


# Water Event Categorization Using Sub-Metered Water and Coincident Electricity Data

J. Scott Vitter <sup>1,\*</sup>  and Michael Webber <sup>2</sup>

<sup>1</sup> Department of Mechanical Engineering, The University of Texas at Austin, 204 East Dean Keeton Street Stop 2200, Austin, TX 78712, USA

<sup>2</sup> Energy Institute, The University of Texas at Austin, 2304 Whitis Ave Stop C2400, Austin, TX 78712, USA; webber@mail.utexas.edu

\* Correspondence: scott.vitter@utexas.edu; Tel.: +1-979-587-9904

Received: 18 April 2018; Accepted: 26 May 2018; Published: 31 May 2018



**Abstract:** This study evaluated the potential for data from dedicated water sub-meters and circuit-level electricity gauges to support accurate water end-use disaggregation tools. A supervised learning algorithm was trained to categorize end-use events from an existing database consisting of features related to whole-home and hot water use. Additional features were defined based on dedicated irrigation metering and circuit-level electricity gauges on major water appliances. Support vector machine classifiers were trained and tested on portions of the database using multiple feature combinations, and then externally validated on water event data collected under dissimilar conditions from a demonstration house in Austin, Texas, USA. On the testing data, a trained classifier achieved true positive rates for occurrences and volume exceeding 95% for most categories and 93% for toilet events. Performance for faucet events was less than 90%. Initial results suggest that dedicated sub-meters and circuit-level electricity gauges can facilitate highly accurate categorization with simple features that do not rely on flow rate gradients.

**Keywords:** water end uses; water event classification; high frequency smart metering data; residential water flow trace disaggregation; water flow trace filtering; supervised learning

## 1. Introduction

Water end-use disaggregation has emerged as a promising tool for urban water demand management, accompanied by increasing adoption of smart water meters among water utilities [1]. The goal of water end-use disaggregation is to contextualize water use data by providing information about activities and fixture types related to water use, primarily in the residential sector [2]. Contextualized information about residential water end uses has a variety of applications for consumers, water utilities, and policy makers [2]. For example, appliance-specific information can help end users by improving perception accuracy of water use [3] while identifying pathways to meet efficiency goals or reduce water bills [4]. For utilities, end-use disaggregation has potential to improve day-to-day operations [5], for example by assisting the resolution of billing disputes, facilitating improved pump scheduling to achieve greater system efficiency [6], or reducing uncertainty associated with making long-term planning decisions [2]. For policy-makers, disaggregated water use information can enable efficiency programs for both energy and water by targeting specific appliances within high-usage homes [7,8], or by providing a tool for digital multi-service utility providers to better understand their networks [9].

Prior works on water end-use disaggregation have focused on methods incorporating data for whole-home water flow rate [10–14], water pressure [15–17], or from microphones [18]. Other methods have taken a multi-modal sensing approach [19–21]. These approaches are thoroughly summarized in

recent reviews [1,2]. Although methods related to flow rate differ, they typically focus on characterizing events based on whole-home water data only [1]. With only whole-home data available, flow rate based methods have incorporated human oversight [10] or very high temporal resolution ( $\leq 5$  s) data combined with measurements of flow rate gradients [14] in pursuit of greater accuracy. Other approaches have deployed multiple sensors throughout the home to gather more data to improve classification accuracy [17,20,21]. Depending upon the approach, algorithms are often limited by requirements for human oversight, requirements for data sampling, or requirements for intrusive metering [2].

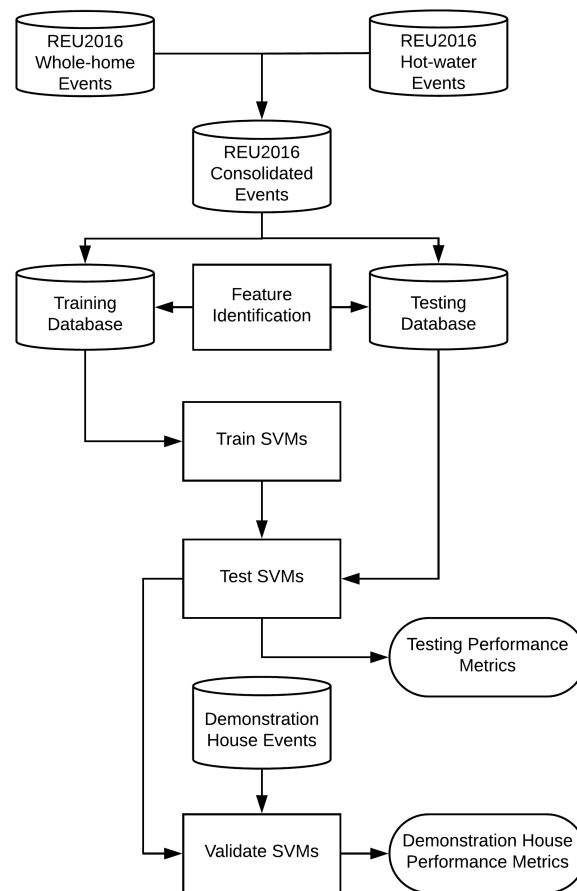
Recently, new approaches have emerged that combine electricity and water data by applying outputs from Non-Intrusive Load Monitoring (NILM) [22,23] or circuit-level electricity use monitoring to the water event categorization problem [24–26]. These approaches hypothesize that information about household electricity use also provides information about water use [25]. For example, electricity consumption by a clothes washing machine provides an indication that the appliance has recently consumed water [25], while an electricity consumption signal from a water heater indicates that hot water use is occurring or recently occurred [24]. However, if the sampling interval increases or gradient-based features are not defined, classification accuracies on events from non-electromechanical appliances are still limited, even with circuit-level electricity data [25].

When dedicated water sub-meters are already installed in the house, for example on the water heater inlet or irrigation branch line, a new approach could extend previous work by incorporating new sources of data to improve the performance of flow rate based disaggregation methods. Using whole-home data, past work identified that fine sampling resolutions (i.e., 1-s or 5-s resolution) are needed to achieve high disaggregation accuracy [27]. This resolution allows the definition of features based on flow rate gradients such as the initial gradient-rise or final gradient-fall for a water event [14]. Adding new streams of sub-metered water data could challenge this paradigm by substituting new information in place of finer resolution, supporting the use of simpler features that are not based on flow rate gradients and are more readily defined for intermediate sampling intervals (i.e., 1-min). Relying on water meters, a well-established technology, could mitigate concerns about intrusiveness [2].

Dedicated metering on water lines has been previously proposed as a tool for water demand management. For example, dedicated irrigation metering is a common practice in residential and commercial sectors [28]. For water utilities, a well-planned dedicated irrigation metering program can facilitate novel pricing strategies that distinguish between indoor and outdoor water use, enable drought response, and improve system modeling [28]. For customers, dedicated irrigation meters remove uncertainty associated with wastewater billing, allowing customers to avoid being penalized for outdoor water use that does not flow through the wastewater system [29]. Dedicated meters have also been proposed for water heaters. For example, the Residential End Uses of Water Study Version 2 (REU2016) included dedicated water heater metering to quantify per-appliance usage and to improve estimations of water-related energy consumption [30].

The authors are unaware of existing water categorization tools that incorporate dedicated sub-meter water data to make event classifications against a known dataset. This analysis evaluates the potential for dedicated water heater and irrigation meters to improve categorization accuracy across all categories without using gradient-based features. The results present accuracy in the context of number of occurrences and cumulative volume. The REU2016 database provides underlying whole-home and hot water data for classifier training and testing. Additional event features are defined to indicate the presence of circuit-level electricity data and simulate data logging from dedicated irrigation meters. Support vector machine (SVM) [31] classifiers are trained and tested on equally-sized portions of the database for up to ten input features, quantifying trade-offs between adding additional metering and improving categorization accuracy. Trained classifiers then make predictions against an external catalog of known water events for a demonstration home with dedicated hot water and irrigation metering in Austin, Texas, USA. By categorizing events from a demonstration house, this study provides a first test

of the general applicability of the trained classifiers on data collected from dissimilar environments. The components of the study are depicted by the flowchart in Figure 1.



