Title
A preliminary investigation of water usage behavior in single-family homes

Journal
Building Simulation, 10(6)

ISSN
1996-3599

Authors
Xue, Peng
Hong, Tianzhen
Dong, Bing
et al.

Publication Date
2017-12-01

DOI
10.1007/s12273-017-0387-7

Peer reviewed

A Preliminary Investigation of Water Usage Behavior in Single-Family Homes

Peng Xue [a,b], Tianzhen Hong [b,*], Bing Dong [c], Cheuk Ming Mak [d]

[a] College of Architecture and Civil Engineering, Beijing University of Technology, Beijing 100124, China

[b] Building Technology and Urban Systems Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

[c] University of Texas at San Antonio, Department of Mechanical Engineering, One UTSA Circle, San Antonio, TX 78249, USA

[d] Department of Building Services Engineering, The Hong Kong Polytechnic University, Hong Kong, China

*Corresponding Author: T. Hong. Tel.: 1(510)486-7082; Fax: 1(510)486-4089; E-Mail: thong@lbl.gov

Abstract:

As regional drought conditions continue deteriorating around the world, residential water use has been brought into the built environment spotlight. Nevertheless, the understanding of water use behavior in residential buildings is still limited. This paper presents data analytics and results from monitoring data of daily water use (DWU) in 50 single-family homes in Texas, USA. The results show the typical frequency distribution curve of the DWU *per household* and indicate personal income, education level and energy use of appliances all have statistically significant effects on the DWU *per capita*. Analysis of the water-intensive use demonstrates the residents tend to use more water in post-vacation days. These results help generate awareness of water use behavior in homes. Ultimately, this research could support policy makers to establish a water use baseline and inform water conservation programs.

Keywords: Water usage behavior; daily water use; data analytics; occupant behavior; residential water consumption

1 Introduction

Most countries around the globe are experiencing a water crisis today. One-third of the global population lives without access to a toilet. A number of people equal to twice the population of the United States live without access to safe water (WHO and UNICEF 2015). Drought conditions in the United States, including threatening drought in California over the last four years, are causing a re-examination of the value of water. Several western states in the United States are surviving the most severe drought conditions in history, with normal, seven-day average stream flows at “extreme hydrologic drought” and “severe hydrologic drought” levels (USGS 2015). The latest 5-year report of the U.S. Geological Survey (Maupin et al. 2010) indicates that total domestic water use, including self-supplied withdrawals and public-supply deliveries, was at 103,709 million liters a day in 2010, with California and Texas ranked the first two in the total water withdrawals among 50 states. The crisis points to a lack of water, but also to poor water management (Cosgrove and Rijsberman 2014). As a nation overall, average domestic daily water use (DWU) *per capita* is reported as 333 liters (88 gallons), which includes potable and non-potable water and includes both indoor and outdoor use. The average per capita use for total domestic water use decreases 10% in last five years and it still represents potential for water conservation (Maupin et al. 2010). Significant energy and associated cost savings are also possible with the reduction in water demand (Malinowski et al. 2015).

Studies over the last decade found that domestic water use is related to many factors (Zhang and Brown 2005; House-Peters and Chang 2011; Ouyang et al. 2013). In an Arizona study, Balling et al. (2008) claimed that 70% of household monthly variance in water use could be explained by atmospheric conditions in the state. In other studies (Praskievicz and Chang 2009; Breyer et al.

2012) also confirmed that the weather condition plays a key role in water use in that country. Household income (Grafton et al. 2011; Kenney et al. 2008) was found to have a positive correlation with water use amount normalized by house size, as well as irrigable lot size (Harlan et al. 2009). Water pricing policy was shown to have influence on single-family residential water use (Polebitski et al. 2010). A study by Wentz et al. (2014) showed that the age of residents was not a significant factor affecting domestic water use, while another study showed that the number of teenagers was a key variable of demostic water use (Aquacraft 2015; Schleich and Hillenbrand 2009). Many other physical building characteristics, including building size (Campbell et al. 2004; Mazzanti and Montini 2006; Tinker et al. 2005), the number of bedrooms per house (Kenney et al., 2008) and housing typology (Fox et al., 2009) are also found to have impacts on water usage. Rosenberg and Madani (2014), in their editorial, suggested that there is a need to think how water interacts with energy. Household water and energy use are heterogeneous and skewed with large variations among households, but for individual appliance shows great energy-water linkage (Abdallah and Rosenberg 2012).

Water use characteristics can only be observed and recorded by a person with relatively long intervals before the installation of data loggers. The output could be just the descriptive results such as the DWU *per capita* (Bullock et al. 1980) and the hourly water use per household (Papakostas et al. 1995). During the mid 1990s, researchers in Boulder, Colorado, started using data logging techniques in data collection (DeOreo and Mayer 1994, DeOreo et al. 1996). With this technique, a computerized sensing device is attached to the water meter and measures flow into the house at 10-second intervals. This makes it possible to obtain and analyze good resolution of water use data from a larger sample. Using new techniques, water conservation could be greatly increased by looking at the logging data in two ways (Cominola et al. 2015), which correlate with two types of information received from meters or sub-meters. The first type identifies the water metering devices, which can be used for end use analysis (Cardell-Oliver et al. 2016). Thus, benchmarks could be established and water-saving devices could be promoted for water conservation. The second type of information is the occupant behavior reflected from the data, which can be used for water policy makers. The behavior model could then be built for water use visualization and prediction, which also helps consumers understand their water use behavior and may influence them to reduce water use through possible behavioral changes.

End-use analysis includes disaggregate water use into end-use components, such as bathing, washing clothes, washing dishes (Richter 2010), and flushing toilets, etc. In a well-known study, *Residential End Uses of Water* (REUWS), published in 1999 by the Water Research Foundation and the American Water Works Association, researchers showed that the average DWU of 262.3 liters *per capita* per day (lpcd) in single-family homes goes into eight end-use components: toilets, faucets, leaks, clothes washers, dishwashers, showers, baths, and other (Mayer et al. 1999). Other studies (DeOreo et al. 2011) show similar findings, which are essential for establishing benchmarks (Mayer 2009) and developing water devices. Other research shows that introducing engineered water efficiency devices could reduce indoor water use by 35% to 50% (Inman and Jeffrey 2006). Heberger et al. (2014) found that the average indoor water use could decrease to 32 gallons *per capita per day* for California residents with efficient water use appliances and fixtures.

Occupant behavior-related water use in residential buildings is a critical issue for water conservation, and water use prediction (Kontokosta and Jain 2015; Chu et al. 2009; Chang et al. 2010; Suero et al. 2012). Occupant behavior is complex and stochastic, causing a high DWU variability both among residences and within the same residence (Lutz 2012). Corral-Verdugo et al. (2003) found that general beliefs could influence specific water beliefs, and in turn could affect water consumption. Willis et al. (2010) investigated the effect of visual display monitors on residents' shower behavior; results confirmed a significant effectiveness with 27% reduction in a shower water use event. Consumer behavior may also be negatively affected by water-saving devices. Inman and Jeffrey (2006) found that residents took longer showers and consumed more water after installation of water-saving devices, due to the belief that their water-saving devices would save water (rebound effect).

In summary, most existing studies on water use behavior models are observed from the perspective of use time of water-consuming devices and lack in-depth behavioral analysis. While energy-related occupant behavior has been studied extensively for residential and commercial buildings (Dong et al., 2015; Hong et al. 2015; Yan et al. 2015), water use behavior is under-researched. Aiming to provide insights into household water use behavior, this paper presents analytical results from monitoring data of DWU in 50 single-family homes in Texas, USA, as well as exploration of root causes behind household water use behaviors.

2 Methodology

2.1 Description of Dataset

This study uses data collected through Pecan Street Dataport (Pecan Street, Inc.), which is the world's largest source of disaggregated customer energy and water use data. The data are stored in a SQL database, which consists of weather, water usage, energy audits, annual surveys, energy consumption and other information (e.g., gas use). The Pecan Street database includes 1338 houses, 1105 of which are still active. The database started collecting data January 1, 2011, and continues up to the time of this study (September 26, 2015). Energy data is recorded in 1-minute time intervals, while water use is recorded as daily sum before May 10, 2013, and by minute from then on.

