Identifying relationships between weather variables and domestic water consumption using smart metering

Maria Xenochristou[1], Zoran Kapelan[2], Chris Hutton[3], Jan Hofman[4]

[1, 2] Centre for Water Systems, University of Exeter, North Park Road, EX4 4QF Exeter, UK
[3] Wessex Water, Claverton Down Road, BA2 7WW Bath, UK
[4] Water Innovation and Research Centre, University of Bath, BA2 7AY Bath Avon, UK
[1]mx220@exeter.ac.uk

ABSTRACT

Satisfying the water supply-demand balance is a major problem of modern societies due to water scarcity, which is expected to be amplified by changes in the climate. In order to fulfil future demands, accurate predictions of water consumption are essential. This paper investigates the relationship between water consumption and air temperature, using a combination of smart demand metering data, household characteristics, and socio-economic data. Results indicate that the correlation between water consumption and temperature increases during working days, evenings, as well as during the summer and spring. In addition, a positive correlation was identified for households that are metered, have bigger gardens, and medium occupancy, as well as residents with higher socio-economic status and high seasonal variations in water consumption.

Keywords: water demand, smart meters, weather

1 BACKGROUND

Water is essential for the preservation of the urban and natural environment and competition over the use and control of water resources is often a source of conflict between countries and individuals. The Government's water strategy for England report, which plans strategies for securing the future of water resources and improving water environment till 2030, highlights that an essential aspect of managing water demand is by ensuring a good forecasting of future patterns [3]. However, this is a very challenging task, as water consumption is highly uncertain, due to the influence of socio-demographic, economic, and environmental factors that vary both spatially and temporally, as well as the lack of data, since more than half of the properties in the UK are not metered [4, 10].

It is a common practice among researchers to regard water demand as the sum of base and seasonal demand [1, 5, 10]. The former is the amount of water that is necessary to cover basic human needs, whereas the latter is weather dependent and consists mainly of water used for outdoor activities, such as gardening or filling water pools, as well as weather induced showering and drinking. Since seasonal demand is weather-influenced and therefore relatively unpredictable, it is what poses a risk for municipalities and network operators, as it challenges their ability to meet peak demands [2], as well as maintain the long-term supply-demand balance under the threat of climate change.

Many studies have investigated the influence of weather and specifically air temperature on water use. The aim of the majority of these studies has been to establish the suspected relationship between water demand and a series of weather variables, but in most cases without accounting for a variety of factors that contribute to or eliminate this relationship. Water demand analysis at group level can provide detailed information about weather induced water use and enable the planning and adaptation of climate change strategies at local scale [2]. Few studies [1, 2, 5, 6, 7, 8, 9, 11, 12, 13] have attempted to explain the sensitivity of seasonal water demand to changes in temperature, based on household and socio-economic characteristics.

The interaction between water use and climate variables varies significantly among census tracts [8, 11, 12]. Communities with large gardens [7, 8, 11, 12], swimming pools [5, 7, 11, 12], sophisticated lawn reticulation systems [7], high proportion of irrigated landscape [10], as well as affluent [7, 8, 11, 12] and well-educated [8] residents with owner occupied [7], newer and larger homes [8], and high property values [8] showed a greater sensitivity to climate with regards to water consumption. On the other hand in one study, large, Hispanic families [11] showed little to no sensitivity to weather changes.

Although the aforementioned studies provide a valuable insight into the sensitivity of seasonal demand to temperature for different demographic groups, limitations exist. Due to data protection and lack of available technologies in the past, data availability is very limited and water demand is most of the times available as an aggregate of groups of properties with non-uniform characteristics, often representing a range of socio-economic backgrounds. In addition, where available, consumption data is collected by conventional water meters, which fail to represent the high seasonal and regional heterogeneity of residential water demand. Lastly, the vast majority of past studies have been conducted in water scarce cities of the US and Australia, where household, as well as socio-economic and climate characteristics are very different to the UK, which also experiences its own weather-dependent demand variation.