**Figure 1.** After datasets are curated and assembled, support vector machines (SVM) are used to train classifiers and test their performance. Following testing, SVMs are validated on data from a demonstration home that were collected under different conditions relative to the training set.

This paper introduces a novel approach to water end-use disaggregation by combining features that can be derived from dedicated sub-meters, a whole-home water meter, and coincident circuit-level electricity data. The results of the study indicate the potential for dedicated sub-meters to support classifier performance using simple input features that are less sensitive to temporal resolution [25]. This analysis can support normative recommendations for household metering layout, meter sampling requirements, and future applications of smart-meter data.

## 2. Materials and Methods

This section describes input datasets, selection of event features, choice of machine learning algorithm, and performance metrics used to evaluate the accuracy of event disaggregation.

### 2.1. Dataset

This study incorporated data from two sources. Data from the REU2016 study were used for classifier training and testing. To conduct a first test of classifier applicability outside of the testing

data, separate data from a demonstration home in Austin, Texas, USA with dedicated sub-meters were used for initial validation.

## REU2016

The Residential End Uses of Water Study Version 2, completed in 2016, collected high-frequency end-use data from 762 homes across nine cities in the United States or Canada. The study updated an earlier study from 1999 [32], providing a snapshot of trends and patterns in residential water usage in North America over the past two decades. A subset of 94 homes were selected for hot water use monitoring [30]. The database of hot water events included faucet, clothes washer, dishwasher events, shower, and bathtub events. Leak events were excluded because they are not attributable to a particular appliance category and can already be identified by commercial leak-detection products [25,33]. Hot water events were merged with whole-home events on a per-house basis by comparing event timing and categories to achieve a dataset with approximately 1000 events per category. Internal clock times of data loggers that recorded whole-home and hot water use were not perfectly synchronized for all houses and required manual curation to resolve discrepancies on a house-by-house basis. Events were discarded if timing discrepancies could not be resolved. The resulting dataset consisted of 1584 clothes washer events, 1014 dishwasher events, 1312 shower events, 77 bathtub events, and 1032 faucet events. Bath events were grouped in the same category as shower events because they are relatively scarce and share a common purpose.

Irrigation and toilet events, which do not consume hot water, were added to the merged database by sampling without replacement from REU2016 data. Toilet events were limited to high efficiency events ( $\leq 2$  gal) to tailor classifiers for homes built after U.S. Energy Policy Act of 1992 standards were enacted [34]. Equally-sized training and testing datasets were formed by sampling 500 events from each category without replacement.

## 2.2. Demonstration Home

Additional water data were collected from a single-family house in Austin, Texas, USA with dedicated meters on the water heater inlet and irrigation branch line. The house had four occupants and consisted of 5 bedrooms, 4 toilets, 3 showers, outdoor space, and a tankless natural gas water heater. The meters are equipped with a BluBand smart register that samples water use once every 7 s. The BluBand, developed by Pecan Street, Inc., attaches to the existing residential meter and detects wobbles by the nutating disk [35]. The logged data were transmitted to a *BluCube*, a central gateway device that uploads data to Pecan Street's data center over the residential wireless network [35].

Residents manually recorded their water use over approximately three weeks, noting appliance type, timing, and duration for each event. Additional irrigation events that occurred outside the three-week period were also included. A validation dataset was formed from a log of known water events by processing underlying flow-trace data from the whole-home meter and dedicated sub-meters. The validation dataset includes 371 events, consisting of 70 events from each category except for the irrigation category, which consisted of 21 events. Note that clothes washer and dishwasher cycles consist of several single events during the overall appliance runs.

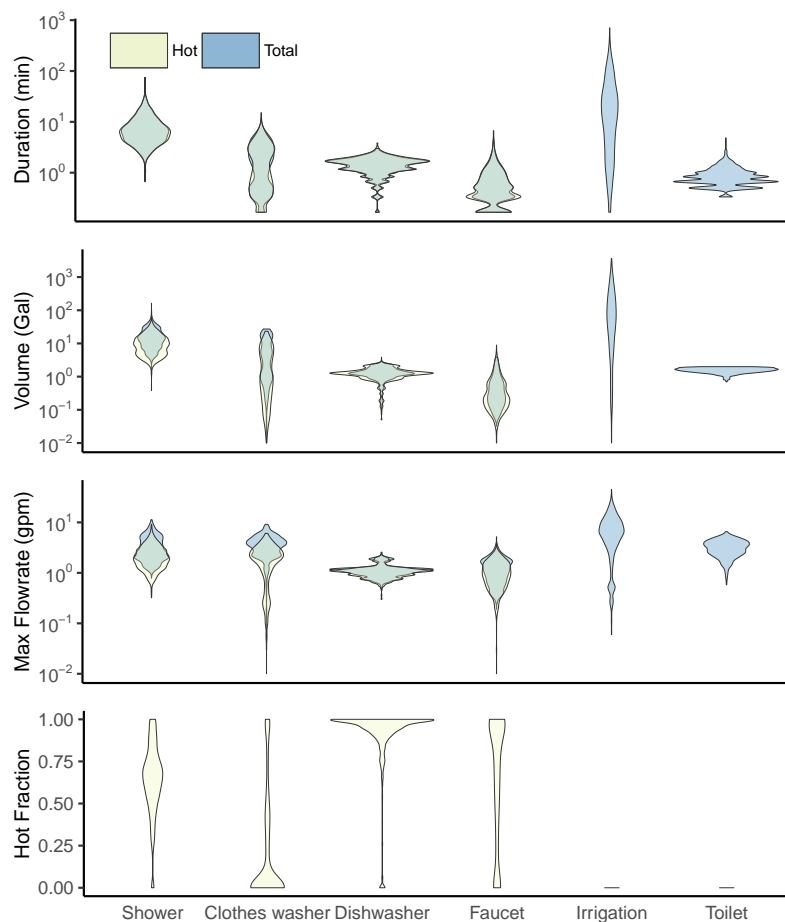
The demonstration home is also equipped with an *eGauge* device for measuring household electricity use [36]. Circuit-level data were not available for the clothes washer or dishwasher circuit. Instead, sub-circuit electricity features were defined based on their proximity to known events [25]. In future work, previously developed NILM algorithms should be implemented to detect electricity consumption by major appliances.

Automatically separating combined water events was not a component of this study. However, recent work introduced an advanced two-step filter that combines an Elitist Non-Denominated Sorting Genetic Algorithm NSCA-II, followed by a cropping algorithm [37]. This approach was tested on data from two dissimilar case studies to establish general applicability. Other work has shown that combined events can be accurately separated into single events using gradient vector filtering

techniques [14]. For demonstration house data, logged combined events were manually separated by following a similar procedure. First, the whole-home flow rate of the base event was recorded before and after the secondary event. The base event was assumed to continue to use water during the secondary event. After usage was assigned to either the base event or a secondary event, features were defined based on the separated profiles. The same procedure was used to separate combined hot water events into base events and secondary events. Once separated, whole-home and hot water events were linked according to their time stamp. Manual disaggregation of combined events was part of early water event categorization methods [10] and is relatively simple to execute, however, it is expensive to adopt at scale due to human labor requirements [14]. As discussed in Section 3.4, automating the combined event disaggregation process and demonstrating its use on hot water events is an important element of future work needed to scale the method to a real-world application.

