socioeconomic factors, as opposed to a function of time as in existing approaches. The prediction by the model proposed in this thesis is compared with those by two other existing models, in an application to the city of Brossard in Quebec. Time series of predicted daily urban water use captures observed characteristics very well and improves the results of the weighted coefficient of determination, the relative index of agreement and the root mean square error from the existing approaches. Water use in the city exhibits a downward trend possibly due to an increasing annual charge for water use. The analysis procedures reported in this thesis can be applied to analyze water use in any other cities. The new approach would be a useful tool for decision makers to manage water use by adjusting water consumption policies and price.

Acknowledgements

I would like to sincerely thank my supervisors, Drs. Li and Haghighat, for all of the time and effort they have put into helping me pursue my Master's program, and for their continuous support and immense knowledge during my research. Without their guidance and motivation, this research wouldn't have been possible.

I would also like to thank Mr. J. P. Richard of the City of Longueuil for providing data of daily water use, and Mr. T. Aldea for his help in the data acquisition process.

Lastly, I would like to express my gratitude to my parents and to my sister for always supporting and believing in me.

Table of Contents

List of Figures

List of Tables

List of Symbols

The following symbols are used in this thesis:

B	Average household annual payment for water consumption (CAD/household)
D	Index of agreement
D_r	Relative index of agreement
I	Average household income (CAD/household)
J	Rainfall occurrence dummy variable
K	Number of harmonics
N	Number of observations
O_i	Observed value
P	Precipitation amount (mm)
P_c	Rainfall Threshold (mm)
P_i	Predicted value
P_o	Population
R	Daily water use residual (litre per person per day)
R^2	Coefficient of determination
R_w^2	Weighted coefficient of determination
S	Weekday/weekend dummy variable
T	Daily maximum air temperature (^oC)
T_c	Threshold of air temperature (^oC)
T_n	Normal daily maximum air temperature (^oC)
U	Mean value of observed time series
W_b	Daily base water use (litre per person per day)
W_c	Daily calendrical water use (litre per person per day)
W_s	Daily seasonal water use (litre per person per day)
W_T	Total daily water use (litre per person per day)
$\overline{W}_r$	Residual of socioeconomic regression (litre per person per month)
$\overline{W}_T$	Grand mean of total water use (litres per person per month)

$\hat{W}_b$	Estimated base water use (litre per person per day)
$\hat{W}_{ce}$	Estimated water used due to climatic effects (litre per person per day)
$\hat{W}_{pc}$	Estimated persistence component of water use (litre per person per day)
$\hat{W}_{sc}$	Estimated water use with seasonal cycle (litre per person per day)
$\hat{W}_w$	Estimate water use due to weekday/weekend effect (litre per person per day)
a_j	Fourier coefficient
b	Slope of the trend line
b_j	Fourier coefficient
d	Daily time index
m	Month number
n	Order of polynomial function
p	Order of autoregressive procedure
r	Geometric growth rate
t	number of years from year 2011
y	Year number
α_0	Regression coefficient (litre per person per month)
α_1	Regression coefficient (litre per person per day)
α_2	Regression coefficient (litre per person per day)
β_0	Regression coefficient (litre per person per day)
β_1	Regression coefficient (litre per person per day)
β_2	Regression coefficient (litre per person per day)
γ_0	Regression coefficient (litre per person per day)
γ_1	Regression coefficient (litre per person per day)
γ_2	Regression coefficient (litre per person per day)
δ_0	Regression coefficient (litre per person per day)
δ_1	Regression coefficient (litre per person per day normalised by income)
δ_2	Regression coefficient (litre per person per day normalised by water-consumption charge)
δ_3	Regression coefficient (litre per person per day normalised by population)

ς_0 Regression coefficient (litre per person per day)

ς_1 Regression coefficient (litre per person per day per degree Celsius)

ς_2 Regression coefficient (litre per person per day per millimetres)

ς_3 Regression coefficient (litre per person per day per millimetres)

ς_4 Regression coefficient (litre per person per day per millimetres)

ς_5 Regression coefficient (litre per person per day per millimetres)

ς_6 Regression coefficient (litre per person per day)

η_0 Regression coefficient (litre per person per day)

η_1 Regression coefficient (litre per person per day)

φ Autoregression coefficient

List of Abbreviations

The following abbreviations are used in this thesis:

API	Antecedent Precipitation Index
CTA	Climatic Thresholds Approach
FA	Factor Analysis
FE-IV	Fixed Effect-Instrumental Variables
GDP	Gross Domestic Product
GMM	Generalised Method of Moments
IV	Instrumental Variables
LMAWUA	Lowest-Monthly-Average Water Use Approach
OLS	Ordinary Least Squares
PACF	Partial Autocorrelation Function
RMSE	Root Mean Square Error
SFA	Socioeconomic Factors Approach

1. Introduction

1.1. Motivation of Study

Water is essential for all life on the planet, and therefore is one of the most important natural resources. However, such important resources of fresh water are finite and do not equally distribute globally. Furthermore, with the ever growing population and rapid urbanization, along with climate change and pollution contributing to increasing water demand, stress on limited resources of fresh water is increasing day by day. Traditionally, water shortages have been deemed as an issue for drier, less water-abundant regions. Now, they have become a global concern, even for water-abundant countries. For example, Canada has one fifth of the world's fresh water and this perception of fresh water abundance in the society has resulted in an increase in the rate of water consumption. However, a large amount of fresh water in Canada is trapped inside distant glaciers and ice caps or is found in remote water bodies that are not easily accessible to populated regions thus, making it very costly and difficult to exploit (Brandes and Ferguson, 2003).

In order to avoid/alleviate problems related to water shortages, it is necessary to establish and implement scientifically sound measures, which requires a good understanding of water use behaviour. This means that factors with the biggest impact on water use must be identified, and accurate predictions of future water demand must be made.

1.2. Statement of the Problem

Through the analysis of water use time series, models for the prediction of water demand can be developed. Urban water use series may be considered as a sum of stochastic, periodic and trend components (Araghinejad, 2014; Hyndman, 2014). The challenge is how to realistically decompose a time series of water use into the components mentioned above. A realistic decomposition provides key information needed for the development of models for predicting water demand.

The detection of trend in time series has recently become one of the most significant components of time series analysis in water resources and environmental engineering (Araghinejad, 2014) and has received a great deal of research attention. In the field of urban water use, the trend of water use time series is regarded as base water use. It is the component of water use that changes gradually over long periods of time with variations in socioeconomic factors such as population size and water price. Base water use is believed to be climate-independent, as opposed to seasonal water use, the component of water use that is dependent on climatic variations. However, a water demand prediction model that can accurately capture the trend of water use as a completely climate-independent base water use has not yet been developed. Furthermore, the existing models commonly assume base water use as a function of time; while in reality, it is dependent on socioeconomic factors. Therefore, there is still the need for improvement of the water demand prediction models, especially in their estimation of base water use.

1.3. Objectives of the Research Work

The objectives of this study are:

- To develop a more reliable water demand prediction model than the existing models, which gives more accurate estimations of base water use.
- To reveal the structure and the behaviour of urban water use, and the effect of different factors on each of its components.
- To compare the performance of previously developed water demand prediction models with that of the model developed in this study, in order to determine the most accurate approach for the prediction of base water use.

1.4. Scope of the Work

The chosen city for this study is the city of Brossard, a part of Montreal metropolitan area located in the south shore of St. Lawrence River. The study requires a number climatic and socioeconomic data which were acquired from Environment Canada and Statistics Canada respectively. The daily water consumption data for the city of Brossard was obtained upon requesting from the City of Longueuil. Using the acquired data as input, statistical methods such as linear regression, Fourier series fitting and autoregression are employed to develop a model that can accurately predict each component of urban water use.

With the aim of effectively assessing the performance of the newly developed model, its performance is compared with that of two other previously proposed models. A number of model accuracy criteria are chosen and employed in order to achieve this.

1.5. Highlight of Research Contribution

The water demand prediction model developed in this study separates urban water use into distinct components and offers the most accurate yet simple method for predicting each of the components. The model can be applied for any other cities as well, providing an effective tool for understanding water use behaviour and decision making regarding the implementation of effective water demand management measures.

In the next section, a review of previous studies in the area of urban water use and some of their shortcomings is done. The area of study and the data used in the research are described in section 3. In section 4, the methodology is introduced and explained. Results are shown and discussed in sections 5 and 6 respectively, and conclusions are highlighted in section 7.

2. Literature Review

In this section, research previously conducted in the area of urban water use is explored. The main purpose of majority of the studies in this area has been to facilitate the prediction of future municipal water demand; either by identifying the various factors that may have an impact on water use, or by analyzing the urban water use time series.

There are simpler methods for estimating future water demand, such as using demand growth models as done by Van Rooijen et al. (2009). However, these methods assume that water use per capita stays constant throughout time, which is almost never true. Furthermore, according to these models, change in water consumption is only dependent on population growth. These basic assumptions make estimates of future water use very unreliable, especially for long-term water use predictions.

Unlike statistical methods such as regression where the goal is to predict water use based on a wide variety of factors (section 2.1), some studies focus on optimization techniques. Examples of optimization in the literature are determining what percentage of irrigators need to be retrofit in order to maximize benefit as a single family residential irrigation demand management strategy (Friedman et al., 2014); or minimizing total operation and scarcity cost in a region (Medellin-Azuara et al., 2007).

Other purposes of study in urban water use include: evaluating the efficiency or effectiveness of water use measures (Bian et al., 2014; dos Santos and Benetti, 2014) and simple case studies for analyzing urban water use conditions in different areas (Cahill and Lund, 2013; Furumai, 2008; Hurd et al., 2006; Rai, 2011; Wu and Tan, 2012).

Since the focus of this thesis is on developing a new urban water demand prediction model, the literature review focuses on studies that aimed to make the prediction of future urban water demand possible; which as mentioned earlier, make up the majority of studies in urban water use.

2.1. Identifying Factors That Impact Urban Water Use

Since climate change and socioeconomic driving forces have the biggest impact on future global water stress (Alcamo et al., 2007), the factors that have received the most attention range from climatic variables, e.g. air temperature, precipitation (Adamowski et al., 2012; BallingJr. and Gober, 2006; Guhathakurta and Gober, 2007; Gutzler and Nims, 2005; House-Peters et al., 2010; Mini et al., 2014; Stoker and Rothfeder, 2014; Wong et al., 2010), and seasonality (Ben Zaied and Binet, 2015; Polebitski and Palmer, 2010; Rathnayaka et al., 2015) to socioeconomic variables such as water price and household income (Abrams et al., 2012; Ben Zaied and Binet, 2015; De Mouche et al., 2011; Kenney et al., 2008; Martinez-Espineira, 2002; Mini et al., 2014; Panagopoulos, 2014; Romano et al., 2014; Schleich and Hillenbrand, 2009; Sohn, 2011; Turner and Ibes, 2011). Another set of factors that are often included are physical property characteristics such as lot size and dwelling age (House-Peters et al., 2010; Kenney et al., 2008; Rathnayaka et al., 2014; Sohn, 2011; Stoker and Rothfeder, 2014). Identifying the effect of different factors on water use in urban areas has become the predominant aim of study in researches involving urban water use. Because of that, when reviewing the previous work in this area, studies are often categorized by the variables they've taken into account (Arbués et al., 2003; Ferrara, 2008).

2.1.1. Climatic Factors

Climatic factors refer to weather-related variables. These factors are known as the variables that are beyond control, since they are independent of human manipulations. However, this doesn't mean that they have no effect on water use. Climatic factors usually appear as controlling factors when investigating the effect of certain variables on urban water use. Studies in the past decade that solely focus on the effect of climatic factors on urban water use are discussed in this sub-section.

Gutzler and Nims (2005) showed that the year-to-year change of water demand is highly correlated to the year-to-year change of temperature and precipitation in Albuquerque, New Mexico. In their study, only data from the summer months were included in the ordinary least square (OLS) model and the relation between changes in water demand and changes in temperature and precipitation wasn't determined for winter months. In addition, they couldn't determine which climate variable is the optimum predictor of urban water use. The data used in the study was based on monthly averages whereas using daily averages might help better determine the relation between climatic factors and water consumption.

BallingJr. and Gober (2006) investigated how annual water use in the city of Phoenix, Arizona is influenced by climatic variables. They argued that since mean annual temperature and total annual precipitation are highly correlated, they can't be used in multiple regression with both as predictors. Therefore, de-trended per capita annual water use was regressed against mean annual temperature and total annual precipitation separately. The results show that water use increases as temperature increases and precipitation reduces. Since the data used is on an annual timescale and the predictors of the regression analysis are very limited (only a few climatic factors such as air temperature and precipitation amount), the actual amounts of correlation between water use and the climate variables obtained in the study are very unreliable, even for the city of study itself. Furthermore, the city was undergoing water restriction policy implementations during the time period of study, but this factor was not taken into account, making the analysis inadequate.

Guhathakurta and Gober (2007) studied the effect of urban heat island on water use in the city of Phoenix, Arizona during June 1998 on a census-tract level. Using log-linear regression, the effects of nighttime temperature on residential water use, controlling for the presence of pools, vegetation type, size of house and lot, number of residents, and other socioeconomic, demographic, and housing variables were analyzed. The results showed that the effect of temperature (difference between high and low temperatures) was statically significant. The data included in the study is cross-sectional, making results valuable for analytical purposes rather than water demand prediction purposes.

Adamowski et al. (2012) chose the daily maximum temperature, daily total precipitation and daily water demand for the current day, the previous day, two days before, three days before and four days before, during the summer period, as the variables for forecasting future water demand in Montreal. Furthermore, the performance of different models including linear regression, nonlinear regression, artificial neural networks and wavelet artificial neural network were compared. Precipitation was not found to be a predictor of water demand in any of the models which is in contradiction with many other studies, where precipitation is usually proven to have a rather significant negative correlation with water consumption (Billings and Day, 1989; Bougadis et al., 2005; Gaudin, 2006; Gutzler and Nims, 2005; Polebitski and Palmer, 2010; Romano et al., 2014; Woodard and Horn, 1988). The models are for short-term water demand forecasting, since only climatic variables are accounted for. For long-term water demand prediction, socioeconomic variables need to be added (Adamowski et al., 2012). Furthermore, no models were developed to forecast water demand for the winter season. It has been reported that the occurrence of rainfall or even its forecast, regardless of its amount, affects water use (Maidment and Miaou, 1986; Martinez-Espineira, 2002; Rathnayaka et al., 2015; Schleich and Hillenbrand, 2009; Woodard and Horn, 1988) (but not always, see Bougadis et al. (2005)); therefore, rainfall occurrence may also be an important factor for short-term water demand forecasting.

2.1.2. Physical Property Characteristics

Since a big portion of urban water use belongs to residential water use, the effects of different variables that are related to the built environment are often of interest. It should be noted that built environment and demographic factors are usually included when the data is cross-sectional, as the change in property characteristics are usually not very significant throughout time. This makes the inclusion of this group of variables unnecessary for some studies that aim to predict future urban water demand.

Rathnayaka et al. (2014) used a set of physical property characteristics to investigate their effect on household water use. The variables were typology of dwelling, appliance efficiency, tenancy, dwelling age, the presence of swimming pool, evaporative cooler, and dishwasher, along with household size and the presence of children with respect to their age. OLS regression analysis showed that all variables except for tenancy, dwelling age and the presence of children between 12 and 18 years old affect water use in households. However, only few demographic factors' effect on household water use was investigated and no climatic variables were included in the study.

