Extracting household water use event characteristics from rudimentary data

Bettina Elizabeth Meyer, Heinz Erasmus Jacobs and Adeshola Ilemobade

ABSTRACT

Household water end-uses have been extracted from high-resolution smart water meter data in various earlier studies. However, research on end-use disaggregation from rudimentary data is limited. Rudimentary data is defined as data recorded in intervals longer than 1 min, or data recorded with resolutions larger than 0.1 L/pulse. Developing countries typically deal with rudimentary data, due to the high cost and high resource investment associated with high-resolution data. The aim of this study was to extract useful event characteristics from rudimentary data, without identifying the actual end-uses per se. A case study was conducted in the City of Johannesburg, South Africa, where 63 homes were equipped with iPERL smart water meters. The meters recorded flow measurements every 15 s at a 1 L/pulse resolution, rendering the recorded data rudimentary. A total of 1,107,547 event pulses were extracted over the 217-day study period. Although the method presented is limited in the sense that water use events cannot be identified, the method allows for disaggregation of event pulses in the presence of rudimentary data. Using this tool, it is possible to lift valuable information from rudimentary data that would subsequently benefit service providers in setting water demand strategies.

Key words | data resolution, end-use, household water demand, smart meters, water use

Bettina Elizabeth Meyer
Heinz Erasmus Jacobs (corresponding author)
Department of Civil Engineering,
Stellenbosch University,
Private Bag X1, Matieland 7602,
South Africa
E-mail: *hejacobs@sun.ac.za*

Adeshola Ilemobade
School of Civil and Environmental Engineering,
University of the Witwatersrand, Johannesburg,
Private Bag 3, WITS 2050,
South Africa

CHARACTERISTCS OF WATER END-USE EVENTS

End-uses of water, such as the shower, toilet, tap, and washing machine, are considered the building blocks of the residential water demand pattern (Buchberger & Wu 1995). The relationships between an end-use event's characteristics, namely duration, intensity, and volume, create a unique end-use 'fingerprint'. Each end-use event 'fingerprint' is typically represented by a rectangular pulse (Buchberger & Wu 1995; Alcocer-Yamanaka *et al.* 2012). Extracting and identifying end-uses from high-resolution data was pioneered by De Oreo *et al.* (1996), and subsequent investigations include Mayer *et al.* (1999), Loh & Coghlan (2003), Beal *et al.* (2011), DeOreo *et al.* (2011), Beal & Stewart (2013), Arregui (2015), Nguyen *et al.* (2013, 2018), and Pastor-Jaboloyes *et al.* (2018). However, extracting end-use events from rudimentary data sets, and utilising the relationships between event characteristics to categorise extracted end-use events, has yet to be explored.

DATA RESOLUTION

Developing effective demand management strategies requires a clear understanding of household water consumption (Jorgensen *et al.* 2013). Water consumption at a home is typically measured using water meters. Meter readings could be time-based or event-based. In the case of time-based recordings, flow volume through the meter would be averaged over time and recorded at fixed intervals of say 1 s (Buchberger & Wells 1996; Kowalski & Marshallsay 2003), 5 s (Roberts 2005; Beal & Stewart 2013), 10 s

doi: 10.2166/aqua.2020.153

(Mayer et al. 1999; Stewart et al. 2009), 15 min (Pretorius et al. 2019), or 1 h (Cardell-Oliver et al. 2016). Disaggregation of end-uses from a time series requires water end-use data to be collected at a sub-minute resolution (Cominola et al. 2018).

Alternatively, event-based recording involves meter readings taken per water meter pulse – a water meter producing one pulse per litre would not be able to record end-use events smaller than 1 L, for example. Domestic consumer water meters typically used in South Africa, where the case study was undertaken, provide one pulse per litre; the smallest pulse volume commercially available in South Africa at the time of this study was 0.5 L/pulse. At the time of this study, two of the most accurate pulse volumes reported were 0.014 L/pulse (Beal et al. 2011) and 1,000 pulses/L, or 1 mL/pulse (Otaki et al. 2011). To date, the lowest data resolution used for end-use disaggregation and classification was found in a study conducted by Pastor-Jaboloyes et al. (2018), employing volumetric water meters generating a pulse every 0.1 L. Data obtained from meter readings with pulse volumes higher than 0.1 L/pulse (which was the case for this study) were thus deemed rudimentary.