In this study, household information comes from the survey tables. DWU value is calculated from the water usage table, which shows a household's total water use within a specific time interval (by day or minute). Energy consumption data are from the hourly energy-use table, which contains 67 columns showing energy consumptions of different appliances in a house. Water use data in the database is sparse and not always continuous. After data processing (excluding the ones without water use data), 50 single-family houses are selected for this study.

2.2 Pre-processing of Dataset

To exclude outliers and unexplained noise, and improve the quality and reliability of the data, the dataset was pre-processed with a series of procedures.

2.2.1 Translating

The data of houses were first downloaded from the database as comma-separated-value (CSV) files. The main purposes were to calculate the DWU and daily energy use (DEU) from the cumulative data (by hour and minute) for each house, and to convert units of the measured data. All pre-processed data were further processed in the following steps.

2.2.2 Cleaning

The second step was to clean all the translated data obtained from the previous step, which includes summarizing all household data into one sheet with outdoor air temperature in chronological order and removing data (cumulative raw data) with gaps of more than one day.

2.2.3 Checking

After the translating and cleaning steps, 11852 logging data points for 60 houses were collected in one Excel sheet. Some zero values of DWU were also included, which reflected that no residents were home and consumed no water on those days. As the zero values may have a significant influence on calculating the average and DWU values, small values such as 0 and 1 liter/day were excluded in the study of water use behavior. After applying the above criteria, 10 of the households with valid data had data for less than a month. These 10 households were excluded from the originally selected 60. In the end, water use data for the remaining 50 households were used in the study.

2.2.4 Summarizing data

After all data were pre-processed, a dataset of 10659 valid DWU values from 50 houses was built. Combined with the house information, the data were summarized by different objectives and shown in Table 1.

2.3 Overview of Data Analysis

The procedure and methods of data analysis are shown in Fig. 1. The first step is to investigate typical water use patterns through the time and frequency distribution of DWU among houses. Then, correlation analysis based on Spearman's rank coefficient method is conducted between

DWU and outdoor air temperature, day of the week, and season are also studied. A sudden (anomaly) peak of DWU is found as a common phenomenon in many homes, which will be discussed with possible causes in Section 3.3. The second step is to find water use patterns between weekdays and weekends based on normalized data from all 50 homes, and to establish a baseline water usage model of DWU for single-family homes. By comparing the results among different houses, the third step is to find related factors affecting DWU, namely residents' income, education, age and daily activity. The information for all of the selected 50 households is shown in the appendix (except for the exact house ID which was anonymized due to privacy concerns).

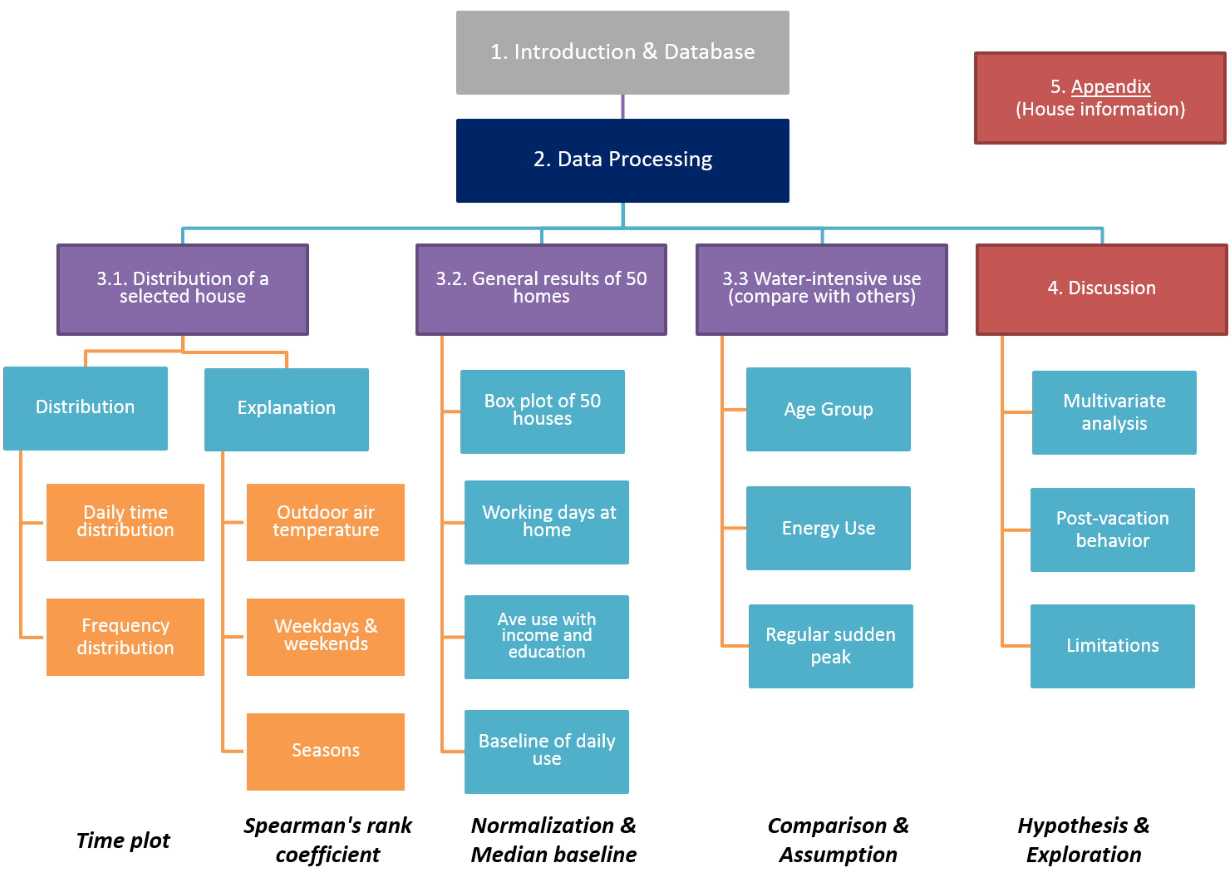

Fig. 1 The procedure of data analysis

2.4 Description of Statistical Methods for Data Analysis

2.4.1 Spearman's rank correlation coefficient and Stepwise regression

Spearman's rank correlation coefficient (Zar 1972) is adopted to describe the relationship between two variables by assessing the monotonic function. A perfect value of +1 or −1 occurs when one variable is a perfect monotone function of the other. The coefficient ρ could be computed from:

$$\rho = 1 - 6 \sum d_i^2 / n(n^2 - 1) \tag{1}$$

where d_i is the difference between ranks of two variables; i is the case number; n is the total number of cases. This correlation coefficient was applied to investigate the relationship between the DWU per house and the age groups of occupants in the house.

Stepwise regression is a step-by-step selection model for multivariate analysis, which involves automatic selection of independent variables. This analysis was further used to find the most parsimonious set of predictors for DWU.

2.4.2 Frequency distribution

In this study, frequency distributions are displayed as graphs that show the frequency of DWU in a house or the whole dataset. A frequency distribution shows a summarized grouping of DWU values divided into mutually exclusive intervals and the number of occurrences in an interval.

2.4.3 Median for baseline

Water use distribution should be studied with medians, not averages, as the feature is not symmetrical (Lutz 2012). In this study, median values of all logging data can be obtained in three different ways. The first way is to obtain the median values from all logging data directly, the second one is to obtain median value from all household median DWU values, and the third one is to obtain median value from all household average DWU values.

The first way chooses the median value from all the data but ignores the fact that the number of data points from each household is not the same (as shown in appendix). The second way is more appropriate, which considers the differences between households and obtains the median values of each house first. However, the median value of a house can only be explained as the most likely condition. The value itself ignores the high water use condition and sudden peak, which should be considered as the behavior of the residents. Therefore, the third way is most appropriate to establish the baseline, which calculates the average values of DWU for each household first and then finds the median DWU for the entire dataset of 50 households.

3 Results and Analysis

3.1 Statistical analysis of a single house

We started studying the residents' water use behavior in a single house. House No. 7 is selected with the most logging data points—538 days of valid data—from the 50 monitored homes.