### 2.3. Feature Selection

To help understand the distribution of water events in the REUWS dataset by category, violin plots of whole-home and hot water volume ( $V_{tot}$ ,  $V_{hot}$ ), duration ( $D_{tot}$ ,  $D_{hot}$ ), and maximum flow rate ( $M_{tot}$ ,  $M_{hot}$ ) for each event are shown in Figure 2 on a log-10 scale. Hot-water fraction, defined by dividing hot water volume by total water volume for each event, is plotted on a linear scale.



**Figure 2.** Distribution of water event features on the basis of duration, volume, and maximum flow rate for total and hot water use. Hot-water fraction is also plotted. Distributions for hot water features have similar shapes relative to total water use features, although natural groupings between appliances emerge on the basis of hot-water fraction.

Hot water use is only a fraction of total water use for a given event. In general, hot and cold water consumption occur together, with the relative share of each determined by the user or by the appliance. This behavior is reflected in the violin plots for hot water volume and maximum flow rate, which are slightly closer to zero in Figure 2 relative to plots for total volume and total maximum flow rate (excepting toilet and irrigation events, which do not consume hot water). However, distributions for hot water event duration closely match total water event duration. Overall, there is little distinction between hot water versus total water events on the basis of aggregate distributions for event volume, duration, and maximum flow rate. The lack of differentiation suggests using directly measured hot water features might not introduce new useful information into a classification algorithm. However, natural groupings emerge for indoor water events on the basis of hot-water fraction, particularly between clothes washer, shower, and dishwasher categories. According to Figure 2, shower events tend to use 50–75% hot water, clothes washer events typically use less than 25% hot water, and dishwasher events most commonly use only hot water.

Because the REU2016 study did not include dedicated irrigation meters, an irrigation fraction feature ( $F_{irr}$ ) was defined as the expected fraction of irrigation volume divided by total water volume for each event. For irrigation events,  $F_{irr} = 1$ . For non-irrigation events, small non-zero values were assigned to  $F_{irr}$  to prevent classification algorithms from incorrectly making a binary distinction between zero and non-zero values of  $F_{irr}$ . Such a distinction would be inappropriate when working with empirical data that could include small irrigation leaks or irrigation meter measurement error during non-irrigation events. A similar approach was taken when defining  $F_{hot}$  for toilet and irrigation events.

Similar to a preceding study, an important aspect of this work was incorporating simultaneous circuit-level electricity data to improve the classification accuracy across all categories of water events [25]. When electricity data are available, information related to electricity use can help clarify drivers of water use. For example, data for when clothes washers or dishwashers consume electricity provide information that can be used to differentiate between events with similar water-related features. Similarly, data for electric water heater operation in conjunction with water use data could help differentiate shower events that use large amounts of hot water from other event types. In past work, NILM techniques have been developed to identify electricity appliance signatures for electric smart meter data, including clothes washer, dishwasher, and water heater events [23]. When circuit-level electricity data are recorded, major appliances are often allocated their own circuit, making appliance events trivial to differentiate from underlying load profiles [35]. However, circuit-level electricity data for major appliances does not completely simplify water event categorization. For example, shower events might commonly occur before, during, or after clothes washer events; whereas several faucet events might occur while the dishwasher consumes electricity.

In previous work, the feasibility and usefulness of binary event flags were evaluated as input features for water end-use categorization [25]. These features,  $E_{cw}$  and  $E_{dw}$ , were introduced and defined for each water event in the present study. The value of each flag indicates whether a clothes washer or dishwasher event would be observed following a water event. For each house, water events that occur within a specified timespan,  $dt_{cw}$ , after a clothes washer event have an assigned value of  $E_{cw} = 1$ . Otherwise,  $E_{cw} = 0$ . Similar rules apply for setting  $E_{dw}$ . The values of  $dt_{cw}$  and  $dt_{dw}$  were set equal to 20 and 30 min, respectively, to cover the expected duration of clothes washer and dishwasher activity observed in a previous study [25]. Ongoing work is evaluating simple threshold techniques for setting clothes washer and dishwasher related flags on circuit-level data that are available from 350+ homes in the Pecan Street database. In future work, NILM techniques should be adopted to automate clothes washer and dishwasher event detection for households with only whole-home electricity data. An example of  $E_{cw}$  and  $E_{dw}$  assignment for twelve events from a single house in the REUWS dataset is shown in [25]. The 10 features defined for each water event in the REUWS and demonstration house are summarized in Table 1.



**Table 1.** A summary of the 10 input features defined for each event in the merged REU2016 dataset. Combinations of these features were used for classifier training.

Utility	Meter/Circuit	Name	Description
Water	Whole-home	$V_{tot}$	Volume of total event (gal)
		$D_{tot}$	Duration of total event (min)
		$M_{tot}$	Maximum recorded event flow rate (gal/min)
	Hot water	$V_{hot}$	Volume of hot water in event (gal)
		$F_{hot}$	Hot water volume divided by total volume
		$D_{hot}$	Duration of hot water flow during event (min)
		$M_{hot}$	Maximum recorded event hot water flow rate (gal/min)
	Irrigation	$F_{irr}$	Irrigation volume divided by total volume
Electricity	Clothes washer	$E_{cw}$	Binary indicator of circuit-level activity
	Dishwasher	$E_{dw}$	Binary indicator of circuit-level activity

#### 2.4. Training and Testing Support Vector Machine Classifiers

The *Matlab Classification Learner* [38] was used to train multi-group SVM classifiers with 10-fold cross-validation [39–42]. These classifiers were selected because they achieve a balance between training time and accuracy [25]. Other classification techniques such as complex trees exhibited similar performance during testing. SVM classifiers are explained in detail by [31]. A gaussian kernel function [43] was observed to perform well relative to a linear or polynomial kernel function [38], achieving a high degree of specificity while limiting overfitting. During testing, performance remained approximately equal over a range of values for the *KernalScale* parameter, ranging 1.2–2.0. A value of 1.6 was used for training SVM models in this study.

SVM models were developed to make event predictions using up to the 10 input features described in Section 2.3. Training SVM models to handle an increasing numbers of inputs allows for features to be added based on an increasing amount of metering at a given house. For example, only features related to total water use are available for training when only whole-home data are being collected. As circuit-level electricity metering is added,  $E_{cw}$  and  $E_{dw}$  become available for classifier training. Discerning the performance increase from adding new inputs into the classification can help prioritize decisions about adding additional water meters or circuit-level electricity instrumentation. Following training, classifiers were evaluated against the testing dataset.

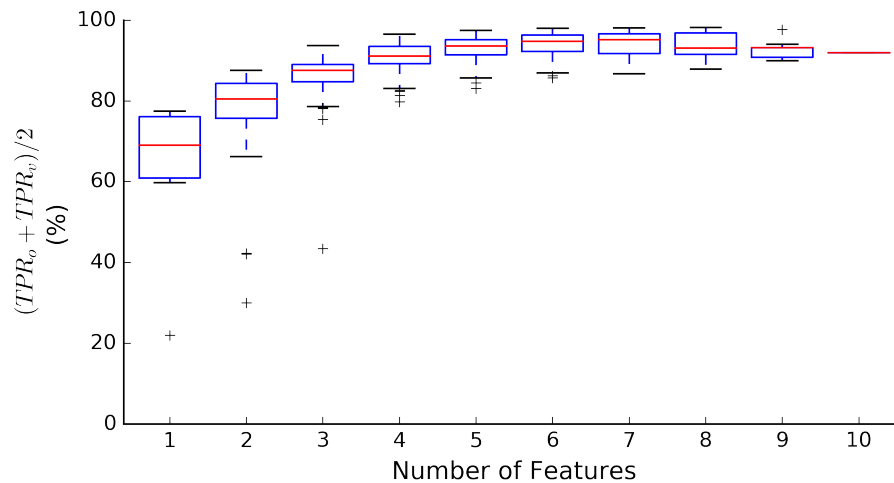
Classification performance is reported in terms of True Positive Rate (TPR), or recall, on the basis of event occurrence ( $TPR_{occ}$ ) and category-specific cumulative volume ( $TPR_{vol}$ ). Recall for a particular category is the fraction of correctly labeled events of the category divided by the total number of events of the category. Results for Positive Predictive Value (PPV), or precision, are also reported. Precision is the fraction of predicted positive cases that are real positive cases for each category [44]. As shown in previous work [25], it is possible to achieve a high level of performance for a particular category on the basis of number of occurrences while simultaneously achieving low performance on the basis of category volume. In practice, it is desirable to achieve high rates for both metrics. Existing literature often focuses primarily on number of events identified without commenting on how much volume was identified correctly [12,14–17,26], although category-specific volume has been discussed in recent work [25,37].