MOTIVATION AND AIM

The required resolution for end-use analysis is not typically available or accessible to service providers in developing countries, due to various constraints (financial, human resources, limited technical expertise, etc.). In order for developing countries to utilise rudimentary data as a vital tool for water demand strategies, a method is needed to classify end-uses into indoor use and outdoor use based on the relationships that exist between basic event characteristics (i.e. duration, intensity, and volume). Before end-use events can be classified, the end-use needs to be extracted from a rudimentary data set. This paper addresses the latter problem. The aim of this study was, therefore, to extract event characteristics from a rudimentary time series data set, without the need to identify the end-use in question. Also, this study set out to develop a procedure that will identify major end-uses from rudimentary data.

STUDY SITE AND CONSUMER SURVEY

The study site was located in Lonehill, a suburb north of Johannesburg, South Africa. The study sample comprised 63 suburban homes, of which nine were stand-alone single family homes and 54 were single-family, semi-detached town houses, located inside a gated community. Gated communities are common in South Africa and earlier studies provide more detail about this relatively high-income dwelling type (Du Plessis & Jacobs 2018).

Following an ethical approval process, the project team embarked on a comprehensive consumer survey and water audit process by visiting selected homes, interviewing selected individuals and distributing survey questionnaires to all homes in the sample. Twenty-one completed survey responses were received, with the team visiting six homes as part of the research process. The average household size for the survey respondents was 1.9 people per household (PPH), with the maximum household size being 4 PPH. Roughly half of the sample reported single-person dwellings, and 29% of the survey respondents reported a household size of 2 PPH.

DATA COLLECTION AND SORTING

Each home in the sample was equipped with a smart water meter, recording the total consumption of each property. In order to identify household end-uses from the recorded flow rate profile, a relatively small volume per pulse and a relatively short time interval would be required. As part of this study, the Sensus iPERL (International) smart water meters were used. The iPERL has integrated bi-directional communications capability and high measurement accuracy.

Data were collected between 5 September 2016 and 29 January 2018 from all 63 homes. The iPERL smart meters measured flow volume to a resolution of 1 L/pulse. The smart meters inbuilt data loggers were programmed to transmit pulse counts at 15 s intervals. Thus, the minimum temporal resolution for recording meter pulses was 15 s. Although the data were metered with a sub-minute resolution, the data set was considered rudimentary due to the relatively large meter pulse volume (1 L/pulse). All water use events smaller than 1 L would thus be reported as part of a larger event, or as part of a set of smaller events, which exceed 1 L when combined. Similarly, events with durations <15 s would be reported at regular intervals of 15 s (not less). Each measurement was reported in terms of the metered volume (≥1 L) and the time stamp (≥15 s), to the nearest 15 s. The recorded values were set to be reported each minute and every 15 s afterwards (00:00:00, 00:00:15, 00:00:30, 00:00:45, and so on).

Ilemobade *et al.* (2018) reported on the complexities of dealing with high-resolution data in the context of a developing country. While the intention with this study was to record only end-use data from the 63 smart meters, in reality, data from various nearby devices (e.g. other household smart meters, security system remotes, and some toys) that were transmitting at the same frequency as the designated smart meters, although unwanted events, were also recorded. The data generation rate (~500 kb/h) led to about 10,000–16,000 records being reported per hour. The raw data were filtered and organised into a format appropriate for analysis, using algorithms developed for the particular purpose. After undertaking several iterations of sorting the data, an algorithm was developed to filter and sort the data, as presented by Ilemobade *et al.* (2018).

After downloading and processing the relevant water use data, the recorded data were sorted chronologically. The final set comprised 63 separate MS Excel files, with each file containing the filtered water meter recordings of a single property. Each MS Excel file contained three fields, namely the unique identifier (meter number), the date-time stamp, and the recorded meter reading (L), from which the pulse volume (L) and intensity (L/s) over the said time interval was deduced. Table 1 summarises the format of the MS Excel files, which were later used as input files for the extraction process.

PROCEDURE FOR EVENT EXTRACTION

A single event was identified by investigating the sequence of measured pulse readings. Event start times (d_0) and event end times (d_e) were derived by evaluating the time difference between recording intervals. If a gap occurred between readings, in other words, if consecutive pulse readings were recorded at intervals larger than the temporal resolution of the meter (15 s), the start/end time of an event was identified. The time passed between measured events was termed a time gap. Figure 1 shows an example of two single events, with a time gap of 45 s. The second event is thus preceded by a 45 s gap.

The event duration (D) was calculated by subtracting d_0 from d_e. The water meter reading difference between two consecutive water meter pulses (Δr) represented the volume consumed between the two pulses. The difference between the event start water meter reading (v_0) and the event end water meter reading (v_e), derived from (Δr), represents the total event volume (V). The average intensity (I) of an identified event was calculated using the total event volume (V) and the event duration (D).