Stoker and Rothfeder (2014) assumed that water use is a function of climate (temperature and precipitation), built environment (number of bedrooms, number of kitchens, total bathrooms, lot size, the year the building was built, number of units in a building, number of stories in a building, turf fraction and number of lots on a commercial property) and demographic variables (value of the property and number of families in each building). They used an OLS regression model to investigate whether the relative importance of these variables changes for common urban land use types (single family residential, semi attached residential, apartments, and commercial). They were able to carry out this research for Salt Lake City, Utah, thanks to the availability of a rich database. Such disaggregate data is not usually available in other regions, making their methodology inapplicable for other areas. Moreover, water price was not included and the use of the value of property as an indicator of income might not always be accurate; thus, the model lacked two socioeconomic factors that had proven to be an important driver of water use (Ben Zaied and Binet, 2015; Kenney et al., 2008; Mini et al., 2014; Schleich and Hillenbrand, 2009). Stoker and Rothfeder (2014) found seasonality to be the greatest driver of urban water use. The characteristic of built environment with the greatest effect was revealed to be the size of the parcel. Seasonality has proven to be the major factor influencing urban water use in other researches as well (Salvador et al., 2011; Wong et al., 2010), along with lot size and outdoor irrigation area (Cleugh et al., 2005; Nakayama, 2011; Turner and Ibes, 2011).

2.1.3. Socioeconomic Factors

Among the socioeconomic factors that have received the most attention are population size, average household income and water price. Since certain socioeconomic variables are under utility control (such as water price), their effect on water use is important from water demand management perspectives. As water has no substitute for some basic uses, the relation between water price and water demand is often inelastic. However, as long as the elasticity is different from zero, water price can play an important role in water demand management (Arbués et al., 2003). Household income plays a role in this effect on urban water use as well. According to a study by Flörke et al. (2013), as average income increases, water users at first tend toward a more water-intensive lifestyle. Eventually, after a maximum level is reached, per capita water use is either stable or declines. National gross domestic product (GDP) has been studied as a driver of water demand as well (Cole, 2004; Wada et al., 2011). While Alcamo et al. (2007) argued that higher GDP increases urban water use, Hughes et al. (2010) showed that, in general, water uses per capita are greater in developing than developed countries due to low-tech water delivery and industrialization (Nazemi and Wheater, 2015). A big portion of studies on urban water use is dedicated to investigating the effect of socioeconomic variables on water use. Some of the most recent studies are discussed in this sub-section.

Kenney et al. (2008) postulated that water demand is a function of two sets of factors: factors under utility control (pricing and non-price strategies), and factors beyond the control of the water utility (weather related and economic-demographic variables) and investigated their effect on water demand in Aurora, Colorado. A major issue when dealing with the price variable is the endogeneity issue. This is because the price of water is often dependent on the amount of water consumed, making the predictor dependent on the dependent variable. Therefore, Kenney et al. (2008) used a fixed effect-instrumental variables (FE-IV). The Instrumental variables (IV) approach was also used by Schleich and Hillenbrand (2009) when dealing with price as a variable in their model. Both studies found water demand to be influenced by water price and consumers' income.

Sohn (2011) attempted to identify the effect of a large quantity of factors, predominantly socioeconomic, on urban water use for the southeastern part of the U.S. Since a large number of variables were to be included in the regression, factor analysis (FA) was carried out at first in order to group the variables that are highly correlated with each other into single factors. Since studies in urban water use often include a large number of regressors, FA has become a somewhat common solution to multicollinearity issues between the independent variables (Panagopoulos, 2014; Turner and Ibes, 2011). The findings in Sohn (2011) show that higher population size brings with it higher water consumption, and that water use decreases as density falls. Results also show that that water price is only effective for cities with light water use and not for cities that are heavy users of water. However, the opposite has been observed in some other areas as well (Ben Zaied and Binet, 2015). What is worth considering is that each region responds to price changes in a unique way with respect to given conditions. For instance, Zhang and Shao (2010) argued that cities with low economic levels that are more prone to water shortage are more sensitive to water price compared to richer, more water abundant cities.

Abrams et al. (2012) investigated the effectiveness of water usage pricing as a water demand management measure by estimating the short-run and long-run price elasticity of water demand in Sydney, Australia. They used household by household panel data and divided households by their tenancy type and size. To overcome endogeneity issues, generalised method of moments (GMM) was employed, which is a prominent method used in dynamic panel models with endogenous regressors. The results showed that long-run price elasticity is higher than short-run price elasticity, and that owner-occupied households are more responsive to price changes. The study was carried out during a time period when mandatory drought restrictions were enforced, which means the real effect of water price on water demand might have not been captured. However, the study shows that price has a bigger effect on log-term trend of water use, which is worth consideration. This has been proven in later studies as well. For instance, in a study for Tunisia, Ben Zaied and Binet (2015) found that at higher water consumption levels, the long-run elasticity of water price becomes significant.

Romano et al. (2014) attempted to estimate the determinants of residential water use in Italy using a linear mixed model. The variables included temperature, precipitation, water tariff, income, altitude and population. Their findings demonstrated that increasing water tariff causes reduction in the consumers' water consumption. A positive relationship was revealed between consumers' income and water use. It was also observed that cities with larger population have a higher water demand. They also investigated whether there is a difference in water consumption between areas where water utilities are publicly owned and areas where water utilities are privately owned. They found that the effect of water price prevailed the effect of water utilities ownership. When price variable was excluded, it was revealed that water consumption was significantly higher for areas where water utilities where publicly owned. However, this finding is not reliable because cities with publicly and privately owned water utilities differed in a variety of other characteristics as well. For instance, cities with privately owned water utilities usually had a higher population.

Mini et al. (2014) investigated the effectiveness of increasing block rate structure of water pricing in Los Angeles, California. The pricing consisted of an increasing block rate structure with a lower first tier rate corresponding to a specified water allotment, and a second higher tier rate for every additional billing unit. In the first tire, water charges were based on lot size and temperature zone; whereas in the second tire, water charges were directly tied to the amount of water consumed. Other controlling factors included income and household size, percentage grass cover, landscape greenness, air temperature and precipitation. Findings showed that price elasticity is higher in the first tire than the second tire. They also noted that higher water user and higher income census groups are slightly more sensitive to increases in the first tire rate than lower water user and lower income groups.

Panagopoulos (2014) employed FA in order to analyze demographic variables such as the number of active urban water connections and the population of the city, socio-economic factors including marginal price of water and the annual income of residents, and climatic variables that affect urban water use. It was found that the number of active connections, population size and annual income of residents have a significant and positive relationship with urban water use.

While the aforementioned studies aim to determine the amount of correlation between different factors and water use (predominantly by multiple regression), none of them reveal the structure of urban water use and its pattern over time. Furthermore, since water use is usually regressed against a large set of factors all at once in these studies, a long record of observed data is required (see Siauw and Bayen (2015)), which might not always be available. In order to be able to predict the pattern of water use over time, time series analysis is required (Maidment and Parzen, 1984b).

2.2. Water Use Time Series Analysis

In studies involving time series analysis, urban water use is usually decomposed into separate components in order to facilitate the analysis, and to illuminate the inherent structure of urban water use as well. In the literature, the most general and common decomposition of urban water use is into base water use and seasonal water use. As mentioned in section 1, base water use is the component of urban water use that is independent of weather and seasonal variations, and only dependent on socioeconomic factors such as population growth and water price; whereas seasonal water use is the component of water use that is dependent on weather and seasonality. Since the focus of this thesis is on the prediction of the trend of urban water use, previous studies are categorized by their assumptions about the base use component of urban water use time series here.

2.2.1. Base Use as Water Use in the Winter Season

House-Peters et al. (2010) defined base water use as the average water consumption in the winter season, equal to indoor water use, and defined seasonal use as total water use minus base use. They also introduced climate sensitivity as the ratio of seasonal use to water use. The dependent variables in the study were base use, seasonal use, drought sensitivity (ratio of 2006 seasonal use to 2004 seasonal use) and inter-annual climate sensitivity, with income, education, household size, population age, outdoor size, property lot size, age of the building and property value as independent variables. The results of OLS regression for Hillsboro, Oregon showed that indoor (base) use is dependent on household size whereas outdoor (seasonal) use is dependent on education level and the size of outdoor space. However, the influence of climate and water price was not analyzed which was a major limitation.

Although the definition of base water use introduced by House-Peters et al. (2010) as the average or lowest water use in the winter months has been used in previous studies as well (Maidment et al., 1985; Syme et al., 2004; Zhou et al., 2000), some studies have shown that indoor water use and water demand in the winter season are actually dependent on climatic variations. Rathnayaka et al. (2015) investigated whether seasonality affects indoor water use in addition to outdoor water use in Melbourne, Australia by regressing different end-uses against a set of climatic variables including minimum and maximum daily temperature, average rainfall and number of days with no rainfall. The end-uses included as dependent variables of the model were shower, bath, toilet, tap, dishwasher, clothes washer and irrigation end-uses. Using an OLS regression model, they came to the conclusion that shower and irrigation, which are the main end-uses, are significantly different between winter and summer. Their results also suggested that shower water use may increase with extreme weather conditions as cooler weather increases shower duration while hot weather increases shower frequency. The results are limited to the city where this research was carried out on, since indoor water use is greatly dependent on culture, attitudes and the built environment which can vary from one region to another. However, it must be considered that indoor use might be dependent on seasonality in other areas as well. In a study carried out by Gato et al. (2003), it was revealed that water use in the winter months may include outdoor water use such as garden watering. Furthermore, Gato et al. (2007) showed that base water use defined as the minimum water consumption in the winter is actually weather-dependent. Therefore, defining base water use as either indoor water use or water demand in the winter season might not be accurate all the time.

Gato et al. (2007) postulated that total water use is made up of base water use and seasonal water use. Base water use was calculated by two different methods and then compared. In the first method, the month with the lowest monthly water use was selected for each year to predict base use as a polynomial function of time, as done traditionally in previous studies (Maidment et al., 1985; Salas-La Cruz and Yevjevich, 1972; Zhou et al., 2000). The second method involved determining the temperature and rainfall thresholds at which water use is no longer dependent on temperature and rainfall. Water use on days with temperature below the temperature threshold and rainfall above the rainfall threshold were then considered as base water use. Base use estimated by the second method proved to be independent of air temperature and rainfall whereas the base use calculated by the first method proved to be dependent on climate variables when regressed against total daily rainfall and maximum air temperature. Base use obtained by the second method was then used in the water demand prediction model. Although base use is said to follow a trend dependent on socioeconomic factors, no socioeconomic variables were used in the regression of base water use. It was assumed that a regression against the day number and a dummy variable representing weekday or weekend, can be a proper substitute. Seasonal use was assumed to be made up of potential water use (smooth variation over the year with normal temperature) and climatic effects. Potential water use was represented by a Fourier series as done by Zhou et al. (2000) and the effects of climate were obtained by deducting potential use from seasonal use and regressing it against daily maximum temperature, daily rainfall, number of days after rainfall, day of the week and days with temperature above 35 °C. An autoregressive equation was fit to the daily residuals to account for the dependence of water use on its past values. The model yielded an R^2 of 83% during validation.

In a modelling study of urban water demand, Wong et al. (2010) introduced calendrical use as a component of total water use in Hong Kong; the other components were seasonal use and base use. Base water use was again equated as the lowest monthly water use in the winter season and modeled as a first-order polynomial function of time, assuming it can represent the trend in base water use caused by socioeconomic factors. Seasonal use was modeled as a combination of seasonal cycle and climatic effects. Fourier technique was used for characterizing the seasonal variations. Climatic effects were formulated by the removal of seasonal cycle from seasonal use followed by a regression against a number of substitute variables for air temperature and rainfall including daily maximum temperature, lagged maximum temperature, number of previous consecutive days with air temperature over 33 ^{o}C, number of days after rainfall and antecedent precipitation index (API). Calendrical water use was assumed to describe the effects of day of the week and holidays on water use. It was obtained by subtracting base use and seasonal use from the daily total water use. The day-of-the-week effect was formulated by regression against seven dummy variables each representing one day of the week. After subtracting base use, seasonality, climatic effect and day-of-the-week effect from water use, the remaining was regressed against dummy variables representing one day, two days and three days before a holiday, the day of holiday, one day, two days and three days after the holiday. Finally, an autoregressive procedure was fit to the water use residuals. Residuals were obtained after all of the previously estimated components of water use were subtracted from daily total water use. The model yielded an R^2 of 76% when validated with an independent dataset. The major factor influencing water use in Hong Kong was found to be seasonality. The results also showed more water use on the weekdays than the weekends, and reduction in water use on holidays as well as on one and two days after it, especially for the lunar new year. The holiday and day-of-the-week effect on water use may be subject to change significantly from a region to another as it depends on the culture, attitude and lifestyle of people. It was also not explored whether certain socioeconomic variables such as water price and consumers' income can improve base water use prediction. The model included the traditional polynomial function of time for base use although it was proven to be possibly climate dependent (Gato et al., 2007).

2.2.2. Base Use as a Function of Socioeconomic Factors

None of the aforementioned studies have used socioeconomic variables to predict base water use. Maidment and Parzen (1984a) estimated the trend of water use by regressing annually-averaged monthly water use against population, water price, household income and the number of water connections. Annually-averaged monthly water use was calculated as

$$\overline{W}_T(y) = \frac{1}{12}\sum_{m=1}^{12} W_T(m,y) \tag{2.1}$$

where $\overline{W}_T(y)$ is the annually-averaged water use and $\overline{W}_T(m,y)$ is the total water use in month *m* of year *y*. However, using average of total annual water use rather than average of annual base water use as the dependent variable in the regression will not give a clear and accurate estimation of the trend of water use. Furthermore, not only it won't ensure a climate independent base water use, but it will also result in the overestimation of base use. Moreover, using this approach, the de-trending process will become more complex which is as follows

$$W_b(m,y) = \frac{W_T(m,y)}{W_T(y)}\left[\overline{W}_T + \overline{W}_r(y)\right] \tag{2.2}$$

where $W_b(m,y)$ is the base water use, $\overline{W}_T$ is the grand mean of total water use and $\overline{W}_r$ is the residual of the aforementioned regression. In addition, this method of detrending is only appropriate for monthly water use data and isn't practical for data with a shorter time interval (i.e. weekly and daily data).

Only a few studies (Maidment and Parzen, 1984a, b) have discussed base water use as a function of socioeconomic factors rather than time; and few attempts were made to acquire a completely climate-independent base use (Gato et al., 2007). The review of the literature has highlighted the following knowledge gaps:

- There is a lack of models needed for reliable estimates of base water use.
- No comparisons of performance between socioeconomic-variable-based models of base water use and time-dependent function types of models have been made.

- It is not known how well the direct use of certain socioeconomic variables as base water use regressors will improve predictions of the trend of future water use.

This thesis will present a new model (referred to as Model C) for predictions of urban water demand, with application to the City of Brossard (in the greater Montreal region) in the Province of Quebec, Canada. The modelling results will help fill the above-mentioned knowledge gaps. The results will include a comparison of performance between the new model and two existing models (referred to as Model A and Model B) of different types.

3. Site of Study and Data

City of Brossard was chosen as a site of study mainly because of the completeness of required data and its characteristic urban development. Its longitude and Latitude are 45^{o}28’ and 73^{o}27’ respectively. The city is a part of the metropolitan area of Montreal on the south shore of the Saint Lawrence River, which is the source of drinking water for the city [Figure 3.1].