### 3. Results and Discussion

Section 3.1 includes results and discussion of classifier performance on REU2016 testing data as the number of input features made available to the training algorithm is increased. Section 3.2 focuses on a specific six-feature classifier and reports performance in terms of true positive rate and positive predictive value for each appliance category, with additional visualization to communicate common modes of mislabeling. Section 3.3 discusses classifier performance on data from the demonstration home.

### 3.1. SVM Classifier Performance on REUWS Dataset

SVM classifiers were trained and tested on  $n = 1023$  combinations of features, representing all unique combinations of the 1–10 input features presented in Table 1. Results of the testing are shown in Figure 3 in terms of a combined metric that averages overall True Positive Rate in terms of occurrences ( $TPR_o$ ) and volume ( $TPR_v$ ) across all event categories.



**Figure 3.** Ranges of classifier performance on a weighted metric that balances overall True Positive Rates on the basis of occurrences and volume. For a given number of features, the highest point in the column indicates the best performance for all feature combinations. Boxes represent the 1st and 3rd quartile of performance and the red line indicates median performance. The whiskers extend to the highest and lowest value beyond the box that is less than or equal to 1.5 times the interquartile range. There is only one possible combination of the 10-feature classifier.

When only one feature is used by the SVM classifiers, combined performance ranges from 22.0% ( $E_{cw}$ ) to 77.4% ( $F_{hot}$ ). As more features are added, the range of performance shrinks. For a given number of features,  $n$ , the highest point in Figure 3 relates to results by the best performing  $n$ -feature SVM model. The best combined performance overall is achieved by an eight-feature SVM classifier, which includes all features in Table 1 except  $V_{hot}$  and  $D_{hot}$ . Beyond seven features in Figure 3, adding additional features is detrimental to median classifier performance. Because some event features are correlated, such as  $V_{tot}$  and  $M_{tot}$ , features can be removed from some SVM models to increase simplicity with only minimal impact on performance. There are 12 distinct six-feature SVM classifiers that are within 0.5% of the best eight-feature classifier in Figure 3. Within this subset, a specific six-feature SVM classifier was selected for additional testing to reduce model complexity. These results are discussed in Section 3.2. Category-specific performance for each of the 1023 SVM models is included in the Supplementary Materials for this work.

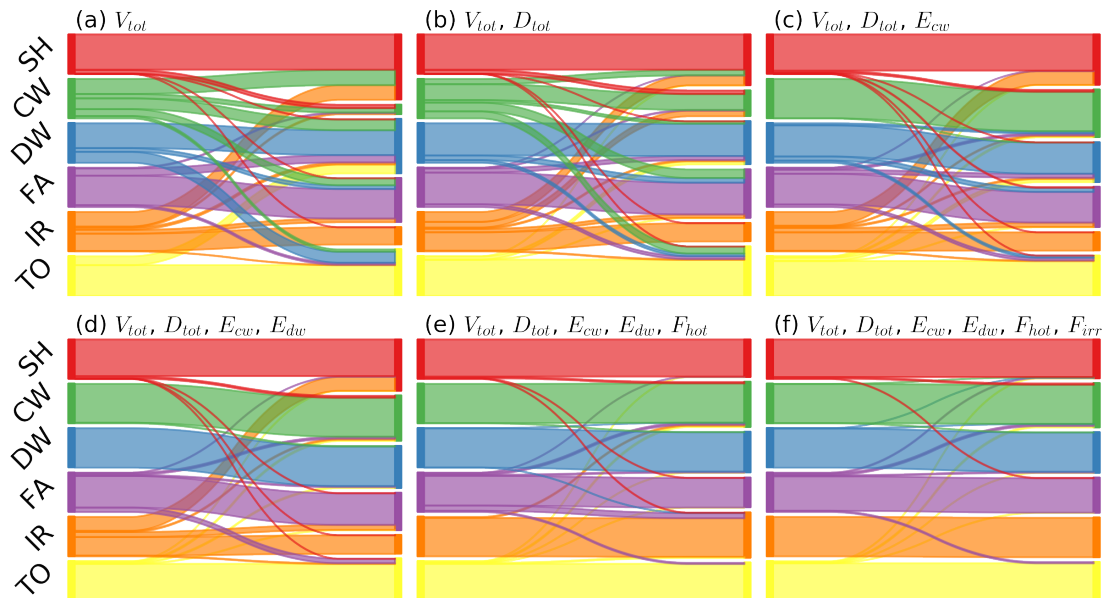
### 3.2. Category-Specific Performance of a Six-Feature SVM Classifier

In the six-feature SVM classifier, features related to the whole-home water meter were incorporated first ( $V_{tot}$  and  $D_{tot}$ ), followed by circuit-level electricity features ( $E_{cw}$  and  $E_{dw}$ ), dedicated hot water meter features ( $F_{hot}$ ), and features related to the dedicated irrigation meter ( $F_{irr}$ ). The testing described in Section 3.1 determined that features related to maximum flow rates ( $M_{tot}$  and  $M_{hot}$ ), hot water volume ( $V_{hot}$ ), or hot water duration ( $D_{hot}$ ) produce limited performance improvement when added as additional features. These features were excluded from additional testing.

Categorization by the six-feature SVM classifier can be visualized by introducing Sankey diagrams, which are commonly used to map material flows between multiple categories [45]. In this



case, the Sankey diagrams are used to map the flow of actual events from each category to their predicted category. These diagrams are introduced by Figure 4 to display appliance-level performance improvement as additional input features are introduced. When used for evaluating classification accuracy, Sankey diagrams use a visually descriptive format to communicate similar information as confusion matrices. For example, the true positive rate for the shower category is the fraction of flow between the actual and predicted shower category divided by the total flow from the actual shower category. The positive predictive value is the fraction of flow between the actual and predicted shower category divided by the total flow into the predicted shower category.