TIME GAP SETTINGS

In some cases, event pulses were lumped despite a delay of ±30 s (preceded by a ±15 s gap), or even ±45 s (preceded by a ±30 s gap). Inspection of the data set confirmed that some lumped readings formed part of a single end-use event. Inconsistencies with recorded meter readings (e.g. lagged meter reading) and data gaps in water meter readings have been reported on in earlier studies (e.g. Cominola *et al.* 2018). Some lagged readings may be superimposed onto the subsequent reading, which is called a lumped reading in this text. Such lumped records would typically be reported

Table 1 | Collected data set format in MS excel

MS Excel field	Field 1 [Column A]	Field 2 [Column B]	Field 3 [Column C]
Data description	Unique identifier (meter number)	Date-time stamp	Recorded meter readings (L)
Data format	1010-001-xxxx	YYYY/MM/DD hh:mm:ss	325 xxx
Variable assigned	–	t	r

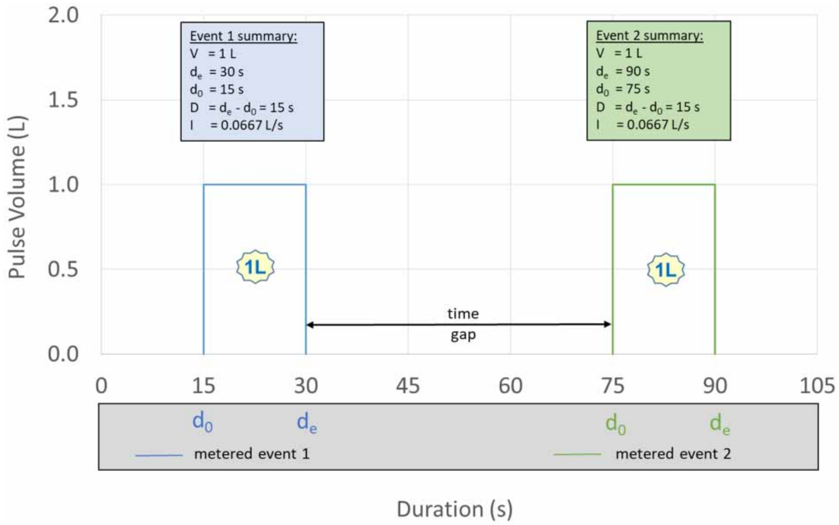

Figure 1 | Two single events with a 45 s time gap.

once during a relatively long water use event, with a relatively constant flow rate. Figure 2 shows a schematic of an interrupted single end-use event (with lumped reading), with the resulting meter spike occurring after a 15 s time gap.

In order to address this problem, a procedure was developed to incorporate lumped values as part of a single event, instead of incorrectly splitting the readings into two or more separate events. Subsequently, a time gap setting (TGS) was incorporated into the extraction tool to determine a suitable time gap between consecutive events. No earlier research was available on which to base an initial time gap estimate. Consequently, time gaps were chosen based on intervals of 15 s. Only three TGS were considered, since a preliminary assessment showed that a TGS >45 s resulted in excessive lumping of end-use events. Thus, the time gaps assessed between separate events were 15, 30, and 45 s.

END-USE EXTRACTION TOOL

A Python End-use Extraction Tool (PEET) was developed as part of this study so that end-use characteristics could be

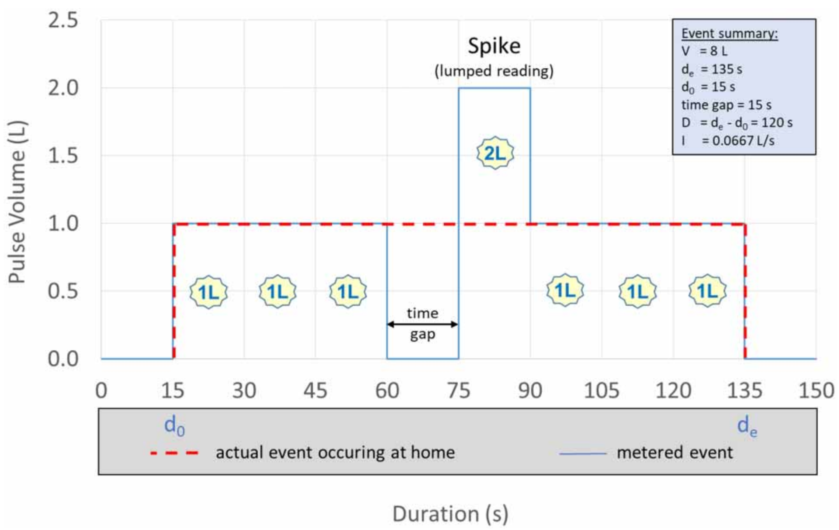

Figure 2 | Schematic of a meter spike/lagged reading.