This research aims to fill the gap in the current literature by utilising an extensive dataset of high temporal and spatial resolution smart demand metering data from the Southwest of England. Spatially, data is available at the property level which enables us to account for the individual characteristics that may influence residents' water use behaviours. In addition, the temporal scale (15-30 min intervals) allows to account for the different water use patterns that occur between weekdays and weekends/holidays or even between different times of the day (e.g. afternoon, evening).

These results can be used in order to improve water demand forecasting, accounting not only for potential changes in future temperatures but also demographic characteristics. They can also be used by network managers and water companies in order to develop targeted water conservation policies and form water pricing structures.

2 Data

The current research utilises a unique dataset of water demand data, along with household and socio-economic characteristics in order to identify relationships between urban water demand and air temperature.

Data used in this study comprises:

- Smart demand metering data collected at 15-30 minute intervals for the period of 10/2014 - 08/2016 from almost 2,000 properties, from the Southwest of England (Dorset, Somerset, Wiltshire, and Hampshire);
- Household characteristics (garden size, rateable value, metering status, occupancy rate);
- Socio-economic data (Acorn groups and types); Acorn is a geodemographic segmentation of the UK's population based on social factors and population behaviour and it is used to

provide an understanding of the different types of people. It segments households into 6 categories, 18 groups, and 62 types [14].

- Maximum daily air temperature for the same time-period (10/2014 - 08/2016) for various weather stations across the Southwest, acquired from the Met office (UK) and NOAA (National Oceanic and Atmospheric Administration – U.S. Department of Commerce).

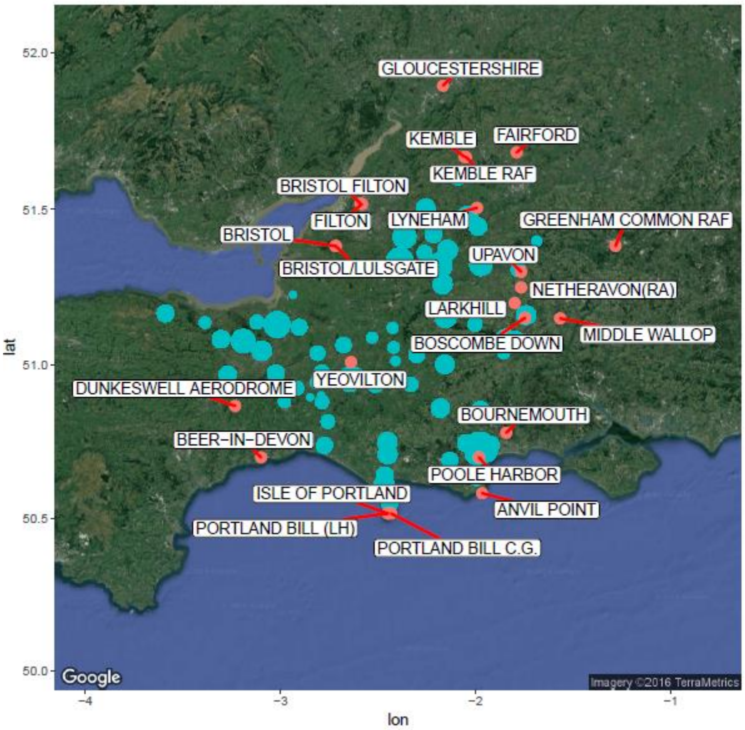

Figure 1. Map of properties and weather stations. Blue dots represent postcode central locations for the properties and size represents the amount of properties in each postcode. Red dots represent weather stations.

3 Methodology

As previous studies have pointed out, water demand patterns are best explored and understood through a theoretical framework of coupled human and natural systems [2, 9]. In order to take full advantage of the available dataset, this study follows a systematic methodology to test the correlation between maximum air temperature and water demand for different segmentations (i.e. groupings) of properties based on household, socio-economic, and temporal characteristics.

The first step was to process and clean the raw dataset of consumption records. As the data included a mix of consumption signals at 15 minute and 30 minute intervals, in order to homogenise the dataset and avoid mistakes, only hourly recordings of consumption were included. In order to ensure the credibility of the data, the following was also excluded from the final dataset:

- Consumption recordings that remain unchanged for more than 24hours. This could be either due to a faulty meter or an empty household.
- Recordings that include leakages. For every property and calendar month in the data, if less than 20% of the total recordings are equal to zero, it is assumed that there is a leakage in the property and the recordings for the corresponding month and household are excluded from the data. This method was tested and found to be effective in excluding leaking properties.