3.1.1 Time distribution of DWU

The valid data lasts for almost 18 months with an interruption of 3 months. The result of time distribution of DWU is shown in Fig. 2.

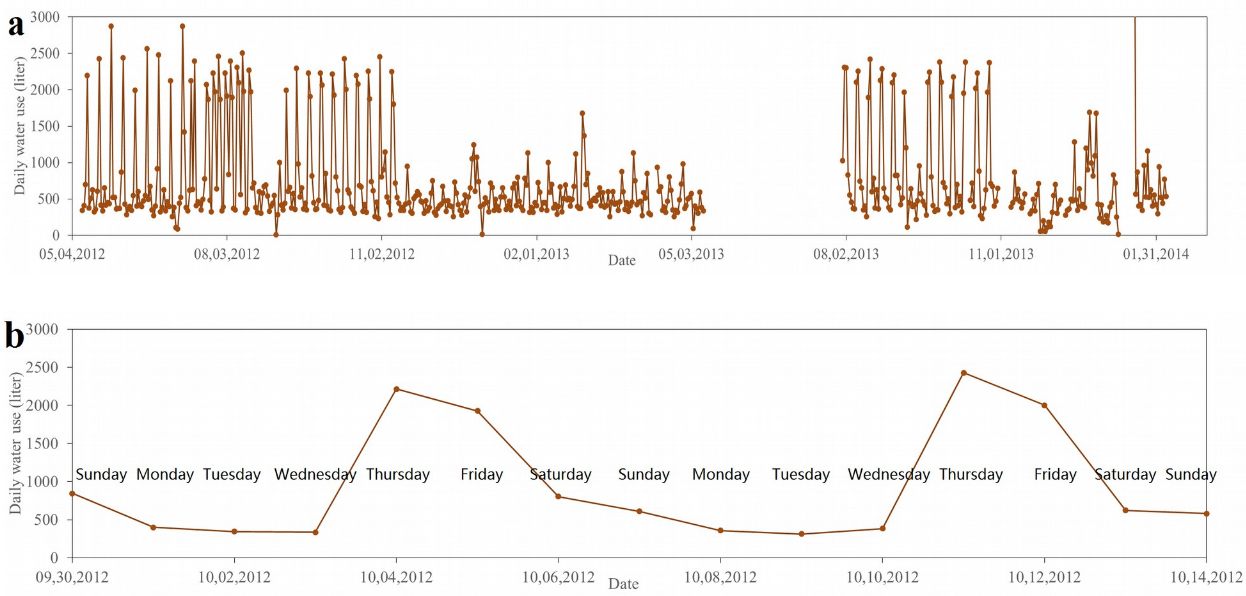

Fig. 2 Daily time distribution of water use (House No. 7): a) with all data points; b) in two weeks

Residents of House No. 7 have a baseline water use of 300 lpd. They used more water on one or two specific days every week, and the peak water use reached around 2000 lpd during May and October, while the peak in other months is almost half at 1000 lpd. Based on the survey information, the residents in this house have a habit of watering. It can be concluded that the DWU behavior is influenced by day of the week and seasons. The major change in peak water use between seasons appears to be due to irrigation.

3.1.2 Frequency distribution of DWU

With the 538 valid data of DWU, the frequency distribution is shown as Fig. 3. The X-axis interval is set as 40 or 75 liters per day (lpd) and Y-axis shows the occurrence number of days.

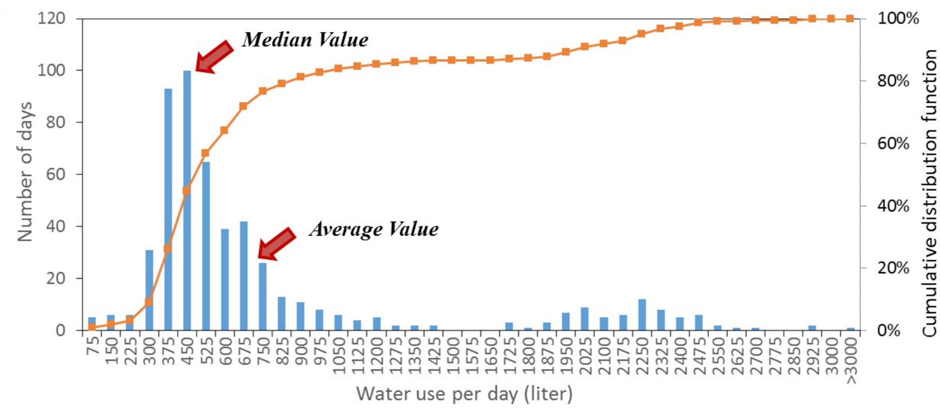

Fig. 3 Frequency and cumulative distributions of DWU (House No. 7)

The frequency distribution of DWU shown in Fig. 3 is neither symmetrical nor normal distribution. The curve has a long tail, it features a striking peak around 450 lpd, and most of the data are equal or greater than 300 lpd. However, there are still 189 days when the DWUs are much more than the average of 730.28 liters per household per day (lphd). A second peak appears around 2250 lpd, which indicates another behavior pattern of high water use that needs further study. It is worth noting that this is a typical feature of DWU frequency distribution: almost all 50 homes show a distribution with two or three peaks (Fig. 11 and Fig. 12).

3.1.3 DWU impacts by outdoor air temperature, day of the week, and seasons

DWU differs from day to day and has large variations. Monitored data from House No. 7 are shown in Fig. 4 with the X-axis of outdoor air temperature. The dataset grouped by weekdays and weekends is shown in Fig. 4a, while it is also grouped in seasons as shown in Fig. 4b. The seasons are divided by solstices and equinoxes. The correlations of DWU with week days, seasons and OAT are shown in Table 2.

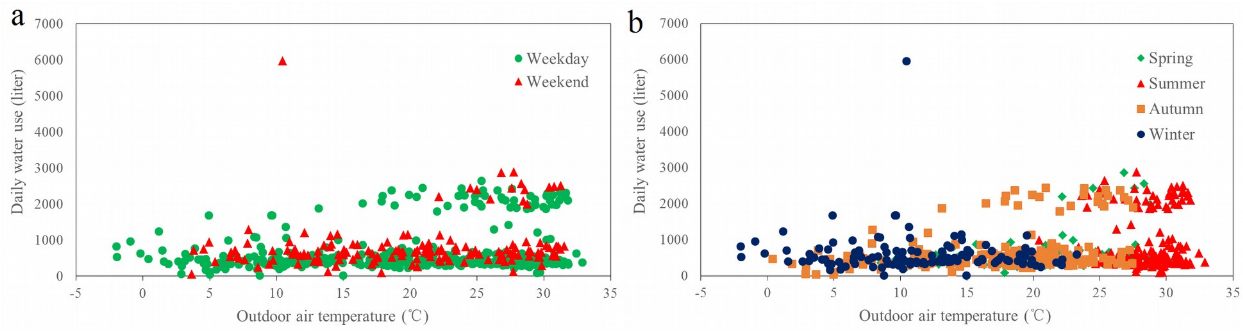

Fig. 4 DWU (House No. 7) with outdoor air temperature: a) by weekdays and weekends; b) by seasons

The phenomena of summer peak in Fig. 2a and the dual peaks in Fig. 3 can be reflected as the two-layer feature in this figure. Table 2 reflects this house has a relative steady water use pattern among the week. Fig. 4 shows that the relationship between household DWU and outdoor air temperature is not linear and Table 2 shows the similar result. However, the two-layer feature indicates that residents keep basic requirements of water use and do not use much water for irrigation when the outdoor air temperature drops below 15℃. As seen from the higher layer in Fig. 4a, the high water use behavior occurred in both weekdays and weekends, indicating that the residents have a constant 2250 lpd of water use once or twice a week. Fig. 4b and Table 2 shows that the water behavior has strong seasonality, with the winter months having lower values. The average DWU values from spring to winter are 593.87 lphd, 948.03 lphd, 694.81 lphd, and 607.42 lphd, respectively.

There is also an isolated data point with a very high value in Fig. 4, which is more than twice the value of other data points. This kind of anomaly peak happens in almost half of the 50 homes, which may result from water leaks, watering, car washing or filling swimming pools. This anomaly is considered further in the Discussion section.