extracted from the recorded water meter data series. PEET's input and output are MS Excel files. The format of PEET's input is summarised in Table 1. The TGS also had to be defined, in order to determine which consecutive pulse readings must be lumped together to represent a single event. Figure 3 depicts the decision pattern of PEET that extracts end-use events at a single residential property. The resulting PEET output as an MS Excel file contains five fields (see Table 2). Figure 3 shows the schematic decision pattern for one property; thus, the process was repeated for each of the 63 properties in the data set. For each TGS, PEET generated 63 MS Excel files, one file per property. Smart meter serial numbers were used as unique identifiers, in order to link the extracted end-use events to the different homes and the corresponding consumer survey results, which were available for selected homes only. Table 3 summarises all variables defined during this study.

CHARACTERISATION AND IDENTIFICATION OF MINOR EVENTS

Two types of events were categorised during this study, namely minor events and major events. Figure 4 represents a schematic of four low-flow events ($I < 0.035$ L/s) occurring at a home (say a tap being opened and closed).

Due to the limiting 1 L pulse volume, a single event of 1 L in volume was recorded at 105 s on the time series, with a total duration of 15 s. Multiple low-flow events would be reported by the water meter as a single event, at

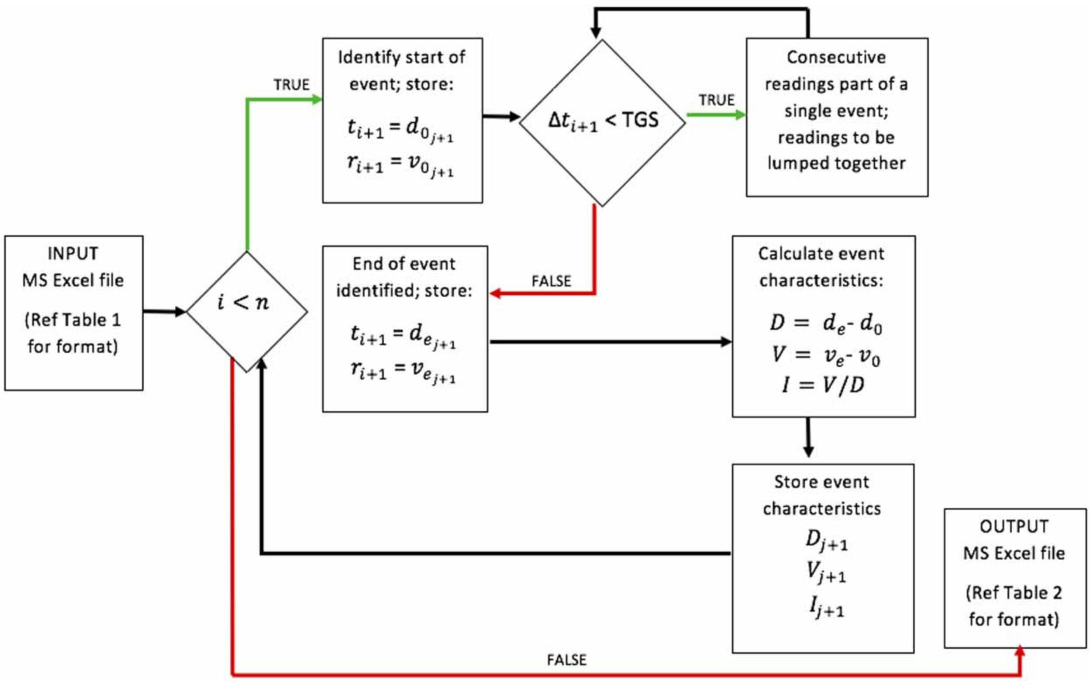

Figure 3 | Schematic of an end-use extraction tool procedure.

Table 2 | MS excel format of PEET output

MS Excel field	Field 1 [Column A]	Field 2 [Column B]	Field 3 [Column C]	Field 4 [Column D]	Field 5 [Column E]
Data description	Date-time stamp	Event volume	Event duration	Event intensity	Unique identifier
Format/(units)	YYYY/MM/DD hh:mm:ss	(L)	(s)	(L/s)	1010-001-xxxx
Variable assigned	t	V	D	I	–

Table 3 | List of variables

Variable	Description
d	Timestamp at start/end of the event
D	Extracted event duration (s)
i	Metered pulse count in time series: $i = 0, 1, 2, \ldots, n$
I	Extracted event intensity (L/s)
j	Event count identified at a home: $j = 0, 1, 2, \ldots, m$
m	Total number of end-uses identified at a home
n	Final pulse reading in MS Excel file
Δr	Volume difference between two consecutive meter readings
Δt	Time difference between two consecutive meter readings
t	Meter pulse reading timestamp
Subscript 0	Start of event
Subscript e	End of event
V	Extracted event volume (L)

a later time. The recorded event is, thus, not a true representation of the actual events occurring at the home. Consequently, all events with a 1 L pulse volume and a 15 s duration (preceded and followed by a delay larger than the time gap) were categorised as minor events and grouped together.