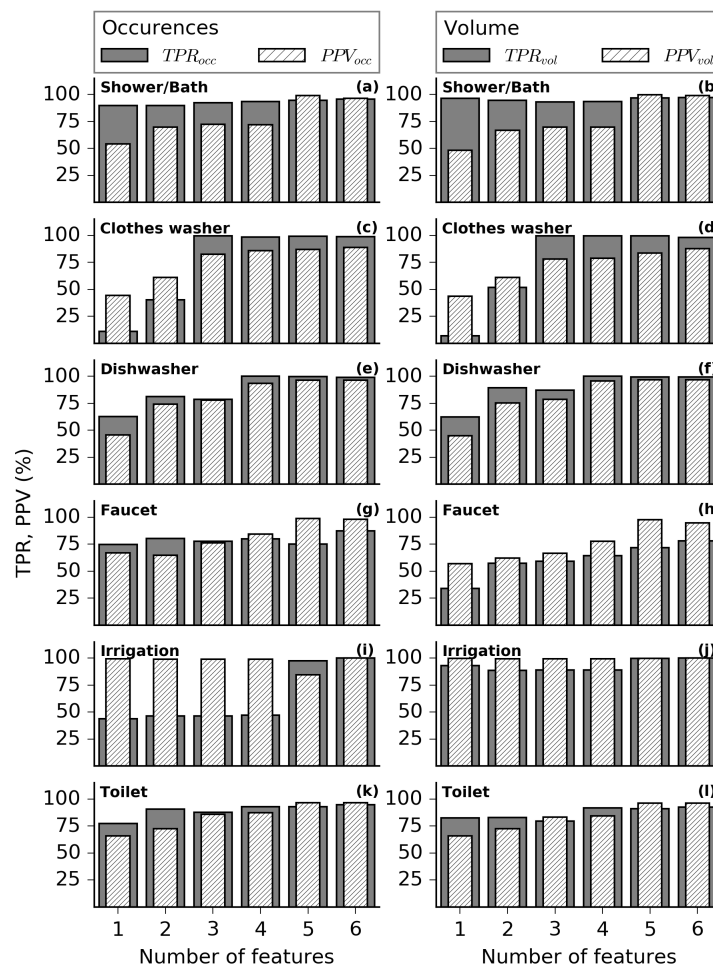


**Figure 4.** Sankey diagram representation of classifier performance on testing data with one to six input features: (a,b) a significant fraction of actual events incorrectly flow to predicted categories for shower, dishwasher, and toilets; (c,d) introduction of  $E_{cw}$  and  $E_{dw}$  more than doubles the flow of correctly predicted clothes washer events; and (e,f) small amounts of incorrectly predicted events are observed when up to six features are included. The severity of remaining misclassifications appears to be mitigated by cancellation error.

In Figure 4a, many events flow between the actual and predicted shower and toilet category. However, both the predicted shower and toilet category receive incorrectly classified events from irrigation, clothes washer, and dishwasher use. In Figure 4c, adding  $E_{cw}$  significantly increases the event flow between the actual and predicted clothes washer category. In Figure 4d, adding  $E_{dw}$  to the classifier has a qualitatively similar effect, although to a lesser magnitude. Classifier performance on the dishwasher category was already relatively high even without  $E_{dw}$  as a feature. In Figure 4e, adding information about hot water fraction ( $F_{hot}$ ) into the classifier more than doubles the true positive rate of irrigation events. Prior to adding  $F_{hot}$ , a portion of short-duration irrigation events resembled shower events on the basis of length, volume, and circuit-level electricity activity. Because shower events use hot water and irrigation events do not, adding  $F_{hot}$  as a feature eliminates those errors. The classification of events in Figure 4f closely resembles Figure 4e, indicating a diminished level of performance improvement from adding a sixth feature. However, performance on the faucet category improves as fewer faucet events consisting of only cold water use are mistaken for irrigation events.

Figure 5 summarizes classifier performance for each category in terms of number of occurrences ( $TPR_{occ}$  and  $PPV_{occ}$ ) and cumulative category volume ( $TPR_{vol}$  and  $PPV_{vol}$ ). The abscissa communicates the number of features used for classification. In Figure 5, achieving 100% represents perfect classifier

performance on a given category. There is a general trend up and to the right within each panel, indicating that classifier performance generally improves when adding up to all six input features.



**Figure 5.** Classifier accuracy generally improves when adding input features, although accuracy gains from one category can come at the expense of another category. For example, adding event duration as the second feature improves (h)  $TPR_{vol}^{faucet}$  while slightly reducing (b)  $TPR_{vol}^{shower}$ .

The six-feature classifier achieves  $TPR_{occ} \geq 95\%$  for all categories except faucets, which is limited to  $TPR_{occ}^{faucet} = 87\%$  (Figure 5g). Clothes washer, dishwasher, and irrigation events are predicted with near certainty. On the basis of volume, toilet events are limited by  $TPR_{vol}^{toilet} = 93\%$  (Figure 5l), with most remaining categories achieving  $TPR_{vol} \geq 97\%$ . However,  $TPR_{vol}^{faucet} = 78\%$  (Figure 5h), suggesting larger faucet events are susceptible to mislabeling. In Figure 4, it appears that some faucet events incorrectly flow to dishwasher or clothes washer categories, suggesting they occur during a washing cycle. Other faucet events are confused with toilet events, suggesting they were cold-water faucet uses of about 1.5 gallons. Fewer faucet events are misclassified as showers, which would likely occur for high-volume events with moderate amounts of hot water use. In general, faucet events are challenging for flow-based methods to account for because they serve a variety of purposes and are ultimately controlled by behavior rather than mechanical settings [10].

A feature set that includes whole-home and sub-meter water data but excludes circuit-level electricity data is also of practical interest. Without circuit-level electricity features, the best performing SVM classifier uses six input features:  $V_{tot}$ ,  $D_{tot}$ ,  $M_{tot}$ ,  $F_{hot}$ ,  $M_{hot}$ , and  $F_{irr}$ . Category-specific results are available in the Supplementary Materials. In general, classifiers without circuit-level electricity

features perform well on most categories but poorly on the clothes washer category in particular, with category-specific True Positive Rates of only about 80% for occurrences and volume. This occurs because clothes washer events in the REU2016 dataset vary widely in terms of whole-home and hot water duration, volume, and maximum flow rate. For example, feature distributions in Figure 2 range from short, small, low-flow rate events to large events that resemble showers. Similarly, although many clothes washer events use primarily cold water, high hot-water fractions are also common.

### 3.3. Classifier Performance on Demonstration House Dataset

Following testing, trained classifiers were used to predict events from the demonstration house dataset with the two-, four-, and six-feature combinations. If algorithms trained by supervised learning can be generally applied, the burdensome process of obtaining ground-truth information for validation might be avoided [2]. The purpose of this testing was to provide an initial test of the general applicability of the trained classifiers on data collected from a single home under dissimilar conditions. In this case study, the geographic location, collection device, and sampling frequency were dissimilar from the training data [30].

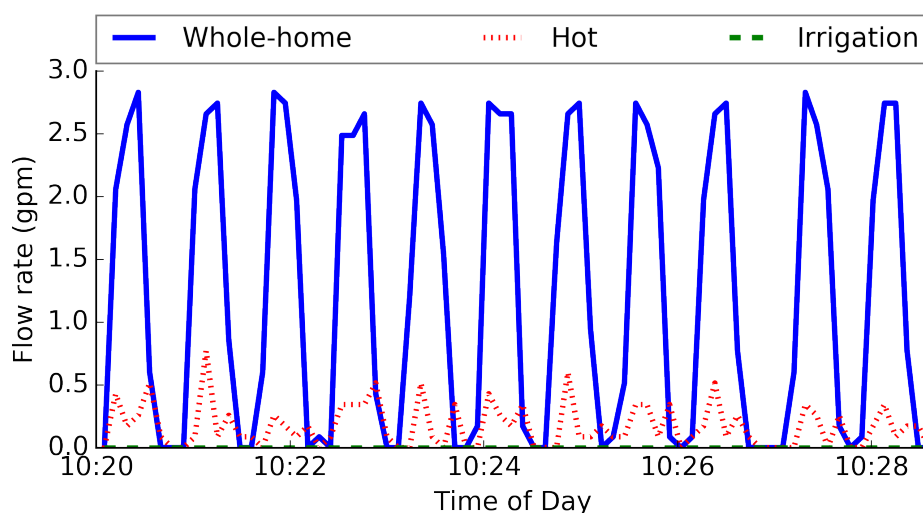
In Table 2, results for  $TPR_{occ}$  and  $TPR_{vol}$  are presented for the demonstration house across all categories. As features are added, improvements in  $TPR_{occ}$  and  $TPR_{vol}$  are similar for each category, with the exception of faucet and irrigation categories. For the faucet category, the true positive rate is typically higher on the basis of event occurrence because small volume events are easiest to classify whereas larger volume faucet events are prone to misclassification. The opposite is true of irrigation events, where identifying a small percentage of large volume events can yield a low occurrence accuracy but a high volume accuracy.