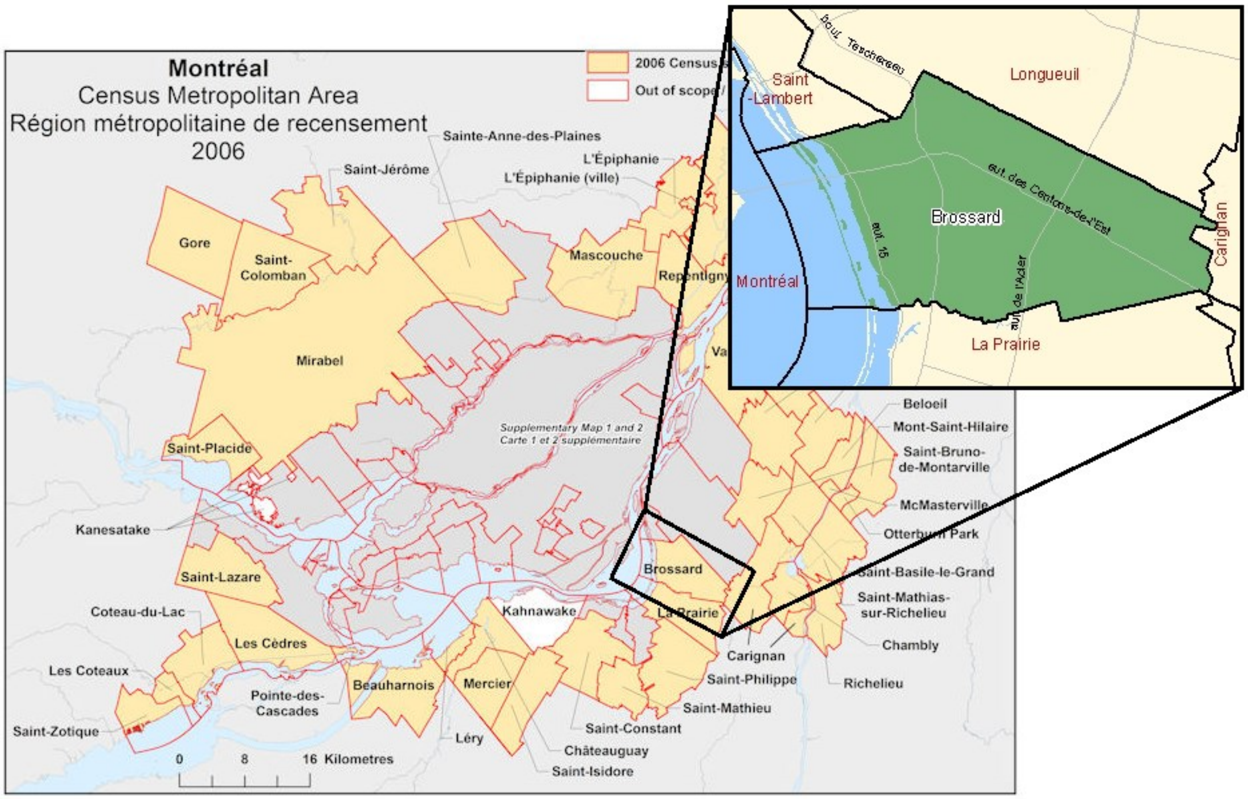

Figure 3.1. Location map of City of Brossard in the metropolitan area of Montreal (Source: 2006 census of Canada. Produced by the Geography division, Statistics Canada, 2011)

According to the most recent census conducted in 2011, the city had a population of 79273, representing a percentage growth of 11.4% from 2006 (the year of the previous census), in comparison to Canada’s national average growth of 5.9%. In 2011, the city had a population density of 1,753.9 persons per square kilometre, compared to Quebec’s provincial population density of 5.8 persons per square kilometre. From 2006 to 2011, there was an increase of 16.3% in private dwellings occupied by usual residents in the city, compared to 7.1% in Canada as a whole. The main land use types in Brossard are residential and commercial with many parks scattered through the city.

For model input and validation, the required data, which has been obtained from a number of sources, includes:

3.1 Water Use Data

Water supply to Brossard has been being served by the Le Royer water filtration plant, located in Brossard's neighbouring City of Saint Lambert [Figure 3.1], and operated by the City of Longueuil. Daily supply of water (in litres) to Brossard, covering a time period of about 54 months (represented by sequential number $m = 1, 2, 3,\ldots 54$) from May 2011 ($m = 1$) to October 2015 ($m = 54$) was provided by City of Longueuil. The water supply data is the total amount of water produced daily by the plant for distribution in Brossard. Presumably there are some water losses in the distribution network. This study has assumed that the losses are negligible.

Volumetric flow meters are installed in individual households as well as in public, commercial and industrial buildings in the City of Brossard. Annual readings of the meters are taken to give the actual amounts of water used by the users (all being referred to as households in this study). Details about the actual amounts of water used in individual households are not available to us. As an approximation, we divide the daily total supply of water to the city by its population to give the daily amount of water used per person, W_T. Time series of W_T is plotted in Figure 3.2(a). The duration of the time series is 1610 days (represented by sequential number $d = 1, 2, 3,\ldots,1610$), including both the starting date (May 15, 2011) and the ending date (October 10, 2015). The first 44 months-or 1327 days (May 15, 2011 to December 31, 2014)-are chosen as the raining period and the remaining (January 1, 2015 to October 10, 2015) as the testing period.

In each of the four calendar years, the lowest water use occurred in November (the corresponding sequential number is m = 7, 19, 31 and 43), whereas the highest value of W_T occurred in July (m = 3, 15, 27, 39 and 52).

3.2. Climatic Data

Daily maximum air temperature, T, and precipitation amount, P, for the Montreal region, covering the same time period as daily water use listed in this section was obtained from Environment Canada. Time series of T and P are plotted in Figures 3.2(b) and 3.2(c) respectively.

3.3. Socioeconomic Data

Population size, P_o, and household income, I, were acquired from Statistics Canada. Population data is only available in 5 year intervals. Therefore, assuming that annual population has a geometric growth rate, population for years with no population data available were calculated as

$$P_o(y) = P_o(2011) * r^t \tag{3.1}$$

with

$$r = \left(\frac{P_o(2011)}{P_o(2006)}\right)^{1/5}$$

where r is the geometric growth rate and t is the number of years between 2011 and year y.

For an individual household, the annual charge for water consumption is calculated by multiplying the actual amount of water used during the whole year by the price for water consumption. In Brossard, the pricing has an increasing block structure; where the price (rate) of water consumption increases for every 200 cubic meters. The water use rates were obtained from City of Brossard.

The average of the annual payments made by the individual households in the city, along with its population and average household annual income are listed in Table 3.1.

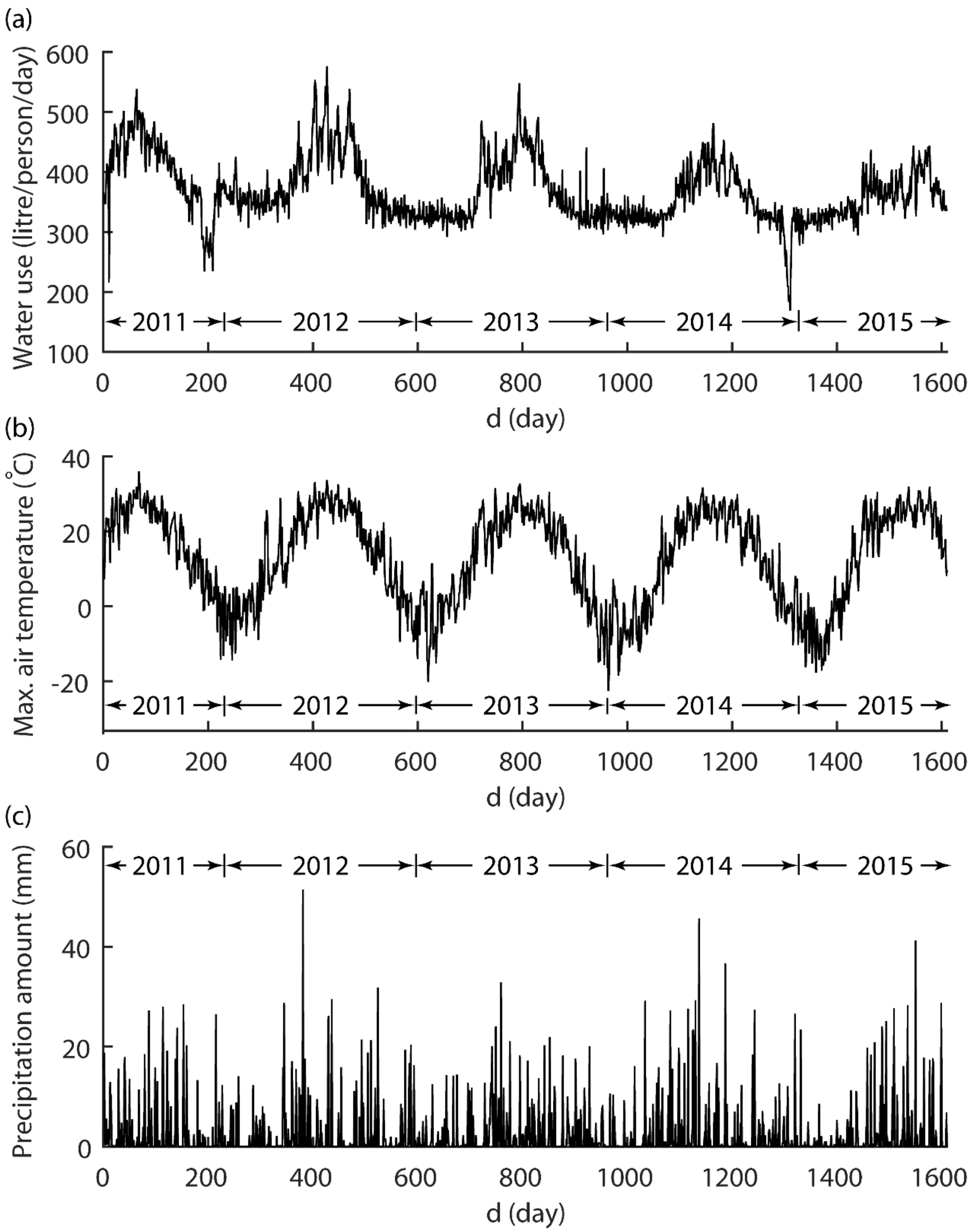

Figure 3.2. Time series of (a) metered daily water use (in litre per person per day) in the City of Brossard in the metropolitan area of Montreal, Quebec; (b) observed maximum air temperature; and (c) observed precipitation amount. The temperature and precipitation observations were made in the Montreal area. The duration of the time series is 1610 days. The starting date is May 15, 2011 and the ending date is October 10, 2015.

Table 3.1. A summary of population size (P_o) in the City of Brossard, average household annual income (*I*) and average household annual payment (*B*) for water consumption.

Year	***Po***	***I* (CAD/household)**	***B* (CAD/household)**
2011	79,273	67,933	191.31
2012	81,005	69,108	197.19
2013	82,774	70,283	201.99
2014	84,582	71,458	205.47
2015	86,430	72,633	227.50

4. Methodology

Following Wong et al. (2010), we express the total daily urban water use, W_T (in litres per person per day) as a sum of three main components, namely the base use W_b, the seasonal use W_s, and the calendrical use W_c

$$W_T(d) = W_b + W_s + W_c \tag{4.1}$$

In Equation (4.1), W_b is intended to describe the long-term trend in water use. W_s is the component of water use that is dependent on weather and climatic effects. It is split into two terms:

$$W_s = \widehat{W}_{sc}(d) + \widehat{W}_{ce}(d) \tag{4.2}$$

where the first term is an estimate of water use with a seasonal cycle, and the second term is an estimate of water use due to climatic effects. The seasonal cycle is the water use pattern due to variations in the normal temperature over the year, whereas the climatic effect is the short-term effect of climatic variability such as the occurrence of rainfall. W_c is also split into two terms:

$$W_c = \widehat{W}_w(d) + \widehat{W}_{pc}(d) \tag{4.3}$$

where the first term is an estimate of water use due to weekday/weekend effect, and the second term is an estimate of water use as the persistence component.

The decomposition of W_T as described above is schematically shown in Figure 4.1(a). The result of the decomposition of W_T as described above is an expansion of Equation (4.1), given by

$$W_T(d) = \widehat{W}_b(d) + \widehat{W}_{sc}(d) + \widehat{W}_{ce}(d) + \widehat{W}_w(d) + \widehat{W}_{pc}(d) \tag{4.4}$$

where $\hat{W}_b(d)$ is the estimated base use. All three water demand prediction models in this study will follow the same assumption regarding the decomposition of daily urban water use, with the main difference between them being the method and the definition employed in determining base water use.

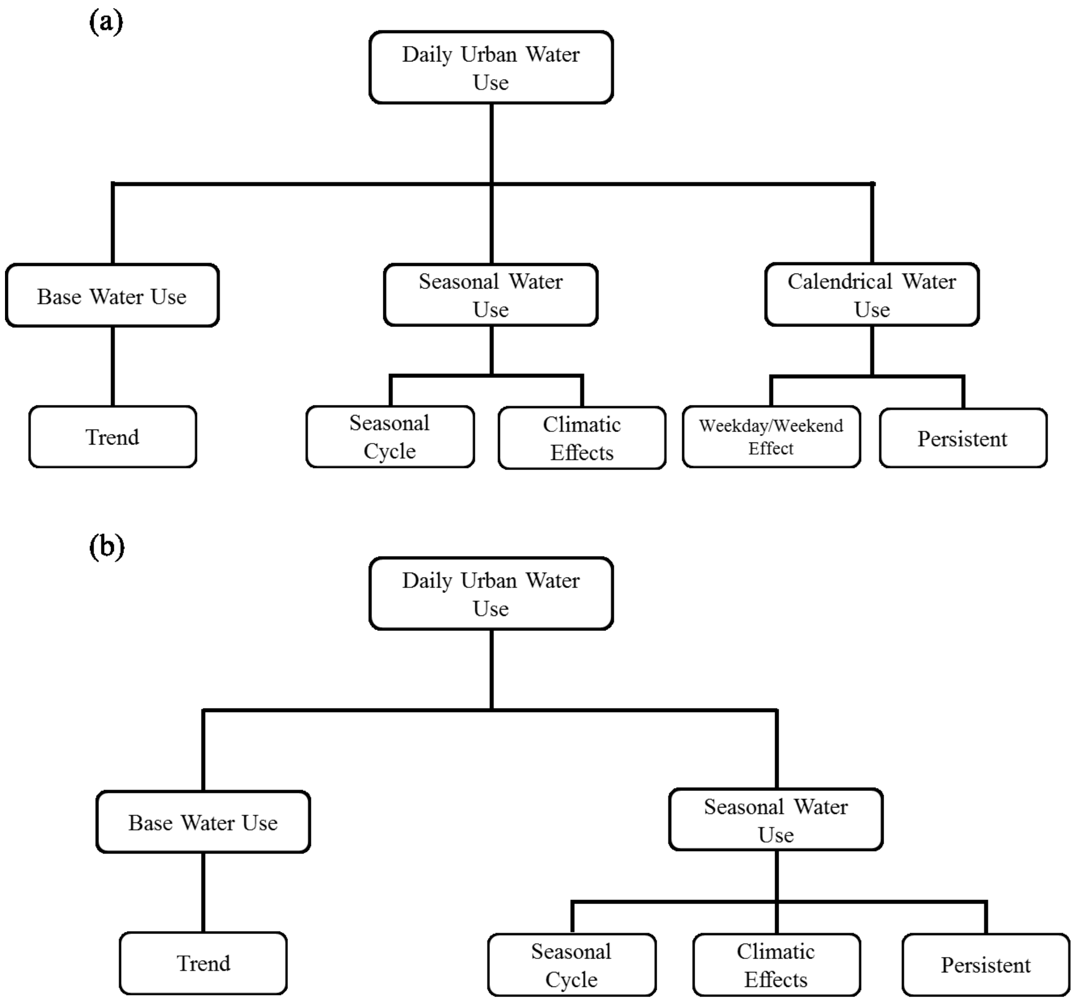

Figure 4.1. Decomposition of daily water use into several components for prediction. The component of base water use is predicted using: (a) model A (the lowest-monthly-average-water-use approach) and model C (the socioeconomic factors approach); and (b) model B (the climatic thresholds approach).