Screening for realistic low-flow events involved assumptions regarding the minimum flow rate of a valid end-use event. Since the flow rate resolution of the meters was 0.067 L/s (1 L pulse over a 15 s recording period), all events with intensities $\leq$0.067 L/s were investigated. Consider an end-use with a constant flow rate of 0.04 L/s being active for a certain period of time (e.g. 75 s) – until (say) the consumer closes the running tap. During the active period, the event would produce one pulse (of 1 L) intermitted at 15 s intervals. Figure 5 represents a schematic of this example.

Due to the TGS incorporated in the extraction process, the three recorded pulses would be lumped together as one single event, with an event volume of 3 L over a duration of 75 s. However, a genuine water use event of 1 L, used in 15 s intervals, would also report one pulse (1 L) at 15 s intervals. The time series of a genuine 1 L event reported over 15 s and that of numerous small events that were reported by the measurement system as 1 L over 15 s, would appear identical. Due to the rudimentary nature of the time series data, the two instances could not be distinguished. Consequently, all extracted events with intensities <0.067 L/s were categorised as minor events and grouped together. Earlier work by Otaki *et al.* (2011) used water meters that were able to measure intensities of 0.0167 L/s (1 L/min), but the

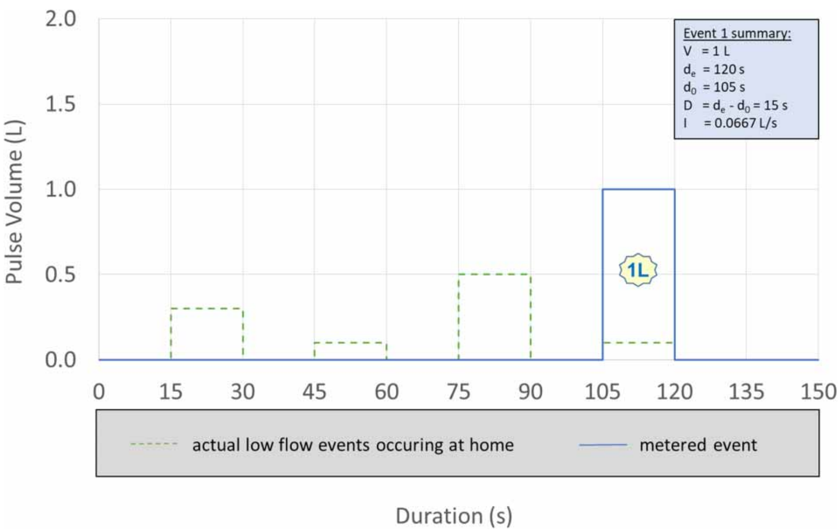

Figure 4 | Schematic of low-flow events.

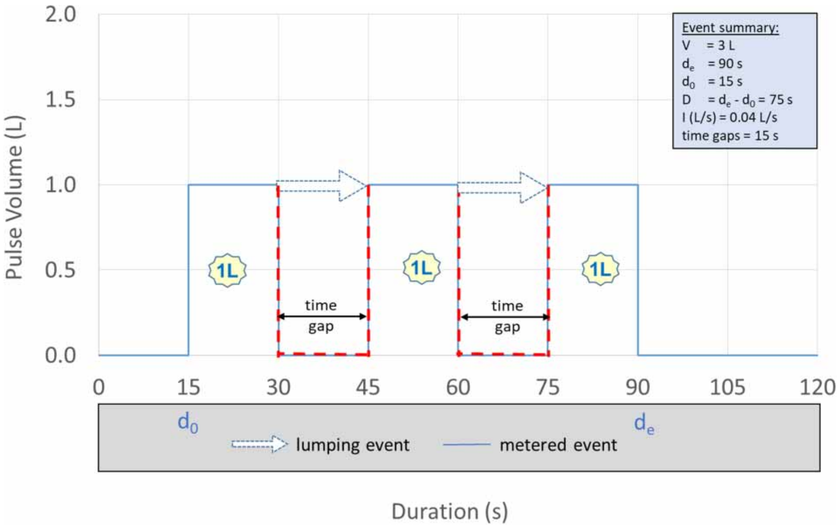

Figure 5 | Schematic example of lumping multiple events.

same authors also note that no in-house activity in the Thailand study area required such a low-flow rate.