3.2 Statistical analysis of 50 houses

3.2.1 Data normalization

The data show large variations in water use from day to day and from home to home. It is important to normalize the water use for a single-family home on the basis of number of persons living in the home and the total floor area of the house. Fig. 5 shows several water use metrics for the 50 homes, including DWU median per household, average DWU per household, average DWU *per capita*, and cumulative distribution function of DWU per household. The results are sorted by the average DWU per household.

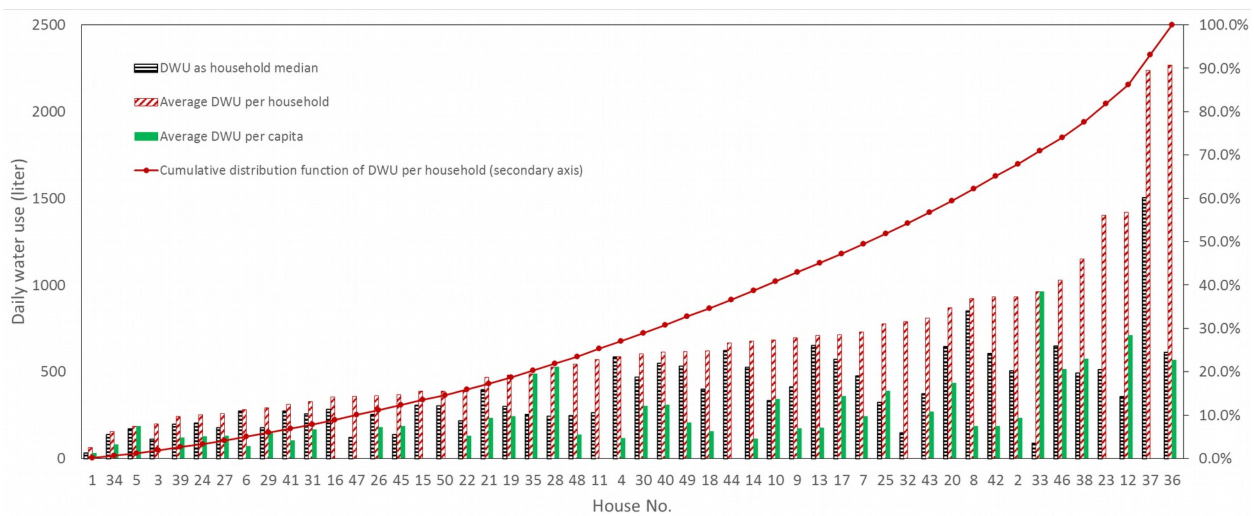

Fig. 5 Water use metrics for 50 houses

As seen in Fig. 5, the average DWU differs significantly from house to house; the largest two houses reach 2250 lpd. This figure also shows that the top 26% of households use 48% of total water. The overall average DWU of all houses is 676.27 liters, and the median DWU of each house is also shown in the figure, which is less than the average DWU. The median DWU in 95% of the houses is between 90 lphd and 650 lphd. This result reflects the frequency distribution curve of DWU is not symmetrical and the long tail is significant. House No. 36 has a median DWU of 613.17 lphd and an average DWU of 2267.84 lphd. The figure also shows the

DWU normalized by the number of residents. The calculated result indicates that nearly 25% of people use 51% of the total water. The overall average DWU *per capita* is 272.81 liters. The DWU, normalized by *capita* and square meter, is also provided in the figure, but it can be much higher for small houses.

In conclusion, this study found that it most effective and appropriate to study the chosen data normalized by *capita*. These results also show that the Pareto Principle (Sanders 1987) is in operation in this water use study. If high water use households (or people) improve (decrease consumption), water will be significantly saved.

3.2.2 Occupied and vacant houses during weekdays

The next analysis had to do with household DWU during the week versus on the weekend. The houses grouped by the types of occupancy are shown in Fig. 6, and the average DWU of each house is separated with the average DWUs for both weekdays and weekends.

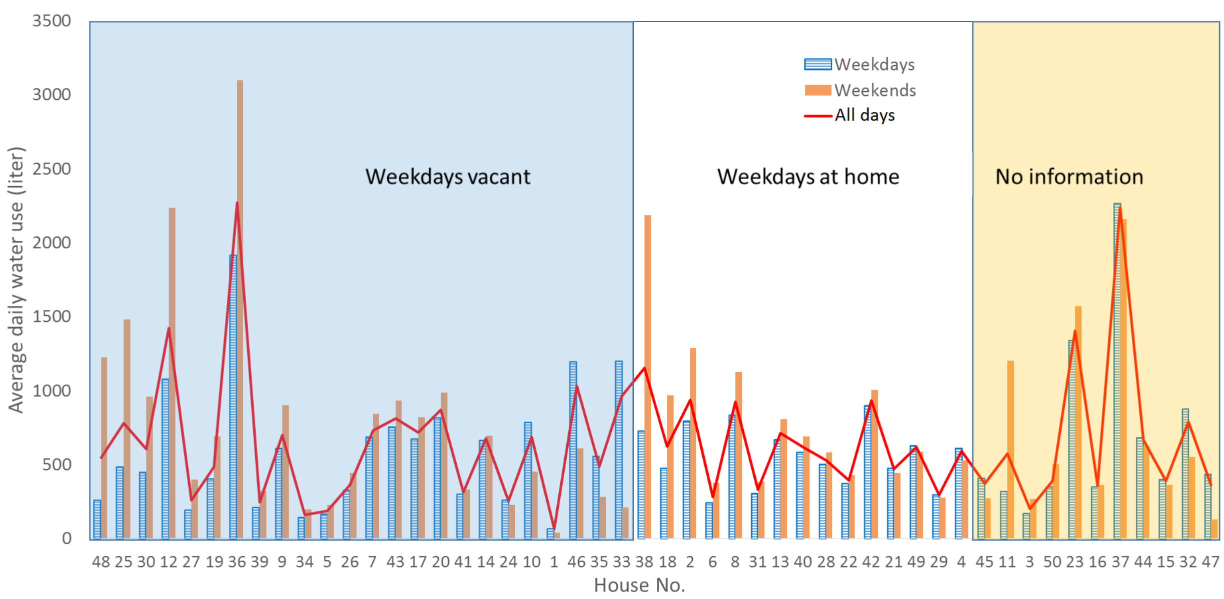

Fig. 6 Average DWU during weekdays and weekends for 50 houses among three occupancy

types

Fig. 6 shows the average DWU values of all houses represented by a solid line. Compared with this solid line, it is clear to show whether residents use more water during weekdays or not. Average DWU is closer to weekdays DWU since weekdays have a higher weighting factor.

Results of this analysis show that 68% of houses consume more water per day on weekends than weekdays. However, some houses have higher average DWU on weekdays. Considering the occupancy on weekdays, no significant relation can be found. According to results of first two groups in Fig. 6, both groups have households using more water on weekdays than weekends. It seems that DWU is less affected by occupancy than by residents' habits.

3.2.3 Income and education

To test assumptions that might explain the correlation between income, education, and DWU, this analysis normalized DWU and income by the number of residents in each house. The assumption was that residents with higher income may have a higher standard of living and consume more water. The personal income is calculated from house information (Appendix) and grouped in seven levels, as shown in Table 3. The relation between DWU *per capita* and personal income is shown in Fig. 7, with the levels of education presented in different shapes.

