1. INTRODUCTION

2. WATER DEMAND FORECASTING ISSUES

2.1 Forecasting Models

2.2 Time Scales

2.3 Water Demand Patterns

3. EXPLORATORY ANALYSIS OF WATER CONSUMPTION DATA

3.1 Weather in Abbotsford

Abbotsford is a city located in the Lower Mainland region of British Columbia, Canada. Abbotsford is proximate to the Pacific Ocean, which provides milder winters, along with much greater rainfall than other regions at the same latitude. Abbotsford’s winter and spring are cool, but relatively mild compared to most of Canada. The temperature starts to drop at the end of October. The rainy season starts at the end of October, and lasts lasts to April of the second year. November and December have the heaviest precipitation. Summer in Abbotsford is relatively dry, and warm. The warmest month is August, which average daily high is 23.8°C, and sometimes as high as over 30°C. August averages only 20% of November’s rainfall, and only about 16% of annual precipitation falls. In Summer, it is common to receive little or no rainfall for weeks at a time.

3.2 Water Use Distribution in Abbotsford

The water consumption data is from Abbotsford city, British Columbia, Canada from September 1, 2012 to August 31, 2013. Daily and hourly time-series data were collected from 25,294 customers, distributed in six customer groups (see Figure 4). There are 2013 215,172,496 data points in total for the hourly time-series data, and 9,232,310 data points for the daily time-series data (roughly 10GB in total). The residential customers are discriminated according to the house types, which are single family residential (SFRES) and multifamily residential (MFRES). As shown, the single family residential has the biggest share (79.61%).

The climate data at the same period were downloaded from Environment Canada (https://weather.gc.ca). The data contains the weather temperature at hourly resolution, and the rainfall at daily resolution.

3.3 Residential Water Use

Approximately two-thirds of the water used in Abbotsford is for residential purposes, 42.68% by the residents of single-family homes. In a spatially weighted regression analysis of single-family residential water demand at the census tract level, Wentz and Gober (2007) found that four variables—average household size, the percent of homes with a swimming pool, average lot size, and average percent of lots covered with mesic (turf) vegetation—explained more than 80% of the spatial variation in metered water use. Larger household size increases indoor water use for such purposes as toilet flushing, showers, laundry, and dishwashing, although Arbués et al. (2003) note the tendency for less-than-proportional increases in use because of economies of scale in water use. Swimming pools, lot size, and vegetation type account for outdoor use from pool evaporation and garden irrigation. We anticipate that residential water use is climate sensitive because of the heavy reliance on outdoor uses in Abbotsford’s water portfolio.

We show the daily time-series of the residential water usage in Figure ??, we could observer that the water usage shows periodicity in the days of a week. From Monday–Friday, the consumptions are lower than the weekend, and the holidays, propably for the reason that in the weekend and holiday people stays more time at home, thus use more water. In addition, we could observer that in summer time (May - October) the consumption are higher than winter time (November–April) which is probably due to the temperature effect. For example, households grow the flowers in outdoors in summer, which need to water the flowers, wash the care, or fill the swimming pools. Another climate effect we might consider is the effect of rainfall. For example, due to the rainfall, the water consumption might be reduced, e.g., people can save the water for watering the flowers in outdoor.

4. MODEL FOR RESIDENTIAL WATER DEMAND FORECASTING

4.1 Basis for Water Demand Forecasting

For the residential water demand, many variables are considered influential and relevant in determining the water usage. we refer to a report prepared for the Water Research Foundation by Coomes
Table 1: Sizes of customer types and the water consumption

<table>
<thead>
<tr>
<th>Customer Types</th>
<th>Size</th>
<th>Percentage, %</th>
<th>Water Consumption</th>
<th>Percentage, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFRES</td>
<td>20,136</td>
<td>79.61</td>
<td>5,158,189.4</td>
<td>42.68</td>
</tr>
<tr>
<td>AGRI</td>
<td>2,612</td>
<td>10.33</td>
<td>2,411,693.3</td>
<td>19.95</td>
</tr>
<tr>
<td>MFRES</td>
<td>1,215</td>
<td>4.80</td>
<td>2,148,880.4</td>
<td>17.78</td>
</tr>
<tr>
<td>COM</td>
<td>561</td>
<td>2.22</td>
<td>758,081.3</td>
<td>6.27</td>
</tr>
<tr>
<td>IND</td>
<td>438</td>
<td>1.73</td>
<td>1,099,533.5</td>
<td>9.10</td>
</tr>
<tr>
<td>INST</td>
<td>332</td>
<td>1.31</td>
<td>509,298.9</td>
<td>4.21</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>25,294</td>
<td>100</td>
<td>12,085,676.9</td>
<td>100</td>
</tr>
</tbody>
</table>
Figure 4: Daily pattern of hourly water demand in Summer

Figure 5: Daily pattern of hourly water demand in Winter

Figure 6: The effect of weather temperature and rainfall on water consumption

Figure 7: Hourly baseload in Summer

Figure 8: Hourly baseload in Winter
et al. (2010), in which the authors tested the effect of 26 variables on average daily water use for 293 residential customers of the Louisville Water Company.

The variables range from socio-economic to various derivatives of weather-related variables. Examples of these weather-related variables and how they are used can be found in Coomes et al. (2010) and Brekke et al. (2002). The availability and choice of these independent variables can also influence the forecasting models used. For instance, whereas population projections and per capita demand are the drivers for unit rate models, these have no consideration when exponential smoothing or Box-Jenkins models are formulated.