READING AND ROUNDING ERRORS

Some extracted events were filtered out of the data set, including all zero values and negative values, which were considered to be reading errors, or rounding errors. Relatively high values could be explained as being either a valid event – possibly spread over a relatively long duration, or a meter reading error. Consequently, a meter verification exercise was conducted to evaluate typical maximum flow rates. The highest flow rate recorded at a single end-use in this study was ~0.4 L/s, but a total flow rate at the consumer meter of ~0.5 L/s was recorded at a home of one of the authors with various taps open simultaneously. Flow rates of >0.5 L/s were reportedly uncommon in Australia, with manual sprinkler systems reporting the highest flow rate of ~0.4 L/s in one study (Roberts 2005). An upper limit of 1.0 L/s was considered appropriate for the study sample, and all readings where the intensity exceeded 1.0 L/s for ≥15 s were filtered out. In other words, events with a total volume difference >15 L in a single recording interval of 15 s were considered to be errors. A summary of all filtered values is presented in Table 4.

RESULTS AND DISCUSSION

Time gap setting

PEET was employed to extract single end-use events from rudimentary data, considering a time gap of 15, 30, and 45 s, before and after a recorded pulse. A comparison of the events extracted, for all three TGS, is tabulated in Table 4. Due to the rudimentary nature of the data, minor events were grouped together. All events not categorised as minor events were considered major events.

As was expected, the 15 s gap setting extracted the most end-use events from the raw data set, with the 45 s gap reporting the lowest numbers. For all three TGS, the total volume of major events comprised >74% of the total volume of all extracted events. Only major events were considered for further analysis. This method was considered acceptable due to the large percentage of total volume representing major events, as well as the uncertainty surrounding minor events.

Characteristics of events

The cumulative distribution functions (CDFs) of event characteristics for the three TGS were compiled and are presented in Figure 6 (event volume), Figure 7 (event duration), and Figure 8 (event intensity).

Table 4 | Comparison between evaluated time gap settings

Description		All extracted events	Reading/rounding errors	Minor events	Major events
TGS = 15 s	# Extracted events	1,288,373	5,377	971,032	311,064
	# Extracted events (%)	100.00	0.42	75.44	24.14
	Total volume (L)	4,429,578	78	971,950	3,457,550
	Total volume (%)	100.00	0.00	21.92	78.06
TGS = 30 s	# Extracted events	1,107,547	5,238	890,249	212,060
	# Extracted events (%)	100.00	0.47	80.38	19.15
	Total volume (L)	4,429,578	86	1,072,735	3,356,757
	Total volume (%)	100.00	0.00	24.22	75.78
TGS = 45 s	# Extracted events	1,022,290	5,177	827,501	189,612
	# Extracted events (%)	100.00	0.51	80.95	18.55
	Total volume (L)	4,596,464	87	1,152,296	3,444,081
	Total volume (%)	100.00	0.00	25.07	74.93

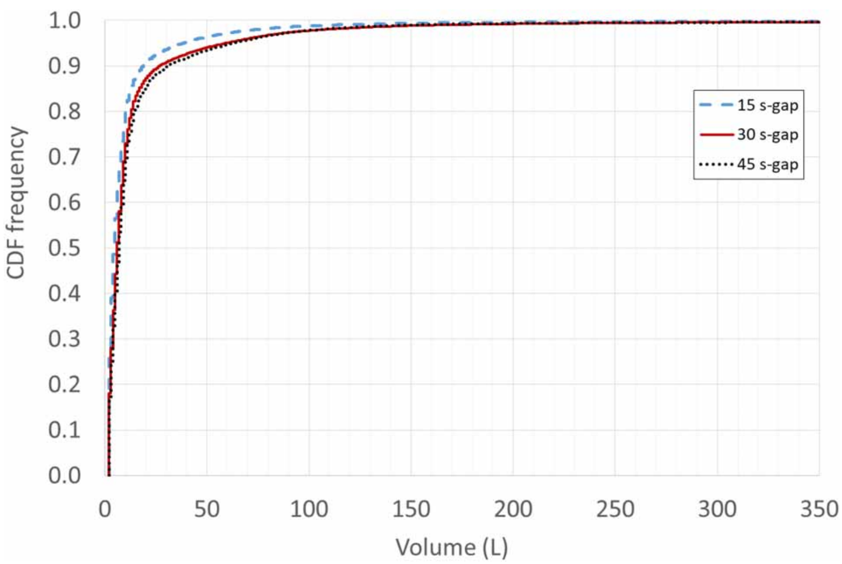

Figure 6 | End-use volume for the three different time gap settings.