4.1. Base Water Use

As mentioned in section 1, base water use is the component of water use which is independent of changes in weather and climate. Base water use represents the long-term trend in water use caused by changes in socioeconomic factors such as change in population and water price.

4.1.1. Lowest-Monthly-Average-Water-Use Approach (LMAWUA Model A)

In model A, it is assumed that the base water use, W_b, in Equation (4.4) is represented by the lowest monthly average water use of each year. This assumption was used in Wong et al. (2010), Salas-La Cruz and Yevjevich (1972), and Zhou et al. (2000). W_b is estimated from a polynomial function of time of the general form

$$\widehat{W}_b(m) = \alpha_0 + \alpha_1 m + \alpha_2 m^2 + \cdots + \alpha_n m^n \tag{4.5}$$

where α_0, α_1, α_2,…, α_n (in litre per person per day) are regression coefficients, m is month number and n is the order of the function. Regression analysis can be performed to estimate the regression function with the lowest monthly-averaged water use in each year as the dependent variable, and the corresponding month number as the independent variable. One should try different orders in the regression analysis to minimize the residual.

In the case of the City of Brossard, the daily water use in November [Figure 3.2(a)] gives the lowest monthly average in each calendar year; the corresponding sequential number m is 7, 19, 31 and 43, representing the month of November of the years of record. Applying the regression model [Equation (4.5)] to the case will produce the following specific equations

$$\begin{cases} \widehat{W}_b(7) = \alpha_0 + \alpha_1 \times 7 + \alpha_2 \times 7^2 \\ \widehat{W}_b(19) = \alpha_0 + \alpha_1 \times 19 + \alpha_2 \times 19^2 \\ \widehat{W}_b(31) = \alpha_0 + \alpha_1 \times 31 + \alpha_2 \times 31^2 \\ \widehat{W}_b(43) = \alpha_0 + \alpha_1 \times 43 + \alpha_2 \times 43^2 \end{cases} \tag{4.6}$$

where the term on the left hand side of each of the equations is the monthly-averaged daily water use in each November of the years of record. The use of the second order (or n = 2) gives the lowest residual. The system of equations (4.6) describes a linear regression problem. Once estimates of the regression coefficients α_0, α_1 and α_2 are obtained, the regression model [Equation (4.5)] can be used to determine base water use for other months. Following the same procedure for daily water consumption, daily equivalent of equation (4.5) is then obtained as

$$\widehat{W}_b(d) = \beta_0 + \beta_1 d + \beta_2 d^2 + \cdots + \beta_n d^n \quad (4.7)$$

where β_0, $\beta_1,\ldots,\beta_n$ (in litre per person per day) are the new coefficients in the daily equivalent of equation (4.5) and d is the day number.

4.1.2. Climatic Thresholds Approach (CTA Model B)

In model B, the base water use, W_b, in Equation (4.4) is treated as a function of time and is estimated by determining climatic thresholds. For any given day d, the daily base water use, W_b (in litre per person per day) is expressed as

$$\widehat{W}_b(d) = \gamma_0 + \gamma_1 d + \gamma_2 S(d) \quad (4.8)$$

where γ_0, γ_1 and γ_2 (in litre per person per day) are regression coefficients; $S(d)$ is a weekday/weekend dummy variable. When the daily time index d is on a weekday, $S = 0$; otherwise, $S = 1$. Since $\hat{W}_b$ is regressed against S at this stage, there is no separate calendrical water use component in model B. Thus, Equation (4.1) will be reduced to

$$W_T(d) = W_b + W_s \quad (4.9)$$

And, equivalently, Equation (4.4) to

$$W_T(d) = \widehat{W}_b(d) + \widehat{W}_{sc}(d) + \widehat{W}_{ce}(d) + \widehat{W}_{pc}(d) \quad (4.10)$$

Equation (4.8) is intended to give estimates of the base water use that are independent of weather parameters (daily maximum air temperature, T, and precipitation amount, P, in this study). It has been assumed that there exist a certain threshold of temperature, T_c, and a certain threshold of precipitation amount, P_c. When both the daily maximum air temperature T is lower than T_c and the precipitation amount P is higher than P_c, the base water use, $\hat{W}_b$, is no longer sensitive to the actual values of T and P, as suggested by Gato et al. (2007). The determination of T_c and P_c is discussed below.

One way to identify T_c is to make a scatter plot of metered W_T versus observed T. Note that time series of W_T and T are shown in Figures 3.2(a) and 3.2(b), respectively. The scatter plot of (W_T, T) as data points uses Cartesian coordinates to display W_T values on the horizontal axis and T values on the vertical axis. A typical feature of the data points in the plot would be that W_T tends to increase with increasing T, at T exceeding a certain value. This value can be visibly identified from the plot, as is the case in this study. This value can be taken as the temperature threshold T_c (Adamowski et al., 2013). At $T < T_c$, the data points in the plot are more or less randomly distributed, without any identifiable trend.

To determine P_c, we make a scatter plot of metered W_T versus observed P. In the plot, W_T values are displayed on the horizontal axis, and P values on the vertical axis. Often, it is not feasible to visibly identify P_c from the plot. It is better to determine P_c through curve fitting. In this study, a polynomial function that has the best fit to (W_T, P) data points is constructed. The peak value of the function is taken as P_c (Gato et al., 2007).

From a time series of daily water us, those data points that satisfy the threshold conditions that $T < T_c$ and $P > P_c$ are selected as input data to regression analysis using Equation (4.8). This analysis gives output of γ_0, γ_1 and γ_2 values.

4.1.3. Socioeconomic Factors Approach (SFA Model C)

In model C, the base water use, W_b, in Equation (4.4), is formulated as a function of socioeconomic factors. In this study, the socioeconomic factors include: population, P_o, average household annual income, I, and average household annual payment, B, for water consumption. For any given year y, the annually-averaged daily base water use, $\hat{W}_b(y)$ (in litre per person per day), is formulated as

$$\widehat{W}_b(y) = \delta_0 + \delta_1 I(y) + \delta_2 B(y-1) + \delta_3 P_o(y) \tag{4.11}$$

where δ_0 (in litre per person per day), δ_1 (in litre per person per day normalised by income), δ_2 (in litre per person per day normalised by water-consumption charge), and δ_3 (in litre per person per day normalised by population) are regression coefficients; $I(y)$ is the average household income for year y; $B(y-1)$ is the average household annual payment for water consumption for the previous year; $P_o(y)$ is the population for year y. We can drop the last term on the right hand side of Equation (4.11) by using per capita daily water use, and reduce the equation to

$$\widehat{W}_b(y) = \delta_0 + \delta_1 I(y) + \delta_2 B(y-1) \qquad (4.12)$$

Usually, socioeconomic factors change from year to year (rather than from month to month or from day to day). Thus, regression analysis is performed for estimating the relationship [Equation (4.11)] between the annually-averaged daily base water use as the dependent variable and the socioeconomic factors as independent variables.

It is assumed in model C (like in model B) that air temperature threshold, T_c, and precipitation amount threshold, P_c, exist. The procedures for determining the temperature and rainfall thresholds have been explained in the previous section (Section 4.1.2).

From a time series of daily water use, those data points that satisfy the threshold conditions that $T<T_c$ and $P>P_c$ are selected as input data to regression analysis using Equation (4.12). This analysis gives output of δ_0, δ_1 and δ_2 values.

As noted in section 2, concerns about endogeneity issues arise whenever water price is used as one of the predictors of water use in OLS regression because water price itself is often dependent on water use (Kenney et al., 2008; Schleich and Hillenbrand, 2009). The price variable used in this regression represents the average annual payment per household for the total water consumed throughout the year and therefore is not dependent on the daily base water use. Furthermore, B corresponds to the previous year. Therefore, there is no endogeneity issue between $B(y-1)$ and $\hat{W}_b(y)$ and OLS regression can be applied.

4.2. Seasonal Water Use – Seasonal Cycle

Seasonal water use is associated with a seasonal cycle and climatic effects, expressed as the second and third terms, respectively, on the right hand side of Equation (4.4). The seasonal cycle presumably repeats itself in 365 days.

There are various methods for capturing the seasonal cycle. These methods include: formation of twelfth differences of data (Box et al., 2008), The heat function approach which involves regressing daily seasonal water use W_s during rainless periods against normal daily air temperature (Maidment et al., 1985), and estimation of seasonal cycle by fitting a Fourier series to seasonal water use. A Fourier series can adequately represent a seasonal cycle within the 365-day period, and the Fourier coefficients can easily be obtained from observed data, as pointed out in Zhou et al. (2000). Furthermore, the method is less complicated to apply than the heat function approach.Therefore, this study follows the Fourier series approach as done by Zhou et al. (2000) for the estimation of seasonal cycle for all models. Seasonal cycle represented by a Fourier series is as follows

$$\widehat{W}_{sc}(i) = a_0 + \sum_{j=1}^{K} (a_j cos\frac{2\pi j}{365} i + b_j sin\frac{2\pi j}{365} i) \tag{4.13}$$

where i = 1, 2, 3,… 365, representing the consecutive days of the year; a_0, a_j and b_j are the Fourier coefficients; K is the number of harmonics. The number of harmonics to be included in the analysis should be adjusted to achieve that the data of daily water use fits the statistical model well, with a satisfactory value for the coefficient of determination, R^2.

Data of daily water use, $\hat{W}_{sc}(i)$, as input to regression analysis using Equation (4.13) are derived in three steps: First, calculate the base water use $\hat{W}_b$, using Equation (4.7), (4.8) or (4.12). Then, subtract $\hat{W}_b$ from the total daily water use W_T to yield de-trended daily water use $W_T - \hat{W}_b$. Finally, take the average of $W_T - \hat{W}_b$ values for the same calendar date of different years. These average values give $\hat{W}_{sc}(i)$ with i = 1, 2, 3,… 365. The regression analysis produces output of a_0, a_j and b_j values. For any given date of the year, Equation (4.13) permits the determination of daily water use due to a seasonal cycle.

4.3. Seasonal Water Use – Climatic Effects

The climatic effects component, $\hat{W}_{ce}$, of the daily water use represents the short-term memory effect of climatic condition on water use. This component is expressed as

$$\widehat{W}_{ce}(d) = \zeta_0 + \zeta_1 \Delta T(d) + \zeta_2 P(d) + \zeta_3 P(d-1) + \zeta_4 P(d-2) + \zeta_5 P(d-3) + \zeta_6 J(d) \quad (4.14)$$

With $\Delta T(d) = T(d) - T_n(d)$

where ς_0, ς_1, … ς_6 are regression coefficients; $T_n(d)$ is the normal daily maximum air temperature; $d - 1$ refers to the previous day, $d - 2$ refers to two days before and d-3 refers to three days before; J is a dummy variable. $J = 1$ for days with rainfall, and $J = 0$ for days without rainfall. In this study, T_n is taken as the monthly-averaged daily maximum temperatures for 2000 throughout 2012. Since the effect of variations in normal temperature is assumed to have been captured by seasonal cycle, $\Delta T(d)$ has been used as a predictor here rather than $T(d)$ itself.

Data of water use $\hat{W}_{ce}$ as input to regression analysis using Equation (4.14) are the daily total water use minus the base water use [calculated using Equation (4.7), (4.8) or (4.12)] and minus the seasonal water use [calculated using Equation (4.13)]. The analysis produces output of ς_0, ς_1,… ς_6 values.

4.4. Calendrical Water Use due to Weekday/Weekend Effect

This component, $\hat{W}_w$, of water use refers to the component of water use that depends on the day of the week. This component appears in models A and C, but not in model B, which has incorporated the effect in estimation of the base water use. $\hat{W}_w$ is formulated as

$$\widehat{W}_w(d) = \eta_0 + \eta_1 S(d) \quad (4.15)$$

where η_0 and η_1 are the regression coefficients. Data of water use $\hat{W}_w$ as input to regression analysis using Equation (4.15) are the daily total water use minus the base water use [calculated using Equation (4.7), (4.8) or (4.12)], minus the seasonal cycle [calculated using Equation (4.13)], and minus the climatic effects [calculated using Equation (4.14)]. The analysis produces output of η_0 and η_1 values.

4.5. Persistence Component

The subtraction of the water use components, $\hat{W}_b$, $\hat{W}_{sc}$, $\hat{W}_{ce}$ and $\hat{W}_w$ from the total daily water use produces residuals of daily water use. An autoregressive procedure is fit to the water use residuals to determine the dependence of water use on its past values. The autoregressive model is given below

$$\widehat{W}_{pc}(d) = \varphi_0 + \varphi_1 R(d-1) + \varphi_2 R(d-2) + \cdots + \varphi_p R(d-p) \quad (4.16)$$

where $\varphi_0, \varphi_1, \varphi_2, \ldots \varphi_p$ are the autoregression coefficients, R is water use residual and p is the order of the autoregressive procedure. The optimum number of lags p is determined by using the partial autocorrelation function (PACF).

4.6. Model Accuracy Criteria

In order to compare the performance of the three models, a number of accuracy criteria were chosen.

4.6.1. Coefficient of Determination

The coefficient of determination (R^2) is often employed to show the amount of correlation between the observed and predicted values. However, it is advised to also take into account the slope of the fit trend line between observed and predicted values. This is because R^2 only describes how much of observed dispersion is explained by the prediction; meaning that it might not always capture how much the model over- or under-predicts (Krause et al., 2005). The closer the slope is to 1, the better the predictions are. In other words, the closer the slope of the trend line is to the $y = x$ line, the closer the predictions of the model to actual observed values are. Weighted R^2 (R_w^2) is used to combine R^2 and the slope for a better model assessment (Krause et al., 2005)

$$R_w^2 = \begin{cases} |b|.R^2, & if\ b \leq 1 \\ |b|^{-1}.R^2 & if\ b > 1 \end{cases} \tag{4.17}$$

where b is the slope of the trend line.

4.6.2. Index of Agreement

Proposed by Willmott (1981), The index of agreement (D) represents the ratio of mean square error and the potential error (Willmott, 1984). It was meant to overcome the insensitivities of R^2 to differences in observed and predicted means and variances (Legates and McCabe Jr., 1999). It is defined as

$$D = 1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(|P_i - U| + |O_i - U|)^2} \tag{4.18}$$

where O_i is the i-th observed value, P_i is the i-th predicted value, U is the mean value of the observed time series and N is the number of observations. The index of agreement has the same range as R^2 with 0 implying no correlation and 1 implying a perfect fit. However, D has also proven to not be sensitive to systematic model over- or under-prediction (Krause et al., 2005). Therefore, instead of D, the relative index of agreement (D_r) is used to compare the performance of the models. D_r not only greatly reduces the influence of the absolute differences during high water consumption periods, but it is also more sensitive to systematic over- or under-predictions (Krause et al., 2005). It is formulated as

$$D_r = 1 - \frac{\sum_{i=1}^{N}\left(\frac{O_i - P_i}{O_i}\right)^2}{\sum_{i=1}^{N}\left(\frac{|P_i - U| + |O_i - U|}{U}\right)^2} \tag{4.19}$$

4.6.3. Root Mean Square Error

A good and common measure for comparing the accuracy of models is the root mean square error (*RMSE*). It measures the average difference between observed and predicted values and is formulated as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(O_i - P_i)^2}{N}} \tag{4.20}$$

It should be noted that $RMSE$ is only applicable for comparing different models on the same set of data, as its scale depends on the scale of data (Hyndman and Koehler, 2006). Since the predictions from the three different models are compared to the same set of observations, there is no data-scale issue. A perfect prediction will have an *RMSE* of zero.