- Recordings that correspond to consumption higher than 400 litres/15 minutes. This is mostly due to a fault in the transmission of the data, where the meter is recording but not transmitting, leading to a high jump in consumption when the fault is restored.
- Months that do not include at least 36 hours of recordings. In these cases, the signal is not considered trustworthy and is therefore excluded from the dataset.
- Properties with no gardens, as their very small number was not considered significant for the analysis.

After the pre-processing of the data, 1,592 properties were included in the final dataset with recordings corresponding to a total duration of 660 days. The water demand was aggregated at a daily scale and a mean consumption was calculated for all the properties in the data. Recordings of maximum daily temperature from the Yeovilton weather station, which is located in the centre of the area of interest were used in the analysis. Due to the high correlation of daily temperature records among the weather stations, with an average R^2 value of 95%, this was considered acceptable.

Following the pre-processing of the data, a methodology was developed that systematically evaluates multiple temporal and household aggregations of properties for possible relationships between water consumption and air temperature. The variables that were included in the segmentation, along with the categories considered for each one appear in Table 1 and Table 2.

Table 1: Temporal segmentation of analysed consumption data

TEMPORAL SEGMENTATION		
Weekends & Holidays	Season	Time of Day
All TRUE (Weekends + Bank Holidays) FALSE (Working Days)	All Summer Spring Autumn Winter	All Morning (06.00 - 12.00) Afternoon (12.00 - 18.00) Evening (18.00 - 24.00) Night (24.00 - 06.00)

Table 2: Household segmentation of analysed consumption data

HOUSEHOLD CHARACTERISTICS SEGMENTATION		
Garden_Sizes	ACORN Groups	Occupancy
All Large (> 165 m^2) Medium (61-165 m^2) Small (< 60 m^2)	All Affluent (A, B, C, D, E) Comfortable (F, G, H, I, J) Financially stretched (K, L, M, N, O, P, Q)	All High (> 3 occupants) Medium (2-3 occupants) Low (< 2 occupants)
Status	Seasonal Consumers	Rateable value
All Measured (knowingly on a meter) Unmeasured (not knowingly on a meter)	All High (Mean Seasonal Summer Consumption > 100 l/day)	All High (> 190) Medium (135 - 190) Low (< 135)

Garden sizes are divided into large, medium, and small by the water company, based on size in square meters (Table 2). ACORN groups are divided based on the ACORN user guide [14]. According to this, consumer groups A, B, and C are classified as “Affluent Achievers”, and groups D and E as “Rising Prosperity”. All groups A to E are classified as “Affluent” in the following.

Groups F to J are classified as "Comfortable Communities" in the same guide, whereas groups K to Q are "Financially Stretched".

The threshold values for the "Rateable value" were chosen in order to result in relatively equal number of properties among the groups. The threshold of the 100 litres/day for the high seasonal consumers was chosen in order for this category to include enough properties to create segmentation categories (approximately 550 properties) but at the same time few enough to be able to distinguish this group from the rest of the properties.

For each segmentation category, a second degree polynomial regression model was fitted on the data. The number of properties and days each category relates to were calculated, along with the R^2 correlation coefficient and p-value for the model. The results were considered significant at 95% confidence level. Segmentations that included less than 50 properties or 40 days, or had a p-value greater than 5%, were excluded from the analysis.

4 Results and discussion

The use of the above segmentation methodology, including 9 variables, resulted in 115,200 different groupings, a sample of which is shown in Table 3, along with the correlation coefficient and p-value obtained. Although this is a comprehensive analysis, it might prove difficult to derive conclusions from due to its size and complexity. Table 3 shows that the correlation between water consumption and maximum daily temperature is not a strong one in the UK, at least on the analysed data set.

Table 3: Example of the aggregation table and results obtained. Nine variables were considered, three for temporal segmentation (yellow) and six for household segmentation (green). For each category, the number of properties and days included in it, as well as the correlation coefficient and p value of a second degree polynomial regression model were calculated (grey).