Above a given threshold on Figures 6 and 7, the shortest TGS resulted in the lowest event volumes and shortest event durations, as could be expected. In contrast, the shortest TGS produced the highest intensities. The CDF also showed that the lowest 70% of event volume and duration values were almost identical for all TGS. The lowest 70% event volume values were less sensitive to the TGS, compared to the upper 30%. Half of all extracted events had durations of less than 60 s, and event volumes of less than 7 L. This was true for the three different TGS in PEET. The median and most frequent intensity was 0.14 L/s.

Earlier studies (Buchberger & Wells 1996) considered event volume <210 L to be a reasonable limit for classifying indoor events. Approximately 99% of the extracted end-use events for the three TGS had volumes <210 L. Based on the assumed limits, 99% of the events at the 63 homes would thus be considered indoor events – which was unlikely when compared to the survey responses regarding frequency of outdoor irrigation. Simply apportioning end-use events based on arbitrary values is thus not sufficient, and future research should develop a robust method to classify end-uses.

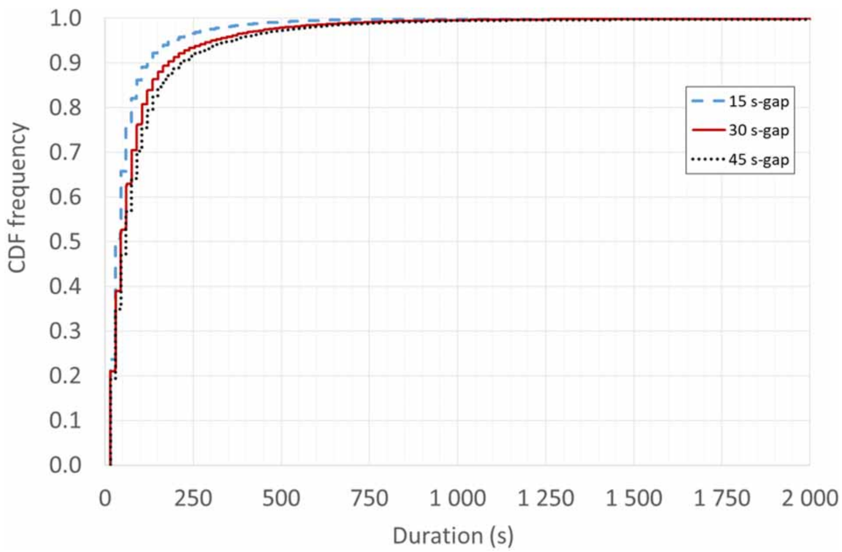

Figure 7 | End-use duration for the three different time gap settings.

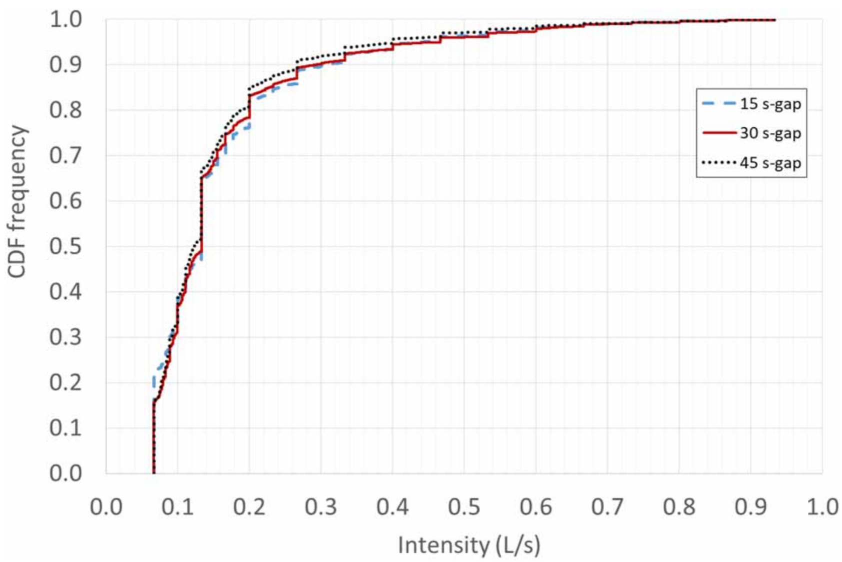

Figure 8 | End-use intensity for the three different time gap settings.

With reference to Figure 6, the largest volume for a single event was 2.6, 3.6, and 4.7 kL for the 15, 30, and 45 s gap settings, respectively. The 45 s TGS reported the longest event of 57,660 s (almost 16 h in duration), while the longest event for the 30 s TGS was 39,136 s. The relatively long durations for the 45 s gap setting were considered excessive, suggesting that the 45 s gap setting may be invalid – in the sense that separate events were combined. The 30 s gap setting showed the most reasonable values for household end-uses when dealing with rudimentary data and was consequently selected as the optimal TGS for this study.