We model the water consumption, \( W_i \), into base load and seasonal load, which is as follows:

\[
D_i = B_i + S_i
\]  

(1)

### 4.2 Base Load

The base load represents the weather insensitive portion of the total water use. The base load mainly are due to the use of daily use of living, such as drinking, toilet, bath and washing. The base use are unlikely to change abruptly, but might change slowly, e.g., placing more water efficient appliance such as new washing machine. Therefore, the base use is assumed to be a fixed percentiled consumption of hourly flows in household for weekdays and weekends/holidays. We use 10 percentiled water usage as the base load, also adopted by [1]. We derive the base load as follows:

\[
\begin{align*}
B_{d} = & \phi_s,iS_{d-i} + \beta_s h(T_d) + \psi_s g(R_d) + \eta_s f(t) + \epsilon_s, \\
& d \in S
\end{align*}
\]

(2)

where \( s \) is the season index; \( \phi_s,i, \beta_s, \psi_s, \eta_s \) are the model parameters, and \( \epsilon_s \) is the white noise; \( S_d \) represents the water consumption at a particular day \( d \) in a week, and \( d = 0, \ldots, 6 \) representing Sunday, Monday, ..., Saturday; \( p \) is the order of auto-regression; \( s \) is the season index; \( h(T_d) \), \( g(R_d) \), and \( f(t) \) are the functions of weather and rainfall effects. Most of the current studies model the weather variables to the water assumption linearly, such as [1]. However, Maidment and Miaou criticized the common linear approach, and suggest that the weather variable, rainfall, has a dynamic effect, which reduces water demand initially, but the effect diminishes over time. Inspired by their suggestion, we set a threshold value for each of the weather variables, which limit the effect of each weather variable.

\[
\begin{align*}
&h(T_d) = \begin{cases} 
T_d - \tau & T_d > \tau \\
0 & \text{Otherwise}
\end{cases} \\
&g(R_t) = \begin{cases} \gamma & R_t > \gamma \\
R_t & \text{Otherwise}
\end{cases} \\
&f(t) = \begin{cases} \zeta - \epsilon & \zeta < \epsilon \\
0 & \text{Otherwise}
\end{cases}
\end{align*}
\]

where \( T_d \) is the maximum temperature at the week day \( d \), and \( \tau \) the threshold of temperature effect. We adopt the 20 cesus degree as the threshold value.

### 4.3 Daily Seasonal Load

The daily seasonal load represents the weather sensitive portion of the total use of a day, i.e., the daily load subtracted by the base load, denoted as \( S_i = D_i - B_i \). We use the periodic Auto-Regression with eXogenous variables (PARX) [5] to model the daily seasonal load. The period is weekly, and the seasons are the week days. Since the social and economic metric are available in our data, such as the price of water, and size of household, we therefore consider the climat impact on the residential water use. There are several environment variables including weather temperature, precipitation and the number of days which precipitation exceed 0.2mm. We take these environmental factors as the exogenous variable in the PARX model. Thus, the model for daily seasonal load can be expressed as:

\[
S_d = \sum_{i=1}^{p} \phi_{s,i} S_{d-i} + \beta_s h(T_d) + \psi_s g(R_d) + \eta_s f(t) + \epsilon_s, \quad d \in S
\]

where \( s \) is the season index; \( \phi_{s,i}, \beta_s, \psi_s, \eta_s \) are the model parameters, and \( \epsilon_s \) is the white noise; \( S_d \) represents the water consumption at a particular day \( d \) in a week, and \( d = 0, \ldots, 6 \) representing Sunday, Monday, ..., Saturday; \( p \) is the order of auto-regression; \( s \) is the season index; \( h(T_d) \), \( g(R_d) \), and \( f(t) \) are the functions of weather and rainfall effects. Most of the current studies model the weather variables to the water assumption linearly, such as [1]. However, Maidment and Miaou criticized the common linear approach, and suggest that the weather variable, rainfall, has a dynamic effect, which reduces water demand initially, but the effect diminishes over time. Inspired by their suggestion, we set a threshold value for each of the weather variables, which limit the effect of each weather variable.

\[
\begin{align*}
&h(T_d) = \begin{cases} 
T_d - \tau & T_d > \tau \\
0 & \text{Otherwise}
\end{cases} \\
&g(R_t) = \begin{cases} \gamma & R_t > \gamma \\
R_t & \text{Otherwise}
\end{cases} \\
&f(t) = \begin{cases} \zeta - \epsilon & \zeta < \epsilon \\
0 & \text{Otherwise}
\end{cases}
\end{align*}
\]

where \( T_d \) is the maximum temperature at the week day \( d \), and \( \tau \) the threshold of temperature effect. We adopt the 20 cesus degree as the threshold value.

### 4.4 Hourly Seasonal Load

– No hourly rainfall data

### 5. RESULTS

#### 5.1 The evaluation statistics

\[
RSMSE = \frac{1}{n} \left[ \frac{1}{b} \sum_{i=1}^{n} \epsilon_i^2 \right]^{\frac{1}{2}}
\]

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} | \epsilon_i |
\]

#### 5.2 Compare with the linear model

We now use daily single family residential (SFRES), and weather data to compute the coefficients. We set the temperature threshold value \( \tau \) as 20 celsius degree, and the threshold value of rainfall \( \gamma \) as 1 mm precipitation.

### 6. REFERENCES