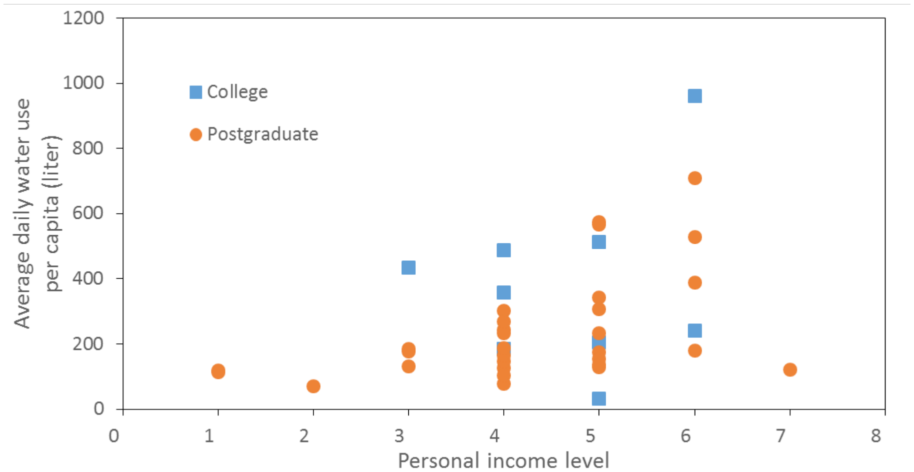

Fig. 7 Average DWU *per capita* with personal income and education

The analysis represented in Fig. 7 seems to show that DWU *per capita* has a positive correlation with the personal income level—residents with higher incomes seem to consume more water. Focusing on the personal income level from 3 to 6, residents with college degrees use more water than the postgraduates, on average. Though the numbers of cases are not equal, people with undergraduate college degrees consume the most in three out of four income levels. Using these 40 cases as a guide, it is reasonable to say that residents with more education are likely to use less water that those residents with less education.

3.2.4 Baseline of DWU

Looking at all valid DWU data points as a whole, the frequency distribution of all 50 houses is shown as Fig. 8. The X-axis interval is set as 40 lpd.

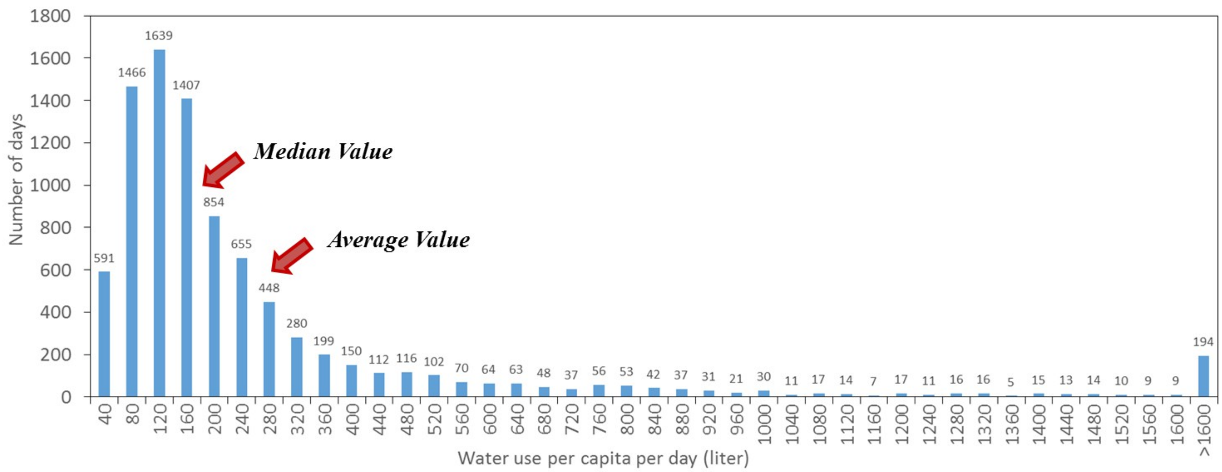

Fig. 8 Frequency distribution of DWU among 50 houses

From Fig. 8, it can be seen that the distribution curve only has one peak and a long tail. The average water use *per capita* per day across all 8949 data points is 272.81 lpcd but the standard deviation can be as high as 521.48 lpcd. When studying the baseline, the median value is often adopted as a fair rule. In this study, median value is obtained from all 50 houses' average DWU value (Fig. 5); the result is 186.00 lpcd. Therefore, the baseline of the DWU for these households can be set as 186.00 lpcd. Compared with the value of 333 lpcd, which is provided from the US Survey Circular, 186 lpcd seems stringent. However, this baseline is established for drought regions and this value ensures half houses are easy to meet the requirement without any guidance. Due to water crisis today in more regions, we should establish a stringent baseline to encourage people's awareness of water savings.

3.2.5 Cross comparison of all houses

After being normalized by the number of residents, the DWU of each house can be studied in more detail. The box plot of DWU *per capita* for 40 houses is shown in Fig. 9 as the other 10 houses have no information about the number of residents. The results are sorted by the median

DWU value.

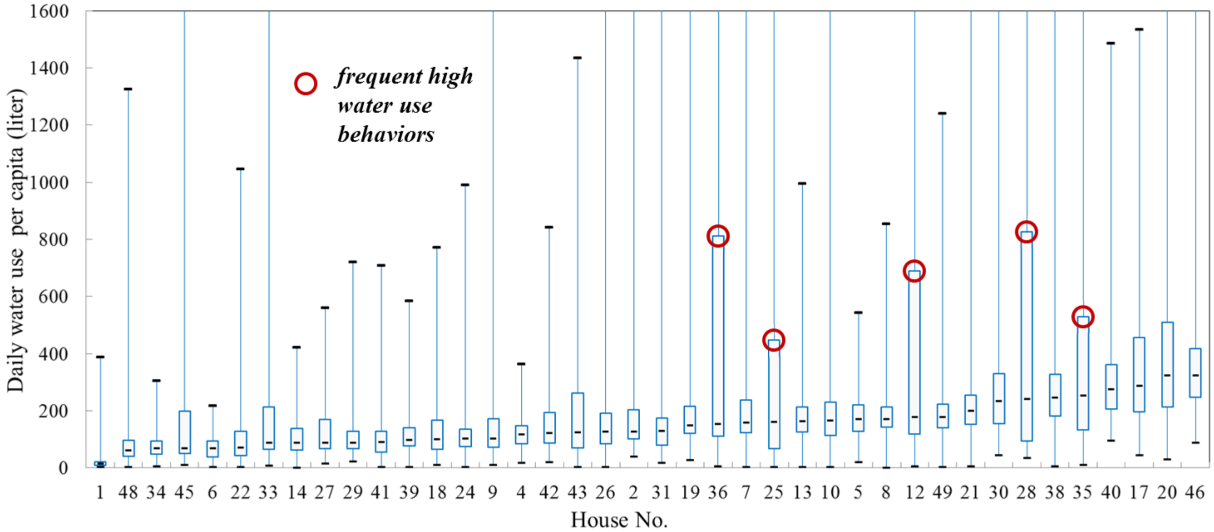

Fig. 9 Box plot of DWU *per capita* for 40 houses sorted by median values

As seen in Fig. 9, the median values of DWU *per capita* among the 40 houses are between 15 liters and 320 liters. The median value of these household median DWU values is 127.76 liters. Focusing on the highest value at each house shows that 50% of households have median DWU values higher than 1500 lpd, which means these data may be experiencing anomaly peak. Some of the interquartile ranges (cubic length) shown in the Fig. 8 are very big. This result reflects that the residents in some households—namely houses 12, 25, 28, 35, and 36—have frequent high water use behaviors compared to their own average DWU. The detailed DWU results on a long interquartile range will be shown the Discussion section.

3.3 Analysis of Water-intensive use

As discussed, the DWU of each household—even the value of DWU *per capita*—differs significantly. This research next looked at which factors could account for higher water use in some households compared to others, factors including higher personal income, better education, teenagers at home, or washing behaviors.

3.3.1 Age group

The next analysis examined the assumption that teenagers use more water than other age groups. It is difficult to separate DWU by age groups since a house may hold people in several different age groups. Therefore, Spearman correlation coefficient is adopted to study the relation of DWU per household with the number of residents in each age group. The key group will be presented with significant coefficient, which means that the corresponding group has more weight to inform the house total DWU. The result of the statistical analysis of 40 houses and 108 residents is shown in Table 4.

As seen in Table 4, no significant value is presented. The results make it clear that there is no significant correlation between DWU per home and age groups; there can be no assumption that any age group uses more water than others.

3.3.2 Correlations with Appliance Energy Usage

Water use was compared to energy use to see if energy use somehow correlates to water use. Among the 50 houses, only four have both daily total energy consumption data and DWU data at the same time. House No. 19 has the longest monitored days among these four houses and its energy use is also sub-metered into three separate data streams, all assumed to have direct relation with water use behavior: bathroom, clothes washer, and dish washer. Fig. 10 shows the DEU of those three data streams and the DWU of House No. 19.