**Table 2.** A summary of TPR classification rates ( $TPR_{occ}/TPR_{vol}$ ) on the demonstration house dataset for an increasing number of input features.

Features	Shower	Clothes Washer	Dish Washer	Faucet	Irrigation	Toilet
$V_{tot}, D_{tot}$	98.6/99.7	12.8/38.3	2.9/9.76	91.4/71.6	35.7/94.9	95.7/95.1
$V_{tot}, D_{tot}$ $E_{cw}, E_{dw}$	100/100	97.1/96.8	100/100	91.4/81.9	28.6/86.8	84.3/84.7
$V_{tot}, D_{tot}, E_{cw}$ $E_{dw}, F_{hot}, F_{irr}$	100/100	97.1/98.8	100/100	90.0/80.0	100/100	87.1/87.9

With all six features, classifier performance on the demonstration house roughly matches the performance observed in Section 3.2. Classifier performance on the toilet category is limited due to unexpected measurements of simultaneous hot water consumption in the demonstration home. An example of this behavior is shown in Figure 6, which shows flow rate data for ten consecutive toilet flushes. Although the events were isolated (i.e., no other events were occurring), non-trivial amounts of hot water use were measured by the *BluBand* on the water heater inlet.

With moderate values of  $F_{hot}$  and active electricity-related flags, several toilet events were misclassified as clothes washer or dishwasher events according to Figure 4. Others were misclassified as shower or faucet events. Because the unexpected behavior is persistent, it appears the measurements of hot water use during toilet events are attributable to the *BluBand* sensor, which potentially detects flow in response to transient pressure drops upstream of the metering device. Notably, the irrigation meter is upstream of all toilet appliances in the demonstration house and did not exhibit similar behavior. This instance exemplifies potential difficulties associated with curating water data for flow-related algorithms and motivates the adoption of data processing or filtering techniques prior to single-event disaggregation [37].



**Figure 6.** Non-trivial amounts of hot water use were measured during isolated toilet events in the demonstration house. This unexpected behavior contributed to slightly diminished classifier performance on the toilet category. This example underscores the importance of robust data curation to facilitate water end-use disaggregation.

### 3.4. Considerations Pertaining to Demonstration Home Results

Although geography, sampling frequency, and collection equipment differed for water measurements from the demonstration home compared to testing data, both datasets were gathered from the North America region (two out of nine REU2016 collection sites are located in Canada [30]). The REU2016 data represent a diversity of climate conditions found on the North American continent; however, future work is needed to determine if the results from this study are valid in other contexts such as Europe or Australia. In addition to climate, user behaviors and end-use fixture characteristics from different contexts could influence how SVM classifiers perform in real-world conditions.

Descriptive statistics published from the South East Queensland Residential End Use Study (SEQREUS) [46] provide an indication for how user behaviors and end-use fixture characteristics from four cities in the Australian context could influence classifier performance on various fixture categories. For example, comparing REU2016 data to SEQREUS data, which is from 2010, reveals that on average shower events in the Australian context are shorter (5.7–6.5 min/event versus 7.8 min/event) and use less water (11.8–14.5 gal/event versus 15.7 gal/event). However, the six-feature SVM model discussed in Section 3.2 still categorizes average Australian showers correctly, assuming typical hot water usage and no clothes washer or dishwasher activity.

Even when end-use fixture characteristics between contexts vary significantly in terms of volume or duration, circuit-level or sub-meter water features can help preserve classifier performance. For example, SEQREUS dishwasher events use 1.1–1.8 gallons on average while REU2016 dishwasher events use 6.1 gallons on average. However, when  $E_{dw} = 1$  and  $F_{hot} \geq 0.95$ , the six-feature classifier correctly classifies even very small volume dishwasher events. The irrigation category has the greatest disparity between the U.S. and Australian context, possibly reflecting increased water restrictions and efficiency attitudes influenced by water scarcity [46]. However, for heavily metered houses with irrigation branch line sub-meters, irrigation events should be easy to identify regardless of geographic context. Without the sub-metered data, there could be an increased frequency of small volume irrigation events that are more likely to be misclassified into a different fixture category.

To more fully address the question of context dependency, future work should evaluate classifier performance on large datasets of end use events gathered across geographic contexts. Making both REU2016 and SEQREUS data freely available could address the North American and Australian context, while ongoing work by researchers at the Politecnico Milano might contribute event data from

the European context [47]. However, to the author's knowledge, only the REU2016 study provides water use information that includes simultaneous hot water consumption, potentially motivating the use of representative data from smaller end-use studies to describe hot water usage behaviors in other contexts [48].

A second, related limitation pertains to the size of the demonstration house dataset, the limited three-week collection window, and its ability to represent variability of end-use categories. For some fixture categories, such as faucet or shower, the four occupants of the demonstration house combined to log a heterogeneous collection of events that differed in terms of duration, volume, hot water usage, and temporal proximity to a clothes washer or dishwasher event. Heterogeneity in flow rate is represented for faucet events due to user behavior, although shower event flow rates are largely consistent in the demonstration home but would exhibit more variation across multiple homes.

For other categories, such as clothes washers and dishwashers, event characteristics are primarily determined by the appliance manufacturer, although users also influence event characteristics. For clothes washer and dishwasher events, appliances from other manufacturers might behave differently in terms of volume, duration, and potentially fraction of hot water use. However, in houses where circuit-level electricity data are measured, values of  $E_{cw}$  and  $E_{dw}$  are strong indicators of appliance usage for a range of volumes and durations.

The demonstration home dataset was not large enough to test for a seasonal impact. Although past work has described limited seasonal variability for indoor water end uses [30,46], outdoor water events are expected to exhibit seasonal variability. Some seasonal variability was incorporated into the REU2016 data, with collection occurring over a three-year period (2010–2013) [30]. However, testing for seasonal influence is a related priority for future work that assembles a larger external dataset for validation.

Based on observed performance in Table 2, the demonstration home appears to be a good match for the six-feature SVM model trained in this study, suggesting the demonstration home is similar to many of the 762 REU2016 homes that provided data used to train the model. However, the trained SVM model might perform poorly for certain types of homes. For example, homes with pre-1992 toilets or urinals might overpredict the faucet category; homes where short showers occur after starting the washing machine might habitually underestimate shower use; and homes that feed cold water to dishwashers with built-in water heaters might confuse dishwasher events and toilet events. Identifying situations that produce feature volatility or classification error, and evaluating their relative frequency, can help more thoroughly evaluate general applicability of the trained SVM model.

The demonstration home is a single-family house, which simplifies challenges associated with differentiating between simultaneously occurring events. As previously mentioned, past work has developed methods to separate single events from combined events [14,37] in single-family houses. Demonstrating combined event disaggregation from a multi-unit housing complex with a common meter would extend the potential reach of end-use categorization techniques. The present study used a manual process to disaggregate combined events from the demonstration home into single events, as described in Section 2.2. This procedure was implemented for both whole-home and sub-metered hot water data. To make the SVM classifiers commercially relevant, future work should test the validity of combined event disaggregation methods on hot water use data and implement a method to automatically link the hot water event with its associated whole-home water use event.