5. Results

5.1. Base Water Use

5.1.1. Estimate of Base Water Use by LMAWUA (Model A)

Fitting the November monthly averages of metered daily water use for the years of 2011, 2012, 2013 and 2014 to model A [Equation (4.5)] yields $\alpha_0 = 323.413$, $\alpha_1 = 0.986$, and $\alpha_2 = -0.024$. This results in a model equation of the form

$$\widehat{W}_b(m) = 323.413 + 0.986m - 0.024m^2 \tag{5.1}$$

Using this model equation, the base water use for the duration of May 2011 to December 2014 [or the first 44 months of the data records shown in Figure 3.2(a)] can be predicted. The base water use prediction by Equation 5.1 is plotted in Figure 5.1.

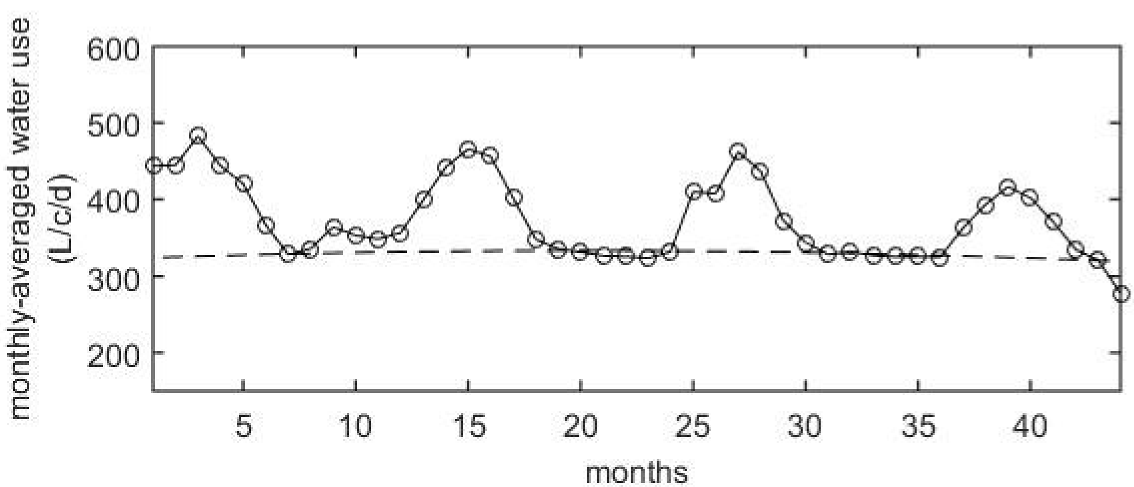

Figure 5.1. Prediction of base water use (dashed curve) using LMAWUA (Model A).

Following the same procedure for obtaining Equation 5.1, one can fit the data of metered lowest daily water use for the same years to model A, and obtain the daily equivalent of Equation (5.1). The resultant polynomial function is

$$\widehat{W}_b(d) = 270.132 + 0.109d - 7.9 \times 10^5 d^2 \qquad (5.2)$$

Equation (5.2) permits the predictions of the base water use for d = 1, 2, 3,…1327. It should be noted that daily water consumption in the month of December was excluded when obtaining Equation (5.2), since water consumption in that month may not really represent the base water use. This is because people tend to go on trips and leave the city during the holidays, increasing vacancy rate, and affecting water use during that period as a result.

5.1.2. Estimate of Base Water Use by CTA (Model B)

Model B for predictions of the base water use requires the determination of threshold values of air temperature and precipitation. A scatter plot of metered W_T [Figure 3.2(a)] versus observed daily maximum air temperature T [Figure 3.2(b)] is shown in Figure (5.2). It can clearly be seen in the plot that when $T > 10$°C, W_T increases with increasing air temperature, and this is not the case when $T < 10$°C. Therefore, the threshold T_c of air temperature can be taken as 10 °C.

A scatter plot of metered W_T [Figure 3.2(a)] versus observed precipitation amount P [Figure 3.2(c)] is shown in Figure (5.3). Fitting the data points to a polynomial function gives

$$P = -1.088 \times 10^{-4} W_T^2 + 0.084 W_T - 13.344 \qquad (5.3)$$

This function is shown as the dashed curve in Figure (5.3). In order to obtain the rainfall threshold, the derivative of P with respect to W_T was taken and set equal to zero. The function gives a maximum P value of 2.7. This peak value is the threshold of precipitation amount or $P_c = 2.7$ (mm).

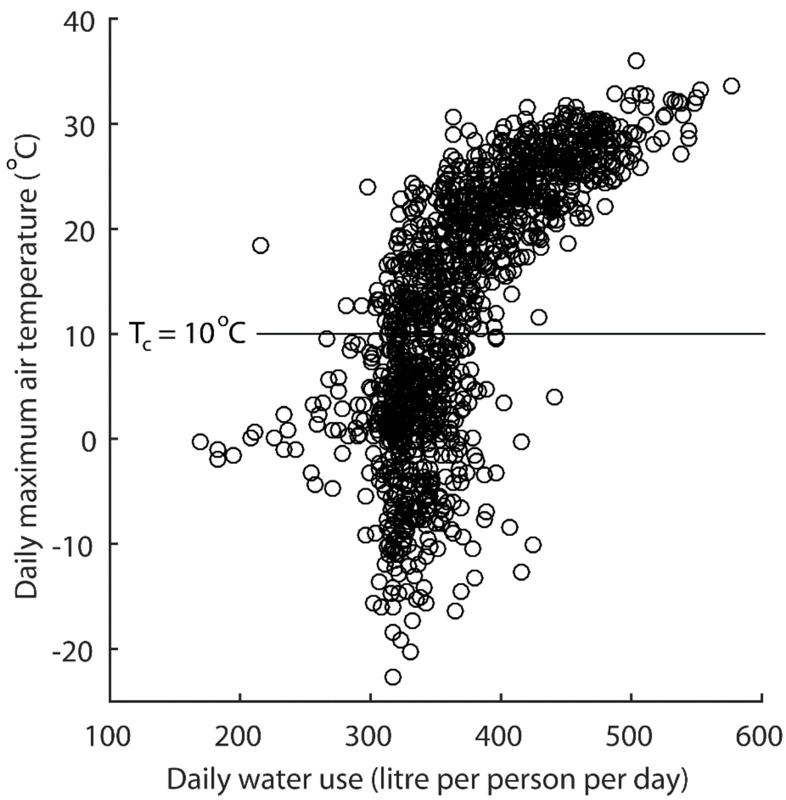

Figure 5.2. Scatter plot of metered daily water use versus observed daily maximum air temperature, based on the data shown in Figures 3.2(a) and 3.2(b) for the City of Brossard.

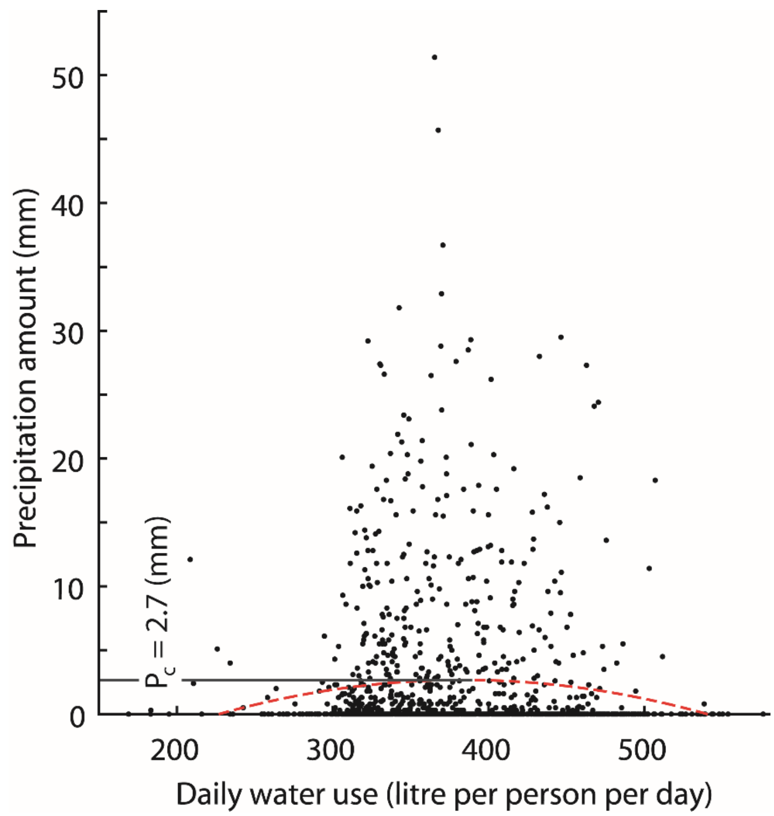

Figure 5.3. Scatter plot of metered daily water use versus observed precipitation amount, based on the data shown in Figures 3.2(a) and 3.2(c) for the City of Brossard. The dashed curve is a graph of the polynomial function [Equation (5.3)] fit through the data points.

Following the procedures described in Section 4.1.2, we determined the regression coefficients as $\gamma_0 = 357.351$, $\gamma_1 = -0.045$, and $\gamma_2 = 20.644$. The resultant model equation for predictions of the base water use is

$$\widehat{W}_b(d) = 357.351 - 0.045d + 20.644S(d) \tag{5.4}$$

The significance of the terms on the right hand side of the regression model is examined through p-value, which is associated with the probability that the null hypothesis is true; null hypothesis being that the variables are not statistically significant in the regression. The p-values are lower than 0.01, meaning that the coefficients obtained from the regression are statistically significant at a 99% confidence level. The regression has produced good and reliable results.

The negative sign of coefficient of d implies a downward trend in water use. According to this model, base water use rises on weekends. Figure 5.4 shows base water use as predicted by Equation 5.4.

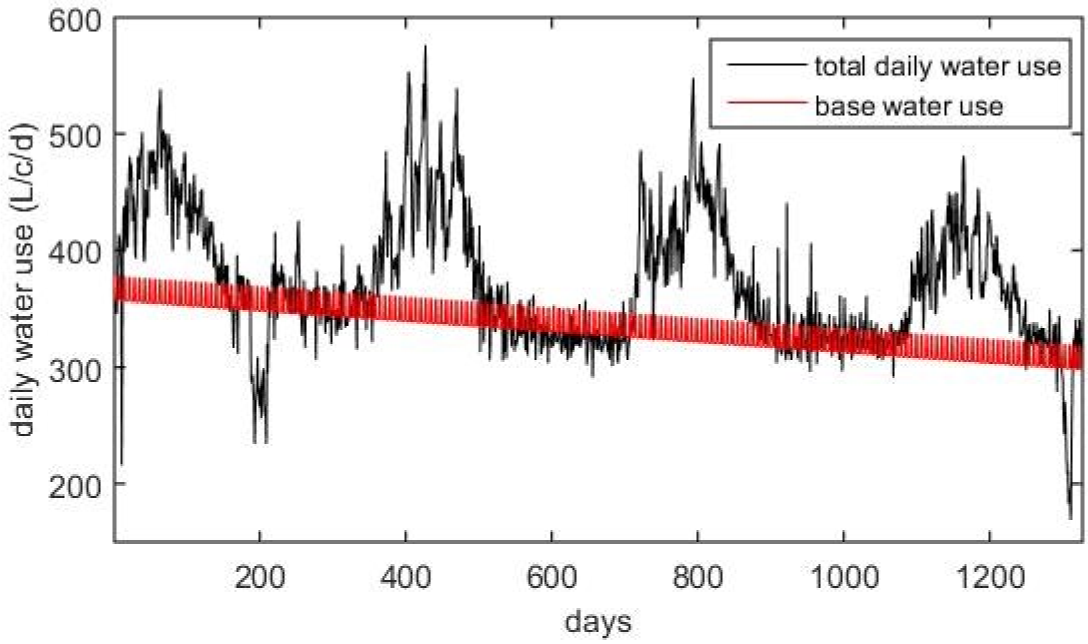

Figure 5.4. Prediction of base water use using CTA (Model B). The predictions show water use peaks for the weekend days

5.1.3. Estimate of Base Water Use by SFA (Model C)

The thresholds of air temperature and precipitation amount have been determined as $T_c = 10$ °C and $P_c = 2.7$ (mm), respectively in the previous section. Regression analysis was performed for each of the years of 2011 to 2014, following the procedures described in Section 4.1.3. The analysis used annually-averaged daily water use on days where the threshold conditions were satisfied as the dependent variable, and the average household annual income I and the average payment B for water consumption as the independent variables. When both I and B are included in the regression analysis, neither of the variables is significant at a 95% confidence level.

Regression analysis using B as the only independent variable gave an R^2 value as high as 0.97. The variable B itself is significant at a 99% confidence level. The analysis produced $\delta_0 = 903.531$ and $\delta_2 = -2.913$. The resultant model equation is given by

$$\widehat{W}_b(y) = 903.531 - 2.913B(y-1) \tag{5.5}$$

where y = 1, 2,…5, with y = 1 corresponding to the year of 2010. The p-values are lower than 0.01 for the terms on the right hand side of Equation (5.5), and thus the coefficients obtained are significant at a 99% confidence level. The negative sign of B implies a downward trend in water use which can be seen in Figure 5.5. Furthermore, the inverse relation between $\hat{W}_b(y)$ and B shows that as the costs for water consumption increases each year, a reduction in base water use is expected to happen in the following year.

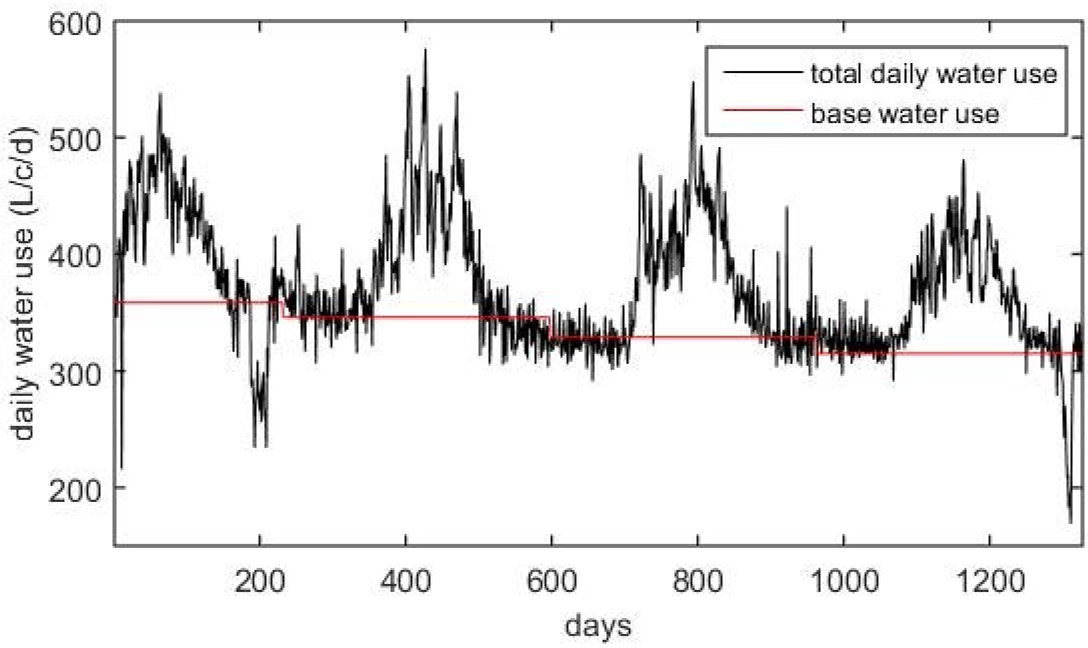

Figure 5.5. Prediction of base water use using SFA (Model C). It predicts the base water use changing from year to year.