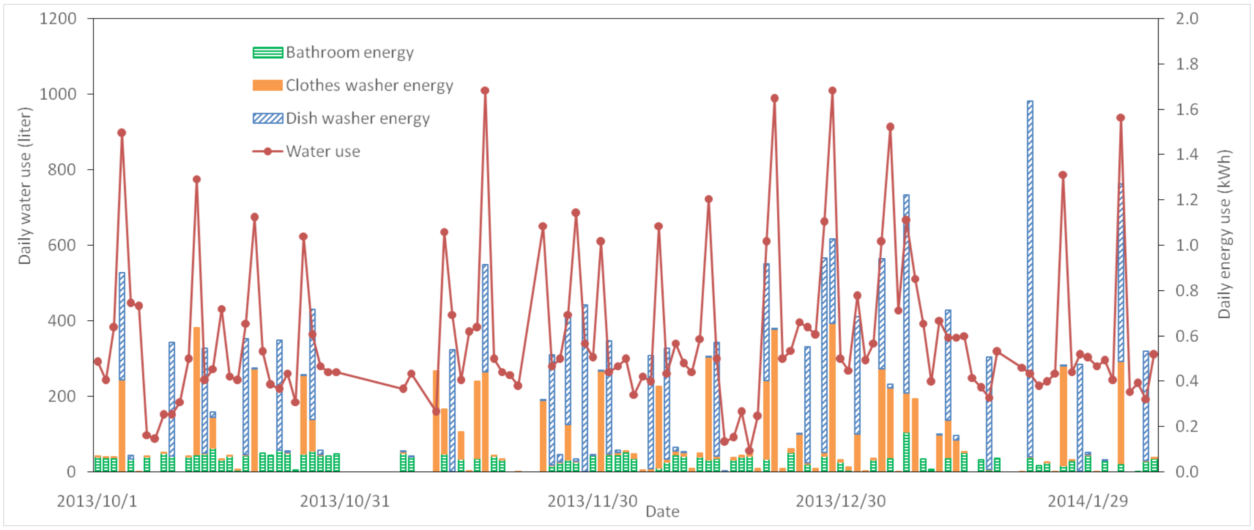

Fig. 10 DWU (House No. 19) and appliances energy use

As seen from Fig. 10, the bathroom shows the most constant and consistent use, at the frequency of 92 days out of a total of 116 monitoring days. While the clothes washer and dish washer are operated in 43 and 37 days, respectively, which are twice a week on average. Energy consumption in the bathroom is much less than the energy consumption of the clothes washer and dish washer, on average. This may result from the fact that light bulbs often have the power level of less than 100 W, while the dish and clothes washers have the power level of more than 2000 W. Given their different power, even if lights are turned on in the bathroom, its overall consumption is lower than that of dishwashers or washing machines. In general, the DWU has ups and downs over the monitored days; it seems higher DWU points have a corresponding higher use of energy.

To test the bivariate relationship between DWU and DEU, Spearman correlation coefficient is adopted. The relevant appliance will be presented with significant coefficient, which means the corresponding appliance has more weight to inform the house total DWU. The result of the statistical analysis for 116 days is shown in Table 5.

From Table 5, a positive correlation is seen between DWU and DEU with a correlation coefficient of 0.463. The disparity may come from the incomplete statistics and residents' different behaviors between water use and energy use. To be more specific, the DEUs of bathroom, clothes washer, and dish washer are also tested and results show that the DEU of the clothes washer has a significant correlation with DWU, with a correlation coefficient of 0.545. However, the DWU per household could not be predicted by the energy use of the bathroom, and this may result from having windows and thus little lighting energy use in their bathrooms.

Not all families use dish washers or other appliances. Restricted by the sample size, this result only proves that energy use can indicate the condition of water use in residential buildings qualitatively. The most important is saving water saves energy.

3.3.3 Regular sudden peak

As outlined in the Fig, 4, half of the monitored houses have anomaly peaks. To investigate possible reasons for this phenomenon, comparing the DWU data of a certain house against its average DWU and visualizing the time-series DWU data is an effective way.

The sudden peak can be also regular in some periods, as shown in Fig. 11.

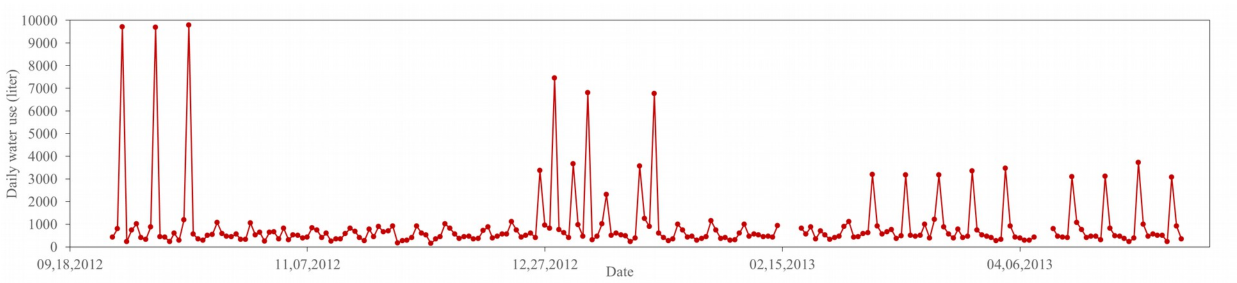

Fig. 11 Regular sudden peak of DWU (House No. 2)

In Fig. 11, three-layer DWU data can be recognized and the sudden peak values at the third layer are two or three times higher than the ones at the second layer. All data points at layer two are around 3500 lpd and happened on Wednesdays, which indicate that residents had a centralized water use habit once a week, such as car washing and landscape watering. All the six sudden peak points happened on Saturdays. This could be explained by refilling or changing the water in a swimming pool.

4 Discussion

4.1 Multivariate regression analysis

After the bivariate analysis, stepwise regression was further used to find a set of predictors that would be effective in predicting DWU. DWU was set as the dependent variable, and six relevant factors (OAT, weekdays, season, EDU of clothes washer, dish washer and bathroom) were chosen as independent variables. A statistically significant model was then selected ($R = 0.828$, $F = 81.313$, $P < 0.001$) from three obtained models. Table 6 shows the results of the analysis.

As can be seen from Table 6, the DEU of clothes washer, Weekdays, and DEU of dish washer all had significant P-values. The standardized beta reveals the relative influence of these three factors. Basically, the degree of DEU of clothes washer had a principal influence on DWU, and the other two factors had a secondary influence.

4.2 Post-vacation water use behavior

Residents' behaviors affect DWU significantly. Some consumers use more water on weekends; some behaviors lead to a two-layer DWU distribution while others a three-layer DWU distribution; and some people have a number of high water use behaviors. This section discusses a special finding—water use behavior after vacation or travel, with results presented in Fig. 12.

Fig. 12 High DWU after vacation or travel (House No. 9)

The dataset of House No. 9 loses DWU data for three months and separates the distribution into two parts. It is clear that the residents had a consistent DWU behavior in the first time period. However, in the second time period, the DWU showed four sudden peak points, which have higher values than the normal DWU value. These four sudden peak points are found to always come after a house vacancy. Therefore, this sudden peak can be interpreted as the post vacation or travel behavior. People shower, wash clothes, clean rooms, water landscaping, and other behaviors after a trip, and this centralized water use makes the sudden peak of DWU. Like energy use, water is not used during travel periods for homes; but unlike energy use, residents will use more water after a trip.

4.3 Limitations

Three limitations are identified in our analysis. The first one is the house information. After the data processing, the valid DWU sample size is 10659. Due to some houses lacking data of floor area and the number of occupants, the normalized data become less and less. Especially with the limitation of the energy consumption data, the valid number of houses comes down to four. Secondly, we could not separate indoor and outdoor water use, due to no sub-metering of water use data is available in this dataset. The irrigation is a key part of outdoor water use, and it is the main pattern of the regular high water use. If the daily water use can be separated into two parts, indoor water use can be analyzed more comprehensively without the sudden peak, and the outdoor water use can be studied with detailed weather factors. Thirdly, we have new interesting findings about residents' water use behavior after coming back from vacation. However with the same reason, we could not know the sub-items clearly without sub-meter data, so the detailed pattern and suggestions could not be provided at this stage. To address these limitations, it is recommended that more efficient water metering and sub-metering devices be installed for improving future water use behavior study.