Weekends	Time of Day	Season	Garden Sizes	ACORN	Occupancy	Rateable	Status	Seasonal Consumers	Properties	Days	R^2	p value
FALSE	all	spring	all	all	all	all	All	high	582	119	0.52	0.00
FALSE	all	summer	all	all	medium	all	All	all	796	104	0.51	0.00
FALSE	evening	summer	all	all	medium	all	All	all	796	107	0.49	0.00
FALSE	all	summer	Medium	all	all	all	All	all	431	104	0.48	0.00
FALSE	all	summer	all	Comfortable	medium	all	All	all	298	107	0.47	0.00
FALSE	evening	summer	Medium	all	all	all	All	all	431	104	0.46	0.00
FALSE	evening	summer	all	all	all	high	All	all	292	104	0.45	0.00

Figure 2 shows six example relationships between maximum daily temperature and average consumption for six different aggregations of properties. Each point in the figure corresponds to one day for which data was available. The horizontal axis shows the maximum daily temperature (°C) that was recorded within that day, while the vertical axis shows the average daily consumption (averaged across all the properties of the corresponding segmentation) for the same day. The red line represents the regression model that was fitted on the data.

In order to determine the influence of each individual variable and considering the number of different aggregations, a summary table was created (Table 4). Excluding segmentation categories with less than a certain number of days and properties favours these categories that contain a larger amount of either properties or days. In order to overcome this bias in the analysis, an indicator value was calculated for each category of each variable as the ratio between the total amount of segmentation categories that include the corresponding variable to the amount of segmentation categories that have an R^2 value of 25% or higher. The higher the indicator value, the stronger the correlation between air temperature and water consumption.