5.2. Seasonal Cycle

The daily water use due to a seasonal cycle is expressed by a Fourier series [Equation (4.13)]. An increase in the number of harmonics (K) in the equation produces different values for individual Fourier coefficients. For example, even if the data of ($W_T - W_b$) as input to regression analysis is the same, a_1 for $K = 8$ is different from a_1 for $K = 5$. For each of the three models (models A, B and C), eight sets of Fourier coefficients were obtained. The predicted seasonal-cycle daily water use for $K = 8$ is not significantly different from that for $K = 7$. Thus, there is no need to include more than 8 harmonics in the analysis.

It should be noted that the predicted base water use W_b differs among models A, B and C. Therefore, the results of W_T minus W_b (known as de-trended water use) are different among the models. Since these different results are the input to Equation (4.13), we obtain three different sets of Fourier coefficients, as shown in Table 5.1. Using each set of the Fourier coefficients, we reconstruct a time series of seasonal-cycle water use and compare it with the de-trended time series [Figure 5.6]. The water use time series de-trended by Model A still shows a slight trend [Figure 5.6(a), dashed curve]. The coefficient of determination, R^2, has values of 0.68, 0.715 and 0.72 for models A, B and C, respectively.

Table 5.1. A summary of the Fourier coefficients [Equation (4.13)] for calculation of seasonal-cycle daily water use using the three different models.

Number of Harmonics (*K*)	**Coefficients**					
	LMAWUA (Mode1 A)		**CTA (Model B)**		**SFA (Model C)**	
	a_j	b_j	a_j	b_j	a_j	b_j
1	24.930	56.370	24.260	52.170	27.920	49.370
2	-17.970	18.050	-17.790	15.460	-20.180	15.290
3	1.268	2.948	1.033	1.135	1.898	2.614
4	0.882	-1.993	0.820	-3.088	1.170	-4.158
5	0.271	2.465	0.334	1.396	-0.551	1.853
6	6.311	-4.350	6.177	1.396	6.901	-4.807
7	0.862	3.068	0.849	2.431	0.681	1.804
8	2.336	2.194	2.302	1.523	1.937	2.026
Mean (a_0)	73.380		36.180		35.060	

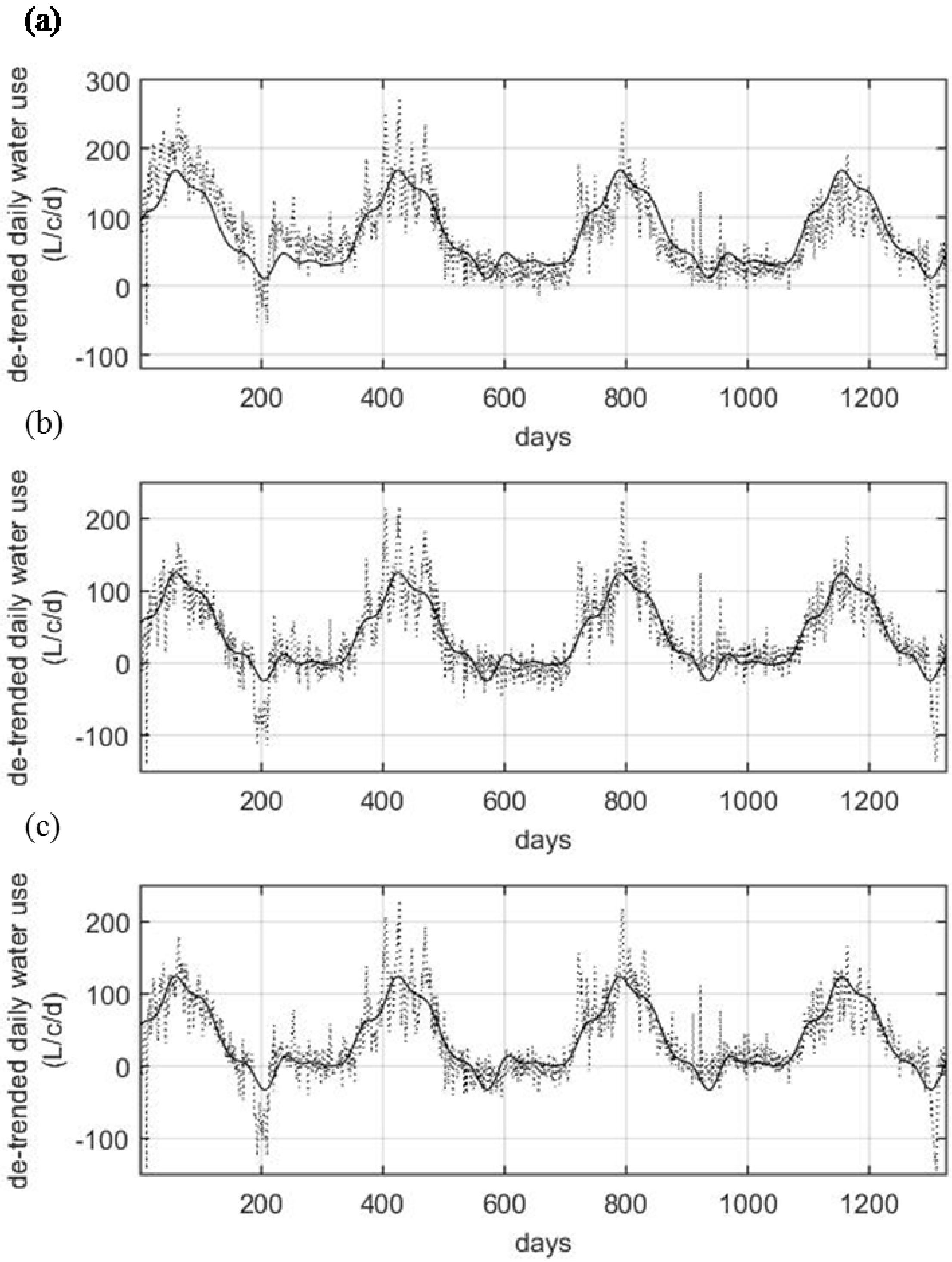

Figure 5.6. Seasonal Cycle (solid curve) fit to de-trended daily urban water use time series (dashed curve) using (a) Model A (b) Model B (c) Model C.

5.3. Climatic Effects

Values of the seven coefficients from regression analysis using Equation (4.14) are listed in Table 5.2. The coefficient ς_0 reflects the basic climatic effect; its values produced by the three models (models A, B and C) are different from each other to a limited extent. Positive values for ς_1 mean that the daily maximum air temperature in exceedance of the normal maximum air temperature causes an increase in water use. The coefficients associated with precipitation have negative values, causing a decrease in the current-day water use. Interesting to note is that ς_3 has a higher value than ς_2, ς_4 and ς_5, meaning that precipitation on the previous day has a bigger effect on water use than precipitation on the current day, two days ago and three days ago. The coefficient ς_6 also has negative values, the magnitude being as large as ς_0. This means that in the city of Brossard, the occurrence of rainfall, regardless of its amount, is expected to cause a significant decrease in water use and to offset the basic climatic effect (ς_0).

Table 5.2. A summary of coefficients from climatic-effects regression analysis [Equation (4.14)]. The differences in the coefficients among the different models result from different base water use produced by the models, and hence different input to the regression analysis. The listed p-values apply to all the three models (models A, B and C).

Coefficient	ς_0	ς_1	ς_2	ς_3	ς_4	ς_5	ς_6
Model A	7.526	1.5	-0.409	-0.815	-0.548	not significant	-7.344
Model B	7.070	0.921	-0.379	-0.837	-0.595	not significant	-6.820
Model C	7.898	0.901	-0.435	-0.874	-0.572	-0.309	-6.719
p-value	< 0.0001	< 0.0001	≤ 0.03	< 0.0001	< 0.001	0.0233	< 0.0001

5.4. Weekday/Weekend Effect

The procedures for regression analysis as well as input data to the analysis have been explained in Section 4.4. Regression analysis was performed using seven dummy variables, each representing one days of the week, as independent variables; the results (not shown) indicated that none of the independent variables were significant at a 95% or 90% confidence level. This was the case for all the three models (models A, B and C).

When the weekday/weekend dummy variable [S in Equation (4.15)] was used as the only independent variable, regression analysis showed that the variable was significant at a 99% confidence level. The weekday/weekend-effect regression result is the following for Model A

$$\widehat{W}_w(d) = -3.227 + 11.3S(d) \qquad (5.6a)$$

and the following for Model C

$$\widehat{W}_w(d) = -3.148 + 11.023S(d) \qquad (5.6b)$$

The large positive η_1 values mean much more water use on the weekend days in comparison to weekdays. This is presumably because people tend to spend more time in their houses on the weekend days, doing laundry, lawn watering and other water-consuming housework.

5.5. Persistence Component

An autoregressive procedure was applied to the time series of water use residual, resulting from the subtraction of all the other water use components (discussed in Sections 4.1 to 4.4) from metered daily water use [Figure 3.2(a)]. PACF was used to identify the optimum number of the lags (or the order p) in the autoregressive model [Equation (4.16)], and to ensure there was no autocorrelation left among the autoregression residuals. The resulted autoregression coefficients are presented in Table 5.3. The blank cells in Table 5.3 mean that the lag number had not been optimum for the corresponding model.

It is interesting to note that the current day water use is influenced, to some extent, by the previous day water use, as indicated by the relatively large values of φ_1 [Table 5.3 and Equation (4.16)].

Table 5.3. Autoregression coefficients from analyzes of water-use residuals produced by models A, B and C

Coefficient	φ_0	φ_1	φ_2	φ_3	φ_4	φ_5	φ_7
Model A	0.006	0.472	0.131	0.116	0.056	-	0.059
Model B	0.020	0.445	0.100	0.165	-	-0.067	0.076
Model C	0.010	0.451	0.124	0.118	-	-	-

6. Discussion

The main difference between the three models in this study is in the approach taken for the prediction of base water use. Figure 6.1 shows the predictions of base water use W_b for the city of study using the three different models. Model C not only estimates base water use that is completely independent of climate variables, but it also formulates the base water use as a function of socioeconomic factors [Equations (4.11) and (4.12)]. This approach offers a new alternative to existing approaches [Equation (4.5), referred to as model A, and Equation (4.8), referred to as model B] reported in the literature about estimates of base water use. The significance of the new approach is two-fold: First, from the management perspective, it helps form a solid scientific base for decision makers to achieve sustainability in urban water consumption and water conservation, to some extent, by adjusting socioeconomic factors (including charge for water consumption). Second, from the technical perspective, the new approach usefully separates different kinds of factors that can affect water use in cities.

Using the Brossard data [Figure 6.2, the black curve, the same as Figure 3.2(a)] for a selected training period as input to the analysis, we obtained Equation (5.2) for Model A, Equation (5.4) for Model B and Equation (5.5) for Model C, for base water use (Figure 6.1), and other regression coefficients [Tables 5.1 to 5.3; Equation (5.6)]. We further predicted daily water use using Equation (4.4) (Equation (4.10) for Model B). The predicted water use by the three models is plotted as a time series in Figure 6.2 (the red curve), along with the observed time series (Figure 6.2, the black curve).

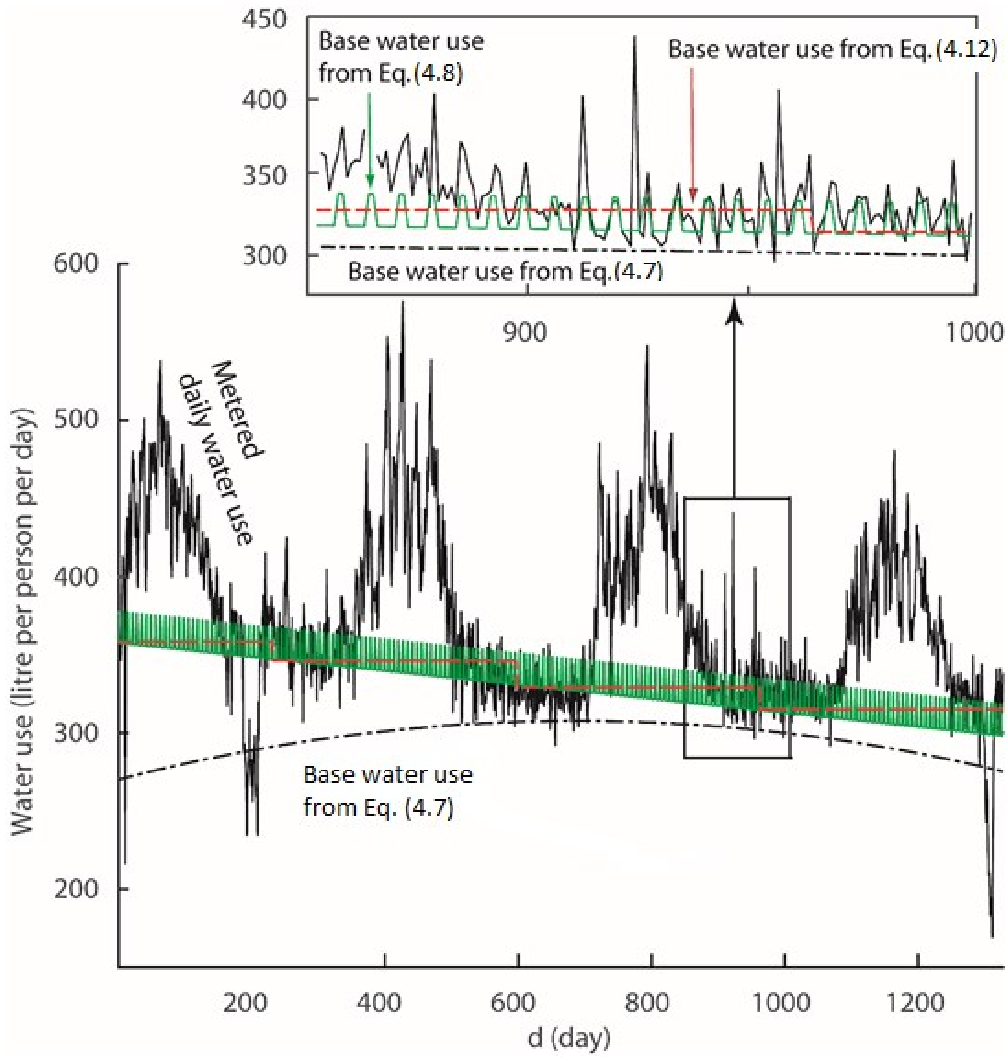

Figure 6.1. Predictions of base water use W_b for the City of Brossard, using the three different regression models. The CTA (model B) predicts water use peaks (the insert panel) for the weekend days. The SFA (model C) predicts the base water use changing from year to year (insert panel).