5 Conclusions

This study presents an integrative analysis towards understanding water use behavior in single-family homes. Analysis of water use and household data from the Pecan Street database draws a conclusion that a quarter of the residents consumed half of the total water, and the baseline of DWU can be set as 186.00 lpcd for this drought region. The analysis also shows the typical frequency distribution curve of the DWU *per household* has dual peaks. It further indicates that the personal income, education level and energy use of appliances have statistically significant effects on the DWU *per capita.* Analysis of the water-intensive use periods demonstrates the residents tend to use more water after returning home from a trip, which is quite different from the energy use behaviors. These findings help generate awareness of water use behavior in single-family homes and assist policy-makers in establishing appropriate guidelines and standards for residential water use. The high water use houses (mostly higher income households) should be engaged and encouraged to save water as they have greater potential for water conservation. Besides, education strategies and customized feedbacks are great long-term policy for water conservation.

Acknowledgments

This work is supported by the Assistant Secretary for Energy Efficiency and Renewable Energy of the U.S. Department of Energy under contract number DE-AC02-05CH11231. It is also part of the research activities of International Energy Agency Energy in Buildings and Communities Program Annex 66, definition and simulation of occupant behavior in buildings. The source data were provided by Pecan Street, Inc. (http://www.pecanstreet.org/), headquartered in Austin, TX. The authors thank this nonprofit research institute for allowing us access to their subscriber water usage database.

Appendix

Table 7 General information about the sample houses (excluding the house id in the database)

Case No.	Days of the valid data	House size (m^2)	People	Education	Total annual income	Workdays at home	Irrigation system
1	145	159.1	2	College	$100-000 - $149-999	No	Yes
2	219	248.0	4	Postgraduate	$150-000 - $299-000	Yes	Yes
3	74						
4	100	147.9	5	Postgraduate	$50000 - $74999	Yes	No
5	140	74.5	1	College	$35000 - $49999	No	No
6	132	153.5	4	Postgraduate	$75-000 - $99-999	Yes	No
7	538	197.1	3	Postgraduate	$100000 - $149999	No	Yes
8	321	248.9	5	Postgraduate	$100-000 - $149-999	Yes	Yes
9	283	217.9	4	Postgraduate	$150000 - $299000	No	Yes
10	135	225.4	2	Postgraduate	$100000 - $149999	No	Yes
11	335	217.4					
12	397	179.7	2	Postgraduate	$150000 - $299000	No	No
13	72	251.8	4	Postgraduate	$100000 - $149999	Yes	Yes
14	277	214.6	6	Postgraduate	$50000 - $74999	No	Yes
15	72						
16	324	159.8					
17	314	159.8	2	College	$75000 - $99999	No	Yes
18	390	264.1	4	Postgraduate	$150-000 - $299-000	Yes	Yes
19	146	173.0	2	College	$150000 - $299000	No	Yes
20	346	143.4	2	College	$50000 - $74999	No	Yes
21	184	117.2	2	Postgraduate	$75-000 - $99-999	Yes	No
22	90	180.8	3	Postgraduate	$75-000 - $99-999	Yes	Yes
23	259	197.1					
24	307	132.1	2	Postgraduate	$75-000 - $99-999	No	Yes
25	363	251.8	2	Postgraduate	$150-000 - $299-000	No	Yes
26	384	251.8	2	Postgraduate	$150-000 - $299-000	No	Yes
27	65	87.1	2	Postgraduate	$100-000 - $149-999	No	No
28	364	159.8	1	Postgraduate	$75000 - $99999	Yes	Yes
29	301	171.1	2	Postgraduate	$75-000 - $99-999	Yes	Yes
30	139	125.0	2	Postgraduate	$75000 - $99999	No	No
31	159	100.3	2	Postgraduate	$75-000 - $99-999	Yes	No
32	56						
33	66	141.4	1	College	$75-000 - $99-999	No	Yes
34	138	116.8	2	Postgraduate	$75000 - $99999	No	No
35	66	113.1	1	College	$35-000 - $49-999	No	Yes
36	204	250.7	4	Postgraduate	$150-000 - $299-000	No	Yes
37	100						
38	242	234.2	2	Postgraduate	$100-000 - $149-999	Yes	Yes
39	391	159.8	2	Postgraduate	more than $1-000-000	No	Yes
40	157	199.6	2	Postgraduate	$100000 - $149999	Yes	Yes
41	368	98.3	3	Postgraduate	$100-000 - $149-999	No	No
42	161	294.0	5	Postgraduate	$150000 - $299000	Yes	Yes
43	337	203.8	3	Postgraduate	$100000 - $149999	No	Yes
44	377	185.9					
45	62		2	College	$75-000 - $99-999	Yes	No
46	149	111.1	2	College	$100000 - $149999	No	No
47	78						
48	163	204.0	4	Postgraduate	$150-000 - $299-000	No	Yes
49	134	178.1	3	College	$150,000 - $299,000	Yes	No
50	37						

References

Abdallah AM and Rosenberg DE. Heterogeneous residential water and energy linkages and implications for conservation and management. *Journal of Water Resources Planning and Management*, 2012, 140(3): 288-297.

Aquacraft Inc. Application of end use study data for development of residential demand models. 2015. Retrieved from: http://www.aquacraft.com/wp-content/uploads/2015/09/Residential-Models.pdf

Balling RC, Gober P and Jones N. Sensitivity of residential water consumption to variations in climate: an intraurban analysis of Phoenix, Arizona. *Water Resources Research*. 2008; 44.

Breyer B, Chang H and Parandvash H. Land-use, temperature and single family residential water use patterns in Portland, Oregon and Phoenix, Arizona. Applied Geography. 2012; 35: 142–151.

Bullock DC, Peebles RW and Smith HH. Water usage patterns in the U.S. Virgin Islands. Water Resources Research Center, Caribbean Research Institute, College of the Virgin Islands, 1980.

Campbell HE, Johnson RM and Larson EH. Prices devices people, or rules: The relative effectiveness of policy instruments in water conservation. Review of Policy Research. 2004; 21: 637–662.

Cardell-Oliver R, Wang J and Gigney H. Smart Meter Analytics to Pinpoint Opportunities for Reducing Household Water Use. *Journal of Water Resources Planning and Management*, 2016: 04016007.

Chang H, Parandvash GH and Shandas V. Spatial variations of single-family residential water consumption in Portland, Oregon. *Urban geography*. 2010; 31: 953-972.

Chu J, Wang C, Chen J and Wang H. Agent-based residential water use behavior simulation and policy implications: A case-study in Beijing City. *Water Resour Manage*. 2009; 23: 3267-95.

Cominola A, Giuliani M, Piga D, Castelletti A and Rizzoli AE. Benefits and challenges of using smart meters for advancing residential water demand modeling and management: A review. *Environmental Modelling & Software*. 2015; 72: 198-214.

Corral-Verdugo Vc, Bechtel RB and Fraijo-Sing B. Environmental beliefs and water conservation: An empirical study. *Journal of Environmental Psychology*. 2003; 23: 247-257.

Cosgrove W J, Rijsberman F R. World water vision: making water everybody's business. Routledge: 2014. Retrieved from: http://www.worldwatercouncil.org/index.php?id=961

DeOreo WB, Heaney JP, Mayer PW. Flow trace analysis to assess water use. *American Water Works Association*. Journal, 1996, 88(1): 79.

DeOreo WB and Mayer PW. Project report: A process approach for measuring residential water use and assessing conservation effectiveness. City of Boulder Office of Water Conservation, Boulder, Colorado. 1994.

DeOreo WB, Mayer PW, Martien L, et al. California single-family water use efficiency study. Aquacraft Water Engineering and Management, Boulder, Colorado, USA. 2011. Retrieved from: http://water.cityofdavis.org/Media/PublicWorks/Documents/PDF/PW/Water/Documents/California-Single-Family-Home-Water-Use-Efficiency-Study-20110420.pdf

Dong B, Li Z and Mcfadden G. An investigation on energy-related occupancy behavior for lowincome residential buildings. Science and Technology for the Built Environment. 2015; 21(6): 892-901.