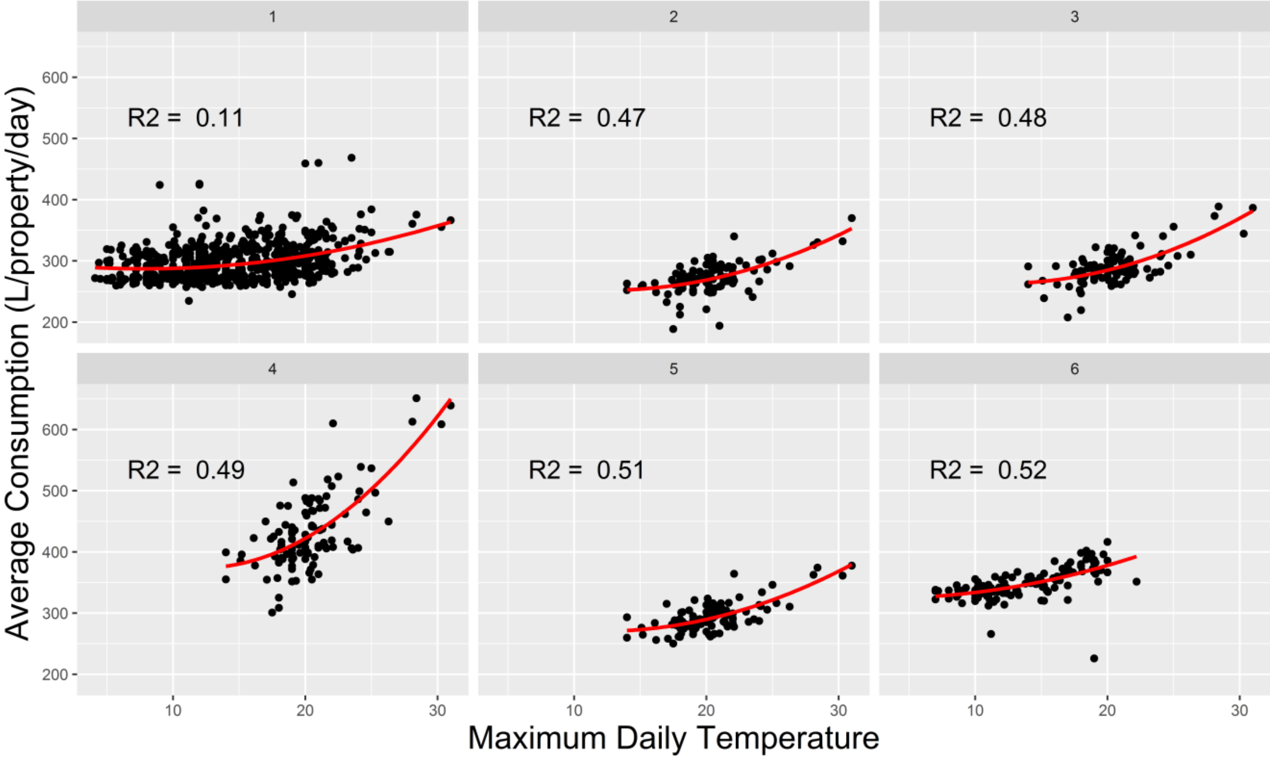

Figure 2: Correlation between maximum daily temperature (°C) and average daily consumption (averaged across all the properties) for (1) all properties, during (2) summer working days and comfortable communities with medium occupancy, (3) summer working days and properties with medium garden sizes, (4) summer working days, evenings, and properties with medium occupancy, (5) summer working days and properties with medium occupancy, and lastly (6) spring working days and residents with high seasonal variations in consumption.

As it can be seen from Table 4, there is some degree of correlation between outdoor temperature and water consumption during working days, evenings, as well as spring and summer. This is likely due to the fact that increased water use in the UK mainly relates to outdoor activities (such as gardening), which is taking place more in spring and summer. Note also from the table that an increase in the garden size as well as household income leads to an increase in seasonal water use, for obvious reasons. In addition to this, increased water use for bathing and showering is more likely to take place in the summer months, when people are at home (mornings and evenings of working days).

Table 4 also shows that water consumption during weekends and holidays, nights and afternoons, in autumn and winter does not correlate well with the maximum air temperature. This is due to the fact that during the weekends and holidays people tend to have less regular schedules and/or are frequently away from home. Also, outdoor temperature is unlikely to have a significant effect on peoples' behaviours towards water use during winter and autumn, when temperatures are generally lower, precipitation higher, sunshine hours limited, and gardening is less likely to occur. However, during spring and summer, an increase in temperature will lead to an increase in evapotranspiration for both humans and plants, which will increase the demand for water.

As shown in Table 4, properties that are metered appear to show a stronger correlation between water consumption and temperature. Since these properties are being charged based on the amount of water they consume, they are likely to use water more sensibly, i.e. when it is necessary due to high temperatures. The rateable value of a property seems to have the least amount of impact compared to all other variables considered in this study and its effect seems unclear. People with a high variation between winter and summer water use also seem to be influenced by changes in temperature, which is to be expected, as this seasonal change is very likely to be influenced by changes in temperature.

Table 4: Ratio (green) between total number of segmentations included in each category (blue) and number of segmentations with an R^2 correlation coefficient higher than 0.25 at 95% confidence interval (yellow).

Weekends				Time of day				Season			
FALSE	4910	196	**4.0%**	Morning	2859	40	**1.4%**	Summer	2625	97	**3.7%**
TRUE	4287	30	**0.7%**	Evening	2850	77	**2.7%**	Spring	2985	92	**3.1%**
				Night	2725	0	**0.0%**	Autumn	2477	12	**0.5%**
				Afternoon	2817	3	**0.1%**	Winter	2960	0	**0.0%**
Garden Size				ACORN				Occupancy			
Large	1894	36	**1.9%**	Affluent	2524	54	**2.1%**	High	1064	6	**0.6%**
Medium	2379	42	**1.8%**	Comfortable	2784	40	**1.4%**	Medium	4296	102	**2.4%**
Small	1961	0	**0.0%**	Financially stretched	1461	2	**0.1%**	Low	660	9	**1.4%**
Rateable value				Status				Summer Consumers			
High	1977	20	**1.0%**	Measured	3613	96	**2.7%**	All	10559	125	**1.2%**
Medium	2278	11	**0.5%**	Unmeasured	3471	54	**1.6%**	High	3548	145	**4.1%**
Low	1021	15	**1.5%**								

Lastly, properties with medium occupancy also show a higher correlation. This could be due to the fact that when there is just one or more than 3 people in a household, it is more likely that their schedules vary a lot, increasing the inherent randomness of water use, while if there is only 2 to 3 people, it's more likely that they follow certain patterns.