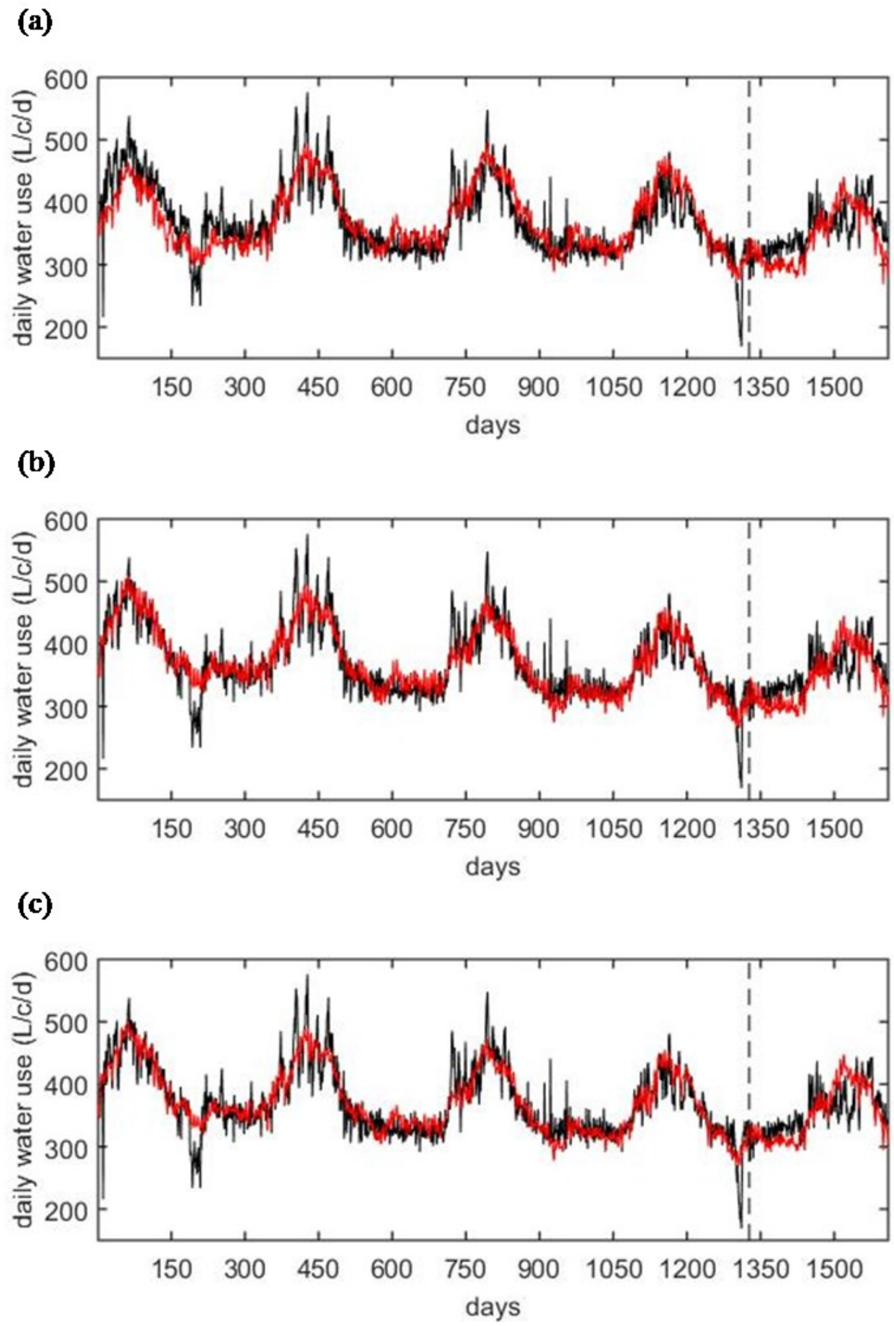

Figure 6.2. A comparison of daily water use time series between observations (the black curve) and predictions (the red curve) for the City of Brossard. The base water use component of the predictions is based on (a) Model A (b) Model B (c) Model C. The training period is from May 15, 2011 to December 31, 2014. The testing period (beyond the dashed line) is from January 1, 2015 to October 10, 2015.

A further quantitative comparison is shown in Figures 6.3(a) to 6.3(f). For each model, the time series of observed and predicted daily water use, shown in Figure 6.2, are presented in two scatter plots for the training and the testing period, respectively. When Model C is used, in both plots [Figures 6.3(a) and 6.3(b)], the data points are seen to cluster tightly about a perfect agreement line (the straight solid line, with a slope of 1); the slope of the data points is 0.78, as indicated by a trend line [Figure 6.3(a), the dashed line] added to the plot, for the training period, and is 0.89 [Figure 6.3(b), the dashed line] for the testing period. This means that the predictions are in a reasonable agreement with the observations, and hence confirms the quality of the prediction techniques that use model C for base water use [Equation (5.5)].

It would be interesting to learn the accuracy of water use predictions, whose base use component is predicted using models A, B and C. To the best of the author's knowledge, a systematic comparison has not been reported in the literature. When model B is used, the slope of the data points is 0.79 [Figure 6.3(c), the dashed line] for the training period, and is 0.88 [Figure 6.3(d), the dashed line] for the testing period. In this regard, model C and model B perform equally well. When model A is used, the slope of the data points drops to 0.73 [Figure 6.3(e), the dashed line] for the training period, and is 0.84 [Figure 6.3(f), the dashed line] for the testing period.

The technical performance of models A, B and C are further investigated by the model accuracy criteria introduced in section 4.6. The results are presented in Table 6.1.

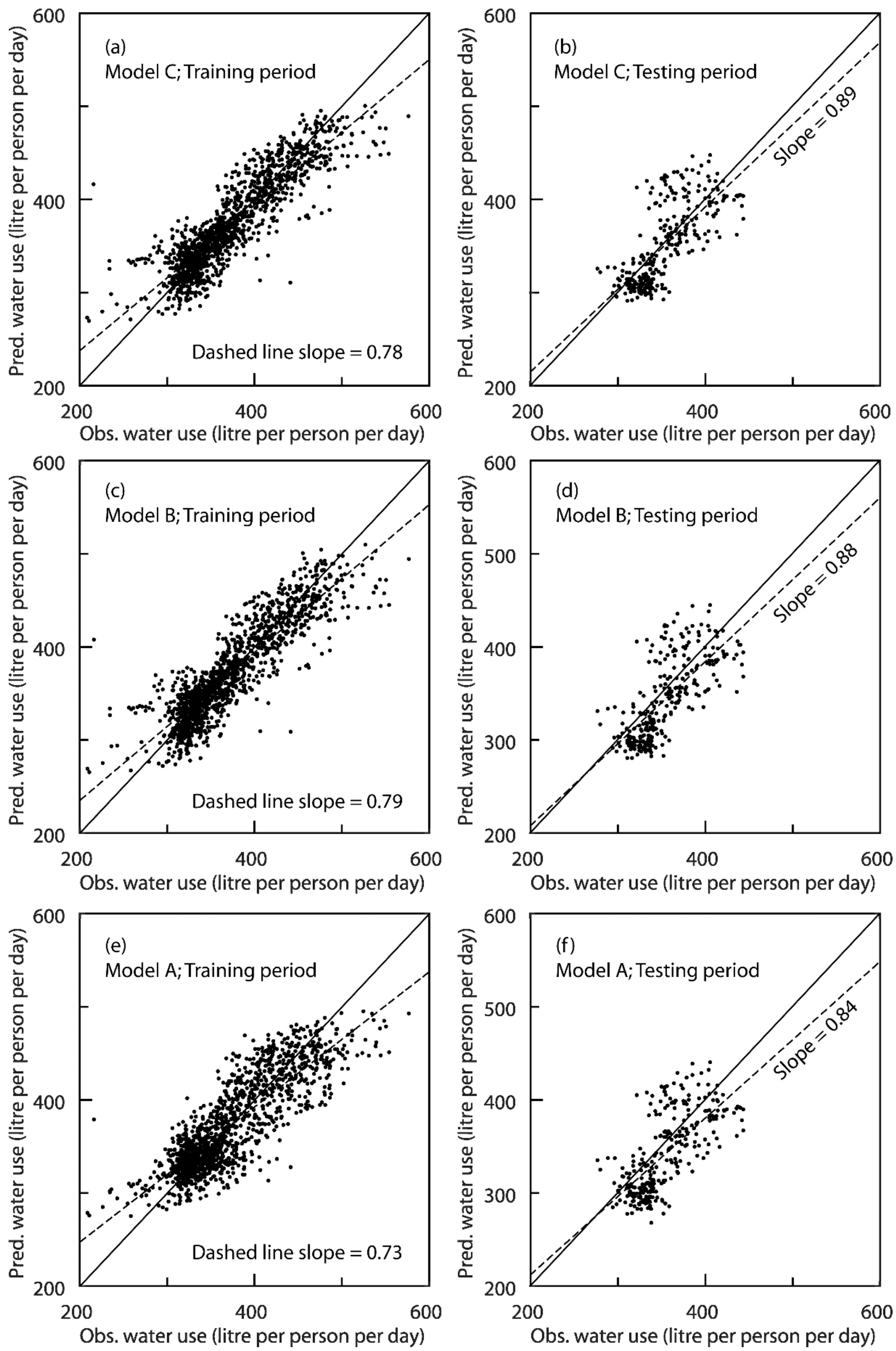

Figure 6.3. A comparison of daily water use between observations and predictions. The base water use component is predicted using the three different models. In each panel, the dashed line fits through data points. The closer to unity the slope of this line, the better the agreement of predictions with the observations.

Table 6.1. A summary of the weighted coefficient of determination (R_w^2), the relative index of agreement (D_r), and the root mean square error *RMSE* (in litre per person per day).

Performance index	LMAWUA(model A)	CTA (model B)	SFA (model C)
		Training Period	
R_w^2	0.52	0.61	**0.62**
D_r	0.90	0.92	**0.92**
RMSE	26.7	24.2	**23.2**
		Testing Period	
R_w^2	0.40	0.45	**0.47**
D_r	0.79	0.81	**0.84**
RMSE	31.2	30.1	**29.4**

According to all the criteria listed in Table 6.1, Model C appears to produce better results. If the basic idea in base water use formulation is to give climate-independent estimates, formulations based on lowest water use in the winter months may be inadequate. Arguably, this is the case when applied to cities like Brossard, where households are required to cover the actual costs of water consumption. In fact, the prediction of base water use, based on winter-month water use (Figure 6.1, the dashed-dotted curve), deviates from the actual trend of the metered daily water use. Model C is simple to use and gives reliable results. It is intended to consider the behaviour of base water use on the annual timescale, as opposed to a monthly timescale (Model A), and a daily timescale (Model B). Such a consideration would be more likely to achieve a climate-independent, accurate estimation. Furthermore, a time-dependent base use model is not able to capture the sudden changes (jumps) in the trend of water use time series. These jumps in the time series can be caused by an abrupt socioeconomic change such as a sudden rise or drop in the water price, or the implementation of water restriction policies. Since the base water use model in Model C is dependent on socioeconomic variables and changes annually, it has a better chance of capturing any jumps in the trend that might happen in the future.

As predicted by using Model C, there is a downward trend in water use in the city of Brossard (Figure 6.1). According to the model, this is possibly because of an increase in the costs of water consumption over the years (Table 3.1). This increase must have played a significant role in motivating households to continue reducing their water consumption. This is in contrast to an upward trend in water use in the neighbouring City of Montreal, as shown in the 1999 to

2002 record (Adamowski et al., 2013); that city has fixed fees for water consumption and included the fees in the property taxes for residential units (City of Montreal, 2016). This contrast evidences that socioeconomic factors, including policies and the price of water, have significant impacts on urban water demand.

Among the models presented in this study, only Model C was able to reveal the possible cause in the downward trend of water use in the city. As opposed to previous base use models, Model C identifies the effect of socioeconomic factors on the trend of water use. This makes the model very valuable for planning water demand management strategies. For instance, using the base use model in Model C (Equation 5.5), it is possible to predict how different price adjustment scenarios might influence the future trend of water use in the next year (2016) for the city:

- *Normal Increase in Water Price:* In this scenario, the price of water would increase at the same rate as the previous years. Model C predicts a higher drop in base water use than in the past years (green curve in Figure 6.4.). This is because the decrease in base water use between the years 2014 and 2015 is lower than that between other previous years (10 litres per person per day compared to about 16 litres per person per day); meaning that water use has not decreased much, while the price for water has kept increasing. This has caused the average charge for water consumption per household to increase significantly, compared to that of the past years (from 205.47 CAD to 227.50 CAD, see Table 3.1). Therefore, it is expected that the residents of the city be urged and motivated to reduce their water use rather significantly for the next year.

- *No Change in Water Price:* In this scenario, the price of water for the year 2015 is the same as that of 2014. According to Model C, base use will decrease in the coming year (Blue curve in Figure 6.4.), but not as much as it would have if water price had increased as always (Green curve in Figure 6.4.). It is even possible that the trend becomes constant or even upward, if the price of water keeps not changing in the following years (the effect of increasing income might prevail over the constant water price).

- *Reduction in Water Price:* Another scenario can be water price reducing to that in 2013. In this case, the model predicts base water use increasing as the trend of water use begins to become upward (Red curve in figure 6.4.). The reason this happens is that in this scenario, the average water consumption charge per household in 2015 becomes actually less than that of 2014. When the residents of the city experience this, they realize that it's not necessary to decrease their water consumption in the next year in order to reduce the amount they have to pay for it. Therefore, they are discouraged from conserving water in the coming year. Interestingly, as seen in figure 6.4, the base use increases back to about the same value it had during 2013 in this scenario.

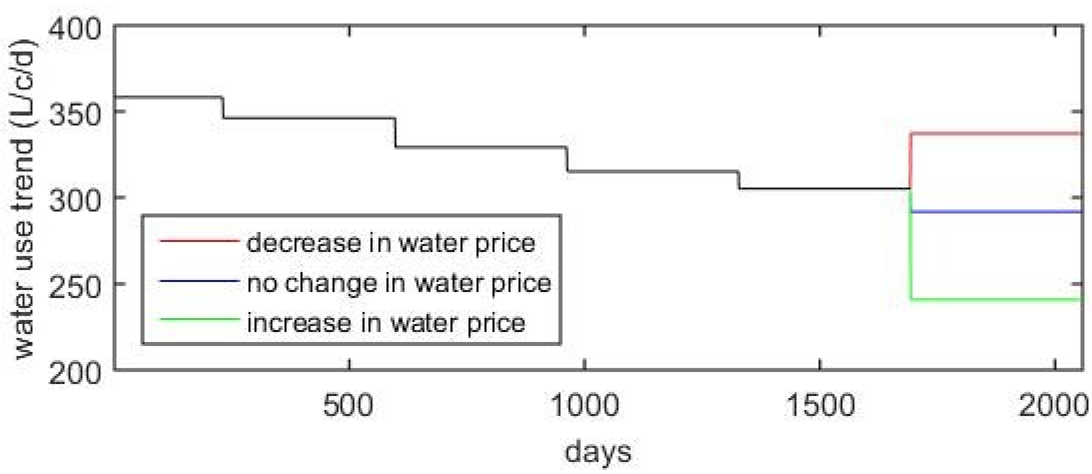

Figure 6.4. The trend of water use in City of Brossard estimated by Model C for years 2011 to 2015 (Black curve) and predicted for 2016 for different water pricing scenarios: (1) Normal increase in water price (green curve); (2) No change in water price (blue curve); (3) Reduction in water price (red curve).