Fox C, McIntosh BS and Jeffrey P. Classifying households for water demand forecasting using physical property characteristics. Land Use Policy. 2009; 26: 558–568.

Grafton RQ, Ward MB, To H and Kompas T. Determinants of residential water consumption: Evidence and analysis from a 10-country household survey. Water Resources Research. 2011; 47(8).

Harlan SL, Yabiku ST, Larsen L and Brazel AJ. Household water consumption in an arid city: affluence, affordance, and attitudes. *Society and Natural Resources*. 2009; 22: 691-709.

Heberger M, Cooley H and Gleick P. Urban water conservation and efficiency potential in California. 2014. Retrieved from: http://www.nrdc.org/water/files/ca-water-supply-solutions-urban-IB.pdf

Hong T, Taylor-Lange SC, D'Oca S, Yan D, Corgnati S. Advances in research and applications of energy-related occupant behavior in buildings. *Energy and Buildings*, 2015. In Press. doi:10.1016/j.enbuild.2015.11.052.

House-Peters L and Chang H. Urban water demand modeling: Review of concepts, methods, and organizing principles. Water Resources Research. 2011; 47: W05401.

Inman D and Jeffrey P. A review of residential water conservation tool performance and influences on implementation effectiveness. *Urban Water Journal*. 2006; 3: 127-143.

Kenney D, Goemans C, Klein R, Lowrey J and Reidy K. Residential water demand management: Lessons from Aurora, Colorado. Journal of the American Water Resources Association. 2008; 44: 192–207.

Kontokosta C E, Jain R K. Modeling the determinants of large-scale building water use:

Implications for data-driven urban sustainability policy. Sustainable Cities and Society, 2015; 18: 44-55.

Lutz J. Hot water draw patterns in single-family houses: findings from field studies. Lawrence Berkeley National Laboratory, 2012: LBNL Paper LBNL-4830E. Retrieved from: http://escholarship.org/uc/item/2k24v1kj

Malinowski PA, Stillwell AS, Wu JS and Schwarz PM. Energy-water nexus: potential energy savings and implications for sustainable integrated water management in urban areas from rainwater harvesting and gray-water reuse. *Journal of Water Resources Planning and Management*. 2015; 141(12): A4015003.

Maupin MA, Kenny JF, Hutson SS, Lovelace JK, Barber NL and Linsey KS. Estimated use of water in the United States in 2010: U.S. Geological Survey Circular 2014.1405, 56 p. Doi: http://dx.doi.org/10.3133/cir1405.

Mayer PW. Water efficiency benchmarks for new single-family homes. Aquacraft Water Engineering and Management, Boulder, Colorado, USA. 2009. Retrieved from: http://www.watersmartinnovations.com/documents/pdf/2009/sessions/T-1009.pdf

Mayer PW, DeOreo WB, Opitz EM, et al. Residential end uses of water. AWWA Research Foundation and American Water Works Association Denver, CO, 1999.

Mazzanti M and Montini A. The determinants of residential water demand: Empirical evidence for a panel of Italian municipalities. Applied Economics Letters. 2006; 13: 107–111.

Ouyang Y, Wentz E, Ruddell B and Harlan S. A multi-scale analysis of singlefamily residential water use in the Phoenix metropolitan area. Journal of the American Water Resources Association. 2013; 1–20.

Papakostas KT, Papageorgiou NE and Sotiropoulos BA. Residential hot water use patterns in

Greece. *Solar Energy*. 1995; 54: 369-374.

Pecan Street, Inc. Pecan Street's Dataport. https://dataport.pecanstreet.org/

Polebitski AS, Palmer RN and Waddell P. Evaluating water demands under climate change and transitions in the urban environment. *Journal of Water Resources Planning and Management*. 2011; 137(3): 249-257.

Praskievicz S and Chang H. Identifying the relationships between urban water consumption and weather variables in Seoul, Korea. *Physical Geography*. 2009; 30: 324-337.

Richter CP. Automatic dishwashers: Efficient machines or less efficient consumer habits? International Journal of Consumer Studies. 2010; 34: 228–234.

Rosenberg DE, Madani K. Water resources systems analysis: A bright past and a challenging but promising future. *Journal of Water Resources Planning and Management*, 2014, 140(4): 407-409.

Sanders R. The Pareto principle: its use and abuse. Journal of Services Marketing, 1987, 1(2): 37-40.

Schleich J and Hillenbrand T. Determinants of residential water demand in Germany. Ecological Economics. 2009; 68: 1756–1769.

Suero FJ, Mayer PW and Rosenberg DE. Estimating and verifying United States households' potential to conserve water. *Journal of Water Resources Planning and Management*, 2012, 138(3): 299-306.

Tinker A, Bame S, Burt Rand Speed M. Impact of non-behavioral fixed effects on water use: Weather and economic construction differences on residential water use in Austin, Texas. Electronic Green Journal. 20015; 1(22).

U.S. Geological Survey. State drought information. U.S. Department of the Interior, 2015.

Retrieved from: http://waterwatch.usgs.gov/index.php?r=us&m=dryw.

Wentz EA, Wills AJ, Kim WK, Myint SW, Gober P and Balling Jr RC. Factors influencing water consumption in multifamily housing in Tempe, Arizona. *The Professional Geographer*. 2014.; 66: 501-510.

WHO, UNICEF. Progress on sanitation and drinking water: 2015 update and MDG assessment. Geneva: World Health Organization, 2015. Retrieved from: http://apps.who.int/iris/bitstream/10665/177752/1/9789241509145_eng.pdf?ua=1

Willis RM, Stewart RA, Panuwatwanich K, Jones S and Kyriakides A. Alarming visual display monitors affecting shower end use water and energy conservation in Australian residential households. *Resources, Conservation and Recycling*. 2010; 54: 1117-1127.

Yan D, O'Brien W, Hong T, Feng X, Gunay B, Tahmasebi F, Mahdavi A. Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy and Buildings*, 2015; 107: 264-278.

Zar JH. Significance testing of the Spearman rank correlation coefficient. *Journal of the American Statistical Association*, 1972, 67(339): 578-580.

Zhang HH and Brown DF. Understanding urban residential water use in Beijing and Tianjin, China. *Habitat International*. 2005; 29: 469-491.

Tables

Table 1 Summary of household data

Objectives	Number of houses	Logging days with valid DWU	Notes
Pre-processed data	50	10659	These households have valid data without “0” and ”1” values
Normalization by house floor area	43	10182	These households have the information of house area
Normalization by capita	40	8949	These households have the information of number of residents
Relation between water use and energy use	4	993	These households have energy consumption data of appliances and DWU data at the same time

Table 2 Spearman correlation coefficients of DWU with week days, seasons and OAT

Age groups	Week days	Seasons	OAT
DWU per house	0.400**	0.173**	-0.158*

** Correlation significant at the 0.01 level (two-tailed).
* Correlation significant at the 0.05 level (two-tailed).

Table 3 Personal income levels

Personal income levels	1	2	3	4	5	6	7
Income *per capita* ($)	10k~15k	15k~25k	25k~35k	35k~50k	50k~75k	75k~150k	>150K

Table 4 Spearman correlation coefficients of DWU per house and age groups

Age groups	≤5	6~12	13~18	19~24	25~34	35~49	50~64	≥65
DWU per house	0.284	0.098	-0.187	-0.117	-0.090	0.140	0.113	0.011

* Correlation significant at the 0.05 level (two-tailed).

Table 5 Spearman correlation coefficients of DWU and DEU of appliances

	DEU of bathroom	DEU of clothes washer	DEU of dish washer	DEU (total of 3 items)
DWU	-0.012	0.545**	0.317*	0.463**

** Correlation significant at the 0.01 level (two-tailed).
* Correlation significant at the 0.05 level (two-tailed).

Table 6 Coefficients of regression

Model 3	Standardized	*t*	Sig.
	Beta		
DEU of clothes washer	.711	11.701	.000
Weekdays	.166	2.739	.007
DEU of dish washer	.126	2.372	.009