The UK has a mild climate with cool summers and rain well distributed over the year. Therefore, although there is an evident relationship between temperature and water demand, strong correlations were not identified in this study. However, it becomes apparent that different demographic groups have different habits with regards to water use. Having said this, the same demographic group could react differently to changes in temperature based on the time of year, the time of the week, or even the time of day.

5 Conclusions

Water supply-demand balance is an issue of increasing concern, especially under the threat of changes in the climate. Accurate forecasts of demand should account not only for changes in future temperatures but also changes in demographics, household and socio-economic factors. The current study aims to address this problem by researching the variation in the correlation between temperature and water consumption, based on socio-economic, household, and temporal characteristics.

Results showed that during working days, evenings, as well as spring and summer, water use relates stronger to air temperature. Properties with medium occupancy, larger garden sizes and socio-economic status, as well as properties that are metered and have high seasonal variations in water use also showed an increase in correlation.

Future work will expand the above analysis to investigate potential differences when looking into different temporal (e.g. monthly) or spatial (e.g. at varying distances from the weather station) scales. Also, the effect of additional weather variables, such as precipitation and humidity will be explored. Ultimately, new, improved demand forecasting models will be developed based on findings obtained.

Acknowledgements

This study was conducted as part of the WISE Centre for Doctoral Training, funded by the UK Engineering and Physical Sciences Research Council.

References

[1] B. Breyer, H. Chang and G. H. Parandvash, “Land-use, temperature, and single-family residential water use patterns in Portland, Oregon and Phoenix, Arizona,” J. Applied Geography. vol. 35, pp. 142-151, 2012
[2] B. Breyer and H. Chang, “Urban water consumption and weather variation in the Portland, Oregon Metropolitan area,” J. Urban Climate. vol. 9, pp. 1-18, 2014
[3] Defra (Department for Environment, Food and Rural Affairs). “Future Water, The Government’s water strategy for England”, 2008
[4] C. J. Hutton and Z. Kapelan, “A probabilistic methodology for quantifying, diagnosing and reducing model structural and predictive errors in short term water demand forecasting,” J. Environmental modelling and Software. vol. 66, pp. 87-97, 2015
[5] D. E. Agthe and R. B. Billings, “Water price influence on apartment complex water use,” J. Water resources planning management. vol. 128, pp. 366-369, 2002
[6] D. R. Maidment and S.Miaou, “Daily water use in nine cities,” J. Water Resources Research. vol. 22, pp.845-851, 1986
[7] G. J. Syme, Q. Shao, M. Po and E. Campbell, “Predicting and understanding home garden water use,” J. Landscape and urban planning. vol. 68, pp. 121-128, 2004
[8] L. House-Peters, B. Pratt, and H. Chang, “Effects of urban spatial structure, sociodemographics, and climate on residential water consumption in Hilsboro, Oregon,” J. American Water Resources Association, vol. 46, pp. 461-472, 2010
[9] L. House-Peters and H. Chang. “Urban water demand modeling: review of concepts, methods, and organizing principles,” J. Water Resources Research, vol. 47.
[10] M. Romano and Z. Kapelan, “Adaptive water demand forecasting for near real-time management of smart water distribution systems,” J. Environmental modelling and Software. vol. 60, pp. 265-276, 2014
[11] R. C. Balling, Jr. P. Gober, and N. Jones, “Sensitivity of residential water consumption to variations in climate: An intraurban analysis of Phoenix, Arizona,” J. Water resources research. vol. 44, 2008
[12] R. C. Balling and H. C. Cubaque, “Estimating future residential water consumption in Phoenix, Arizona based on simulated changes in climate,” J. Physical Geography. vol. 30, pp. 308-323, 2009
[13] S. Gato, N. Jayasuriya, and P. Roberts, “Temperature and rainfall thresholds for base use urban water demand modelling,” J. Hydrology. vol. 337, pp. 364-376, 2007
[14] “The acorn user guide”, CACI Limited, London, 2014