The predictions of base water use in 2016 for different water pricing scenarios as described above were simple examples of employing the new model for future water demand management planning using water price adjustment. However, it should be noted that since water has no substitute for some basic uses, increasing the price will only be effective until a certain point. In other words, as discussed by other researchers as well, the relationship between water use and water price is inelastic (Abrams et al., 2012; Arbués et al., 2003). Therefore, water price adjustment must be consistent with given socioeconomic conditions.

7. Conclusions

This study reports a new approach to the prediction of urban water use. The new approach is successfully applied to analyze daily water use in the city of Brossard in the metropolitan area of Montreal, Canada. The following conclusions have been reached:

(1) The new approach considers changes in base water use on the annual timescale, as opposed to a monthly or daily timescale. The base water use is formulated as a function of socioeconomic factors. Such a formulation is conceptually different from existing formulations of base water use as a function of time. The formulation does not suffer the limitation of climate dependence and is consistent with typical conditions that socioeconomic factors change annually.

(2) The application to the city has quantitatively demonstrated the quality of the new approach. Time series of predicted daily water use captures very well observed fluctuating characteristics and long-term changes in water use. The results of the weighted coefficient of determination, the relative index of agreement and the root mean square error, are improved from the existing formulations.

(3) The new approach usefully separates daily water use into distinct components, including base water use, water use due to a seasonal cycle, water use due to weekend effect and climatic effects, and persistence component. Water use increases due to weekend effect. It decreases in the occurrence of rainfall, and the decrease is more sensitive to previous-day rainfall than current-day rainfall.

(4) The results of regression coefficients presented in this thesis are pertinent to urban water use in a specific city, but the same procedures can be applied to analyze urban water use in any other cities.

(5) According to the new model, base water use exhibits a downward trend possibly as a response to an increase in the annual charge for water use in the city of Brossard. Previously

developed models were not capable of showing what factor(s) have caused this trend in the water use of the city.

(6) The new approach offers a useful tool for decision makers to achieve sustainability in urban water management and water conservation by adjusting policies and price for water consumption.

However, like any other study, this study faces some limitations; most notably having a relatively short record of water use and socioeconomic variables. Since daily water use data was not available from more than five years ago, the analysis was limited to only five years. As the approaches taken in this study were statistical, having a longer time series could help make the analysis more accurate; especially for the base water use estimation of our model, since the data required for it is annually and five years might not always yield satisfactory results.

Possible future researches that this study motivates include: (1) assessing the performance of the models presented in this study for an area where numerous water restriction policies have been implemented or significant economic changes have been experienced throughout time; (2) using a disaggregate socioeconomic dataset (household by household) for the estimation of base use in order to investigate whether the performance of the model will improve; (3) in addition, it would be interesting to investigate how the model performs with a shorter billing cycle (i.e. monthly).

References

Abrams, B., Kumaradevan, S., Spaninks, F., and Sarafidis, V. (2012). An Econometric Assessment of Pricing Sydney's Residential Water Use. *THE ECONOMIC RECORD* **88**, 89-105.

Adamowski, J., Adamowski, K., and Prokoph, A. (2013). A Spectral Analysis Based Methodology to Detect Climatological Influences on Daily Urban Water Demand. *MATHEMATICAL GEOSCIENCES* **45**, 49-68.

Adamowski, J., Chan, H. F., Prasher, S. O., Ozga-Zielinski, B., and Sliusarieva, A. (2012). Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *WATER RESOURCES RESEARCH* **48**.

Alcamo, J., Flörke, M., and Märker, M. (2007). Future long-term changes in global water resources driven by socio-economic and climatic changes. *Hydrological Sciences Journal* **52**, 247–275.

Araghinejad, S. (2014). Data-Driven Modeling: Using MATLAB® in Water Resources and Environmental Engineering. *Water Science and Technology Library* **67**.

Arbués, F., Garc´ıa-Valiñas, M. ı. Á., and Mart´ınez-Espiñeira, R. (2003). Estimation of residential water demand: a state-of-the-art review. *Journal of Socio-Economics* **32**, 81–102.

BallingJr., R. C., and Gober, P. (2006). Climate Variability and Residential Water Use in the City of Phoenix, Arizona. *JOURNAL OF APPLIED METEOROLOGY AND CLIMATOLOGY* **46**, 1130–1137.

Ben Zaied, Y., and Binet, M. (2015). Modelling seasonality in residential water demand: the case of Tunisia. *APPLIED ECONOMICS* **47**, 1983-1996.

Bian, Y., Yan, S., and Xu, H. (2014). Efficiency evaluation for regional urban water use and wastewater decontamination systems in China: A DEA approach. *RESOURCES CONSERVATION AND RECYCLING* **83**, 15-23.

Billings, R. B., and Day, W. M. (1989). Demand management factors in residential water use: The southern Arizona experience. *J. Amer. Water Works Assoc.* **81**, 58-64.

Bougadis, J., Adamowski, K., and Diduch, R. (2005). Short-term municipal water demand forecasting. *HYDROLOGICAL PROCESSES* **19**, 137-148.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. (2008). Time Series Analysis: Forecasting and Control. *John Wiley, Hoboken, N. J.*

Brandes, O. M., and Ferguson, K. (2003). "Flushing the Future? Examining Urban Water Use in Canada," POLIS Project on Ecological Governance, University of Victoria, Victoria BC.

Cahill, R., and Lund, J. R. (2013). Residential Water Conservation in Australia and California. *JOURNAL OF WATER RESOURCES PLANNING AND MANAGEMENT-ASCE* **139**, 117-121.

City of Montreal (2016). Gestion financière - Table of tax rates, http://ville.montreal.qc.ca/portal/page?_pageid=44,137231621&_dad=portal&_schema=PORTAL, accessed on January 11, 2016.
Cleugh, H. A., Bui, E., Simon, D., Xu, J., and Mitchell, V. G. (2005). The Impact of Suburban Design on Water Use and Microclimate. *MODSIM 2005: INTERNATIONAL CONGRESS ON MODELLING AND SIMULATION: ADVANCES AND APPLICATIONS FOR MANAGEMENT AND DECISION MAKING: ADVANCES AND APPLICATIONS FOR MANAGEMENT AND DECISION MAKING*, 2019-2025.
Cole, M. A. (2004). Economic growth and water use. *Applied Economics Letters* **11**, 1-4.
De Mouche, L., Landfair, S., and Ward, F. A. (2011). Water Right Prices in the Rio Grande: Analysis and Policy Implications. *INTERNATIONAL JOURNAL OF WATER RESOURCES DEVELOPMENT* **27**, 314-327.
dos Santos, D. C., and Benetti, A. (2014). Application of the urban water use model for urban water use management purposes. *Water science and technology : a journal of the International Association on Water Pollution Research* **70**, 407-413.
Ferrara, I. (2008). Residential water use: A literature review. *In Household behaviourand environment: Reviewing the evidence. Paris: Organisation for Economic Co-operation and Development*.
Flörke, M., Kynast, E., Bärlund, I., Eisner, S., Wimmer, F., and Alcamo, J. (2013). Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development A global simulation study. *Global Environmental Change* **23**, 144–156.
Friedman, K., Heaney, J. P., Morales, M., and Palenchar, J. (2014). Analytical optimization of demand management strategies across all urban water use sectors. *WATER RESOURCES RESEARCH* **50**, 5475-5491.
Furumai, H. (2008). Rainwater and reclaimed wastewater for sustainable urban water use. *PHYSICS AND CHEMISTRY OF THE EARTH* **33**, 340-346.
Gato, S., Jayasuriya, N., and Hadgraft, R. (2003). A simple time series approach to modelling urban water demand. *In: 28th International Hydrology and Water Resources Symposium, Wollongong, NSW, 10–14 November 2003*.
Gato, S., Jayasuriya, N., and Roberts, P. (2007). Temperature and rainfall thresholds for base use urban water demand modelling. *JOURNAL OF HYDROLOGY* **337**, 364-376.
Gaudin, S. (2006). Effect of price information on residential water demand. *Applied Economics* **38**, 383-393.
Guhathakurta, S., and Gober, P. (2007). The Impact of the Phoenix Urban Heat Island on Residential Water Use. *JOURNAL OF THE AMERICAN PLANNING ASSOCIATION* **73**, 317-329.
Gutzler, D., and Nims, J. (2005). Interannual variability of water demand and summer climate in Albuquerque, New Mexico. *JOURNAL OF APPLIED METEOROLOGY* **44**, 1777-1787.
House-Peters, L., Pratt, B., and Chang, H. (2010). Effects of Urban Spatial Structure, Sociodemographics, and Climate on Residential Water Consumption in Hillsboro,

Oregon. *JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION* **46**, 461-472.

Hughes, G., Chinowsky, P., and Strzepek, K. (2010). The costs of adaptationto climate change for water infrastructure in OECD countries. *Utilities Policy* **18**, 142-153.

Hurd, B., St Hilaire, R., and White, J. (2006). Residential landscapes, homeowner attitudes, and water-wise choices in New Mexico. *HORTTECHNOLOGY* **16**, 241-246.

Hyndman, R. J. (2014). Forecasting: Principles & Practice. *University of Western Australia*, pp 52.

Hyndman, R. J., and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting* **22**, 679-688.

Kenney, D. S., Goemans, C., Klein, R., Lowrey, J., and Reidy, K. (2008). Residential water demand management: Lessons from Aurora, Colorado. *JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION* **44**, 192-207

Krause, P., Boyle, D. P., and Base, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences* **5**, 89-97.

Legates, D. R., and McCabe Jr., G. J. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *WATER RESOURCES RESEARCH* **35**, 233–241.

Maidment, D. R., and Miaou, S. P. (1986). Daily water use in nine cities. *Water Resources Research* **22**, 845–851.

Maidment, D. R., Miaou, S. P., and Crawford, M. M. (1985). Transfer function models of daily urban water use. *Water Resources Research* **21**, 425-432.

Maidment, D. R., and Parzen, E. (1984a). Cascade model of monthly municipal water use. *Water Resources Research* **20**, 15-23.

Maidment, D. R., and Parzen, E. (1984b). Time patterns of water use in six Texas cities. *Journal of Water Resources Planning and Management* **110**, 90-106.

Martinez-Espineira, R. (2002). ResidentialWater Demand in the Northwest of Spain. *ENVIRONMENTAL & RESOURCE ECONOMICS* **21**, 161-187.

Medellin-Azuara, J., Lund, J. R., and Howitt, R. E. (2007). Water supply analysis for restoring the Colorado River Delta, Mexico. *JOURNAL OF WATER RESOURCES PLANNING AND MANAGEMENT-ASCE* **133**, 462-471.

Mini, C., Hogue, T. S., and Pincetl, S. (2014). Patterns and controlling factors of residential water use in Los Angeles, California. *Water Policy* **16**, 1054-1069.

Nakayama, T. (2011). Simulation of the effect of irrigation on the hydrologic cycle in the highly cultivated Yellow River Basin. *AGRICULTURAL AND FOREST METEOROLOGY* **151**, 314-327.

Nazemi, A., and Wheater, H. S. (2015). On inclusion of water resource management in Earth system models – Part 1: Problem definition and representation of water demand. *Hydrology and Earth System Sciences* **19**, 33-61.

Panagopoulos, G. P. (2014). Assessing the impacts of socio-economic and hydrological factors on urban water demand: A multivariate statistical approach. *Journal of Hydrology* **518**, 42-47.

Polebitski, A. S., and Palmer, R. N. (2010). Seasonal Residential Water Demand Forecasting for Census Tracts. *JOURNAL OF WATER RESOURCES PLANNING AND MANAGEMENT-ASCE* **136**, 27-36.

Rai, S. C. (2011). Water Management for a Megacity: National Capital Territory of Delhi. *WATER RESOURCES MANAGEMENT* **25**, 2267-2278.

Rathnayaka, K., Maheepala, S., Nawarathna, B., George, B., Malano, H., Arora, M., and Roberts, P. (2014). Factors affecting the variability of household water use in Melbourne, Australia. *RESOURCES CONSERVATION AND RECYCLING* **92**, 85-94.

Rathnayaka, K., Malano, H., Maheepala, S., George, B., Nawarathna, B., Arora, M., and Roberts, P. (2015). Seasonal Demand Dynamics of Residential Water End-Uses. *WATER* **7**, 202-216.

Romano, G., Salvati, N., and Guerrini, A. (2014). Estimating the Determinants of Residential Water Demand in Italy. *WATER* **6**, 2929-2945.

Salas-La Cruz, J. D., and Yevjevich, V. (1972). Stochastic Structure of Water Use Time Series. *Hydrology Paper No. 52, Colorado State University*, Fort Collins, Colo., 1972, 71 p.

Salvador, R., Bautista-Capetillo, C., and Playan, E. (2011). Irrigation performance in private urban landscapes: A study case in Zaragoza (Spain). *LANDSCAPE AND URBAN PLANNING* **100**, 302-311.

Schleich, J., and Hillenbrand, T. (2009). Determinants of residential water demand in Germany. *ECOLOGICAL ECONOMICS* **68**, 1756-1769.

Siauw, T., and Bayen, A. M. (2015). "An Introduction to MATLAB® Programming and Numerical Methods for Engineers," Academic Press, pp. 202-204.

Sohn, J. (2011). Watering cities: spatial analysis of urban water use in the Southeastern United States. *JOURNAL OF ENVIRONMENTAL PLANNING AND MANAGEMENT* **54**, 1351-1371.

Stoker, P., and Rothfeder, R. (2014). Drivers of urban water use. *Sustainable Cities and Society* **12**, 1-8.

Syme, G., Shao, Q., Po, M., and Campbell, E. (2004). Predicting and understanding home garden water use. *LANDSCAPE AND URBAN PLANNING* **68**, 121-128.

Turner, V. K., and Ibes, D. C. (2011). The Impact of Homeowners Associations on Residential Water Demand Management in Phoenix, Arizona. *URBAN GEOGRAPHY* **32**, 1167-1188.

Van Rooijen, D. J., Turral, H., and Biggs, T. W. (2009). Urban and industrial water use in the Krishna Basin, India. *IRRIGATION AND DRAINAGE* **58**, 406-428.

Wada, Y., van Beek, L. P. H., Viviroli, D., Dürr, H. H., Weingartner, R., and Bierkens, M. F. P. (2011). Global monthly water stress: 2. Water demand and severity of water stress. *Water Resources Research* **47**.

Willmott, C. J. (1981). On the validation of models. *Physical Geography,* **2**, 184-194.

Willmott, C. J. (1984). On the evaluation of model performance in physical geography. *in: Spatial Statistics and Models*, 443–460.

Wong, J. S., Zhang, Q., and Chen, Y. D. (2010). Statistical modeling of daily urban water consumption in Hong Kong: Trend, changing patterns, and forecast. *WATER RESOURCES RESEARCH* **46**.

Woodard, G. C., and Horn, C. (1988). Effects of weather and climate on municipal water demand in Arizona. *Division of Economic and Business Research, College of Business and Public Administration, University of Arizona*.

Wu, P., and Tan, M. (2012). Challenges for sustainable urbanization: a case study of water shortage and water environment changes in Shandong, China. *Procedia Environmental Sciences* **13**.

Zhang, Z., and Shao, Y. (2010). Inequality and polarization analysis of urban water use in the Yangtze River Delta area, China. *WATER SCIENCE AND TECHNOLOGY* **62**, 300-310.

Zhou, S., McMahon, T., Walton, A., and Lewis, J. (2000). Forecasting daily urban water demand: a case study of Melbourne. *JOURNAL OF HYDROLOGY* **236**, 153-164.