#### 4. Conclusions

This study proposes that dedicated sub-meter water data combined with circuit-level electricity data can improve the accuracy of water end-use classification algorithms using simple input features. In particular, hot water data help distinguish natural groupings among appliance categories. However, the marginal benefit of the  $F_{hot}$  feature is diminished when sub-circuit electricity flags are also present. When added as the fifth and sixth feature to SVM classifiers, dedicated sub-meter data results in  $TPR_{occ} \geq 95\%$  for all categories except faucet. The irrigation fraction feature is especially



useful for distinguishing small to medium cold water events between faucet, toilet, clothes washer, and irrigation categories.

This study also provided an initial test of general model validity by applying SVM classifiers to data gathered from a demonstration home that were collected in dissimilar conditions relative to the REU2016 study. Although the demonstration home data were dissimilar in terms of collection equipment, sampling frequency, and geography within the North American context, the demonstration data do not represent dissimilar user behaviors or fixture characteristics that might exist in non-U.S. contexts. However, even though average fixture-level event characteristics are known to differ between U.S. and non-U.S. contexts, circuit-level and sub-meter event characteristics can preserve categorization accuracy over a range of characteristics.

The real-world application of this work is currently limited by the number of homes with dedicated metering for hot water, irrigation, and appliance sub-circuits. This limitation made it infeasible to collect data from multiple locations for demonstration purposes in this study. However, testing the method on additional homes with sub-metered water data is necessary to demonstrate general applicability outside of the U.S. context. One way to overcome this challenge could involve a partnership with interested electric or water utilities, offering the water event categorization method as a potential application that could stack on top of other smart metering initiatives. Applying the SVM classifiers developed in this study across a larger number of houses with circuit-level and sub-metered data could help clarify the potential for a commercial application of the disaggregation scheme. As affordable hardware solutions come to market, electric and water utilities might increasingly embrace sub-metering to improve demand management, further customer education, or introduce novel electricity or water rate structures.

More work is needed to quantify the trade-offs between sampling rate, volume resolution, and categorization accuracy, which influence the hardware requirements of a hypothetical commercial application. Ultimately, water-related features and NILM techniques should merge to create combined methods for water and electricity event disaggregation. Initial work in this area has focused on using clothes washer or dishwasher activity to inform water event disaggregation. By using NILM to disaggregate water heating, additional usage information might be built into water disaggregation tools [24]. Alternatively, NILM algorithms could potentially improve by introducing water data into their formulation [49].

As reviews of water disaggregation methods have noted, unsupervised or semi-supervised disaggregation methods are a promising direction of future work for several reasons, including their ability to avoid the expensive and time-intensive process of gathering ground-truth information [1,2]. By implementing trained classifiers on external data from a demonstration house, this work shows there is potential for classifiers trained on previously collected data to be applied externally. However, results for the toilet category in Section 3.3 hint at potential limitations associated with heterogeneous site conditions, metering equipment, or data resolutions. In future work, results from supervised learning algorithms might be useful for initializing behavior of a semi-supervised approach.

**Supplementary Materials:** The following are available online at [www.mdpi.com/xxx/s1](http://www.mdpi.com/xxx/s1), Table S1: Database of REUWS water events, Table S2: Dataset of validation water events, Table S3: Spreadsheet of complete SVM model results.

**Acknowledgments:** This project was supported by the Texas Emerging Technology Fund. The authors would like to thank Pecan Street and the Water Research Foundation for providing data used in this study.

**Conflicts of Interest:** The authors declare no conflict of interest.



## References

1. Cominola, A.; Giuliani, M.; Piga, D.; Castelletti, A.; Rizzoli, A.E. Benefits and challenges of using smart meters for advancing residential water demand modeling and management: A review. *Environ. Model. Softw.* **2015**, *72*, 198–214. [[CrossRef](#)]
2. Carboni, D.; Gluhak, A.; McCann, J.; Beach, T. Contextualising Water Use in Residential Settings: A Survey of Non-Intrusive Techniques and Approaches. *Sensors* **2016**, *16*, 738. [[CrossRef](#)] [[PubMed](#)]
3. Attari, S.Z. Perceptions of water use. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5129–5134. [[CrossRef](#)] [[PubMed](#)]
4. Hoekstra, A.Y.; Chapagain, A.K. Water footprints of nations: Water use by people as a function of their consumption pattern. In *Integrated Assessment of Water Resources and Global Change*; Springer: Dordrecht, The Netherlands, 2006; pp. 35–48.
5. Nguyen, K.A.; Stewart, R.A.; Zhang, H.; Sahin, O.; Siriwardene, N. Re-engineering traditional urban water management practices with smart metering and informatics. *Environ. Model. Softw.* **2018**, *101*, 256–267. [[CrossRef](#)]
6. Anda, M.; Brereton, F.I.G.; Brennan, J.; Paskett, E. Smart metering infrastructure for residential water efficiency: Results of a trial in a behavioural change program in Perth, Western Australia. In Proceedings of the First International Conference on Information and Communication Technologies for Sustainability, Zurich, Switzerland, 14–16 February 2013; pp. 116–122.
7. Henderson, H.; Wade, J. *Disaggregating Hot Water Use and Predicting Hot Water Waste in Five Test Homes*; Technical Report; National Renewable Energy Laboratory (NREL): Golden, CO, USA, 2014.
8. Escrivá-bou, A.; Lund, J.R.; Pulido-velazquez, M. Modeling residential water and related energy, carbon footprint and costs in California. *Environ. Sci. Policy* **2015**, *50*, 270–281. [[CrossRef](#)]
9. Stewart, R.A.; Nguyen, K.A.; Beal, C.; Zhang, H.; Stewart, R.A.; Nguyen, K.; Sahin, O.; Bertone, E.; Vieira, A.S.; Castelletti, A.; et al. Integrated intelligent water-energy metering systems and informatics: Visioning a digital multi-utility service provider Briefing paper on water-. *Environ. Model. Softw.* **2018**, *105*, 94–117. [[CrossRef](#)]
10. Deoreo, W.B.; Heaney, J.P.; Mayer, P.W. Flow trace analysis to assess water use. *J. Am. Water Works Assoc.* **1996**, *88*, 79–90. [[CrossRef](#)]
11. Kowalski, M.; Marshall, D. *A System for Improved Assessment of Domestic Water Use Components*; Technical Report; Water Research Center: Swindon, UK, 2003.
12. Nguyen, K.; Zhang, H.; Stewart, R.A. Development of an intelligent model to categorise residential water end use events. *J. Hydro-Environ. Res.* **2013**, *7*, 182–201. [[CrossRef](#)]
13. Nguyen, K.A.; Stewart, R.A.; Zhang, H. An intelligent pattern recognition model to automate the categorisation of residential water end-use events. *Environ. Model. Softw.* **2013**, *47*, 108–127. [[CrossRef](#)]
14. Nguyen, K.A.; Stewart, R.A.; Zhang, H.; Jones, C. Intelligent autonomous system for residential water end use classification: Autoflow. *Appl. Soft Comput. J.* **2015**, *31*, 118–131. [[CrossRef](#)]
15. Froehlich, J.E.; Larson, E.; Campbell, T.; Haggerty, C.; Fogarty, J.; Patel, S.N. HydroSense: Infrastructure-mediated single-point sensing of whole-home water activity. In Proceedings of the 11th International Conference on Ubiquitous Computing, Orlando, FL, USA, 30 September–3 October 2009; pp. 235–244.
16. Larson, E.; Froehlich, J.; Campbell, T.; Haggerty, C.; Atlas, L.; Fogarty, J.; Patel, S.N. Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments. *Pervasive Mob. Comput.* **2010**, *8*, 82–102. [[CrossRef](#)]
17. Froehlich, J.; Larson, E.; Saba, E.; Campbell, T.; Atlas, L.; Fogarty, J.; Patel, S. A Longitudinal Study of Pressure Sensing to Infer Real-World Water Usage Events in the Home. *Pervasive Comput.* **2011**, *6696*, 50–69.
18. Fogarty, J.; Au, C.; Hudson, S.E. Sensing from the basement. In Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology (UIST'06), New York, NY, USA, 3–6 October 2006; ACM Press: New York, NY, USA, 2006; p. 91.
19. Kim, Y.; Schmid, T.; Charbiwala, Z.M.; Friedman, J.; Srivastava, M.B. NAWMS: Nonintrusive Autonomous Water Monitoring System. In Proceedings of the Sixth ACM Workshop on Embedded Networked Sensor Systems (SenSys'08), Raleigh, NC, USA, 5–7 November 2008.
20. Srinivasan, V.; Stankovic, J.; Whitehouse, K. Watersense: Water flow disaggregation using motion sensors. In Proceedings of the Third ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings (BuildSys'11), Seattle, WA, USA, 1–4 November 2011.

21. Ranjan, J.; Griffiths, E.; Whitehouse, K. Discerning electrical and water usage by individuals in homes. In Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings (BuildSys'14), Memphis, TN, USA, 3–6 November 2014; ACM Press: New York, NY, USA, 2014; pp. 20–29.
22. Zeifman, M.; Roth, K. Nonintrusive appliance load monitoring: Review and outlook. *IEEE Trans. Consum. Electron.* **2011**, *57*, 76–84. [[CrossRef](#)]
23. Makonin, S.; Popowich, F. Nonintrusive load monitoring (NILM) performance evaluation A unified approach for accuracy reporting. *Energy Effic.* **2015**, *8*, 809–814. [[CrossRef](#)]
24. Ellert, B.; Makonin, S.; Popowich, F. Appliance Water Disaggregation via Non-Intrusive Load Monitoring (NILM). In Proceedings of the EAI International Conference on Big Data and Analytics for Smart Cities (BigDASC), Toronto, ON, Canada, 13–16 October 2015.
25. Vitter, J.; Webber, M. A non-intrusive approach for classifying residential water events using coincident electricity data. *Environ. Model. Softw.* **2018**, *100*, 302–313. [[CrossRef](#)]
26. Nguyen, K.A.; Stewart, R.A.; Zhang, H. Water End-Use Classification with Contemporaneous Water-Energy Data and Deep Learning Network. *Int. J. Comput. Syst. Eng.* **2018**, *12*, 1–6.
27. Cominola, A.; Giuliani, M.; Castelletti, A.; Rosenberg, D.; Abdallah, A. Implications of data sampling resolution on water use simulation, end-use disaggregation, and demand management. *Environ. Model. Softw.* **2018**, *102*, 199–212. [[CrossRef](#)]
28. Tiger, M.W.; Eskaf, S.; Hughes, J.A. Implications of residential irrigation metering for customers' expenditures and demand. *J. Am. Water Works Assoc.* **2011**, *103*, 30–41. [[CrossRef](#)]
29. *Dedicated Irrigation Meter Introduction*; Alliance for Water Efficiency: Chicago, IL, USA, 2016.
30. Deoreo, W.B.; Mayer, P.W.; Dziegielewski, B.; Kiefer, H.; Sawyer, P. *Residential End Uses of Water*, version 2; Technical Report; Water Research Foundation: Denver, CO, USA, 2016.
31. Cortes, C.; Vapnik, V. Support Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
32. Mayer, P.W.; Deoreo, W.B. *Residential End Uses of Water*; Technical Report; AWWA Research Foundation: Denver, CO, USA, 1999.
33. Hawkins, C.A.; Berthold, T.A. A literature review: developing an information feedback interface to encourage water conservation behavior among utility customers. *Texas Water J.* **2016**, *7*, 40–55.
34. *Wastewater Technology Fact Sheet High-Efficiency Toilets*; Technical Report; United States Environmental Protection Agency: Washington, DC, USA, 2000.
35. McCartney, D. *Water Power: Measuring Household Water and Energy Intensity*; Technical Report; Pecan Street Incorporated: Austin, TX, USA, 2016.
36. *eGauge Model EG30xx Owner's Manual (v1.4)*; Technical Report; eGauge: Boulder, CO, USA, 2016.
37. Pastor-Jabaloyes, L.; Arregui, F.; Cobacho, R. Water End Use Disaggregation Based on Soft Computing Techniques. *Water* **2018**, *10*, 46. [[CrossRef](#)]
38. *MATLAB and Statistics Toolbox Release 2017a*; Mathworks: Natick, MA, USA, 2017.
39. Allwein, E.L.; Schapire, R.E.; Singer, Y. Reducing multiclass to binary: A unifying approach for margin classifiers. *J. Mach. Learn. Res.* **2000**, *1*, 113–141.
40. Fürnkranz, J. Round robin classification. *J. Mach. Learn. Res.* **2002**, *2*, 721–747.
41. Escalera, S.; Pujol, O.; Radeva, P. Separability of ternary codes for sparse designs of error-correcting output codes. *Pattern Recogn. Lett.* **2009**, *30*, 285–297. [[CrossRef](#)]
42. Escalera, S.; Pujol, O.; Radeva, P. On the decoding process in ternary error-correcting output codes. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 120–134. [[CrossRef](#)] [[PubMed](#)]
43. Kecman, V.; Huang, T.M.; Vogt, M. Iterative single data algorithm for training kernel machines from huge data sets: Theory and performance. In *Support Vector Machines: Theory and Applications*; Springer: Berlin, Germany, 2005; pp. 255–274.
44. Powers, D.M.W. *Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation*; Technical Report December; Flinders University of South Australia: Adelaide, Australia, 2007.
45. Nuss, P.; Blengini Gian, A.; Hass, W.; Mayer, A.; Nita, V.; Pennington, D. *Development of a Sankey Diagram of Material Flows in the EU Economy Based on Eurostat Data*; Technical Report; Publications Office of the European Union: Luxembourg, 2017.
46. Beal, C.; Stewart, R. *South East Queensland Residential End Use Study: Final Report*; Technical Report 47; Griffith University: Gold Coast, Australia, 2011.

47. Cominola, A.; Ghetti, A.; Castelletti, A. Building an open high-resolution residential water end-use dataset with non-intrusive metering, intrusive metering, and water use diaries. In Proceedings of the General Assembly 2018 of the European Geosciences Union (EGU), Vienna, Austria, 8–13 April 2018.
48. Kenway, S.J.; Binks, A.; Scheidegger, R.; Bader, H.P.; Pamminger, F.; Lant, P.; Taimre, T. Household analysis identifies water-related energy efficiency opportunities. *Energy Build.* **2016**, *131*, 21–34. [[CrossRef](#)]
49. Piga, D.; Cominola, A.; Giuliani, M.; Castelletti, A.; Rizzoli, A.E. A convex optimization approach for automated water and energy end use disaggregation. In Proceedings of the 36th IAHR World Congress, The Hague, The Netherlands, 28 June–3 July 2015.

**Sample Availability:** The database from the Residential End Uses of Water Study Version 2 is freely available with permission from the Water Research Foundation. The point of contact is Maureen Hodgins, Research Manager, 6666 West Quincy Avenue, Denver, CO 80235-3098, [mhodgins@waterrf.org](mailto:mhodgins@waterrf.org), (303)734-3465. Data from Pecan Street, Inc. is freely available to university researches with permission. The point of contact is Grant Fisher, Water and Data Director, 3925 West Braker Lane, Austin, Texas, USA 78759, [info@pecanstreet.org](mailto:info@pecanstreet.org), (512)782-9213.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).