Final data set

Using the 30 s gap setting, a total of 1,107,547 events were extracted from 63 homes over the 217 days, prior to cleaning and filtering the data set. After filtering, the final data set comprised 177,498 single end-use events. About 19% of the total volume of all events was attributed to minor events, representing 76% of the number of events extracted.

The average number of events per home per day was 16 for 1 PPH, 18 for 2 PPH, and 28 for 4 PPH. The number of notable end-uses equates to nine events per person per day, on average over the study period and for all homes. This value was considered realistic, considering that all minor events were filtered out and many homes had a low occupancy of one or two persons.

CONCLUSION

Household water end-use event characteristics were extracted from rudimentary data – in this study, the resolution was 1 L per water meter pulse at a recording interval of 15 s. Various assumptions were employed in the process and three TGS were investigated in attempt to eliminate data lumping problems in the raw data. PEET was developed as part of this study in order to automate the process. PEET was able to extract three water use characteristics, namely event duration, event volume, and event flow intensity, from a rudimentary data set. One of the limitations encountered when dealing with rudimentary data is the fact that minor events had to be grouped together and could not be further analysed. Nonetheless, major end-use events were extracted, and valuable information was deduced from the results. Unfortunately, it was impossible to separately classify background leakage flows in the plumbing system, minor leaks at the point-of-use (e.g. a dripping tap) and relatively low flows from valid water use events (e.g. filling a 200 mL glass with water), so all had to be categorised as minor events. The extracted characteristics of major events could in future be used to classify end-uses as being either indoor events or outdoor events. Such a classification would benefit service providers in setting water demand strategies when faced with rudimentary data.

ACKNOWLEDGEMENTS

The authors would like to thank Johannesburg Water for funding this project, contributing in part to the PhD bursary of Mrs Meyer.

REFERENCES

Alcocer-Yamanaka, V., Tzatchkov, V. & Arreguin-Cortes, F. 2012 Modelling of drinking water distribution networks using stochastic demand. *Water Resour. Manage.* **26** (7), 1779–1792. doi:10.1007.s11269-012-9979-2.

Arregui, F. 2015 New software tool for water end-uses studies. In: *Presented at the 8th IWA International Conference on Water Efficiency and Performance Assessment of Water Services*, 20–24 April 2015, Cincinnati, OH, USA.

Beal, C. & Stewart, R. A. 2013 Identifying residential water end uses underpinning peak day and peak hour demand. *J. Water Resour. Plan. Manag.* **136**, 19–26.

Beal, C., Stewart, R. A., Huang, T. & Rey, E. 2011 South East Queensland residential end use study. *J. Aust. Water Assoc.* **38** (1), 80–84.

Buchberger, S. G. & Wells, G. J. 1996 Intensity, duration, and frequency of residential water demands. *J. Water Resour. Plann. Manage.* **122** (1), 11–19.

Buchberger, S. G. & Wu, L. 1995 A model for instantaneous residential water demands. *J. Hydraul. Eng.* **121** (3), 232–246. *J. Water Resour. Plan. Manag.* 122, 11–19.

Cardell-Oliver, R., Wang, J. & Gigney, H. 2016 Smart meter analytics to pinpoint opportunities for reducing household water use. *J. Water Resour. Plan. Manag.* **142** (6). doi:10.1061/(ASCE)WR.1943-5452.0000634, 04016007.

Cominola, A., Giuliani, M., Castelletti, A., Rosenberg, D. E. & Abdallah, A. M. 2018 Implications of data sampling resolution on water use simulation, end-use disaggregation, and demand management. *Environ. Modell. Softw.* **102**, 199–212.

DeOreo, W. B., Heaney, J. P. & Mayer, P. W. 1996 Flow trace analysis to assess water use. *Am. Water Works Assoc.* **88**, 79–90.

DeOreo, W. B., Mayer, P. W., Martien, L., Hayden, M., Funk, A., Kramer-Duffield, M. & Davis, R. 2011 *California Single Family Water Demand Efficiency Study*. A report compiled for the California Department of Water Resources, 20 July 2011.

Du Plessis, J. L. & Jacobs, H. E. 2018 Analysis of water use by gated communities in South Africa. *Water SA* **44** (1), 130–135.

Ilemobade, A., Jacobs, H. E. & Botha, B. E. 2018 Sorting high resolution end-use data – a pleasurable headache. In: *Proceedings of the 1st International WDSA/CCWI 2018 Joint Conference*, July 23–25, Kingston, Ontario, Canada.

Jorgensen, B. S., Martin, J. F., Pearce, M. W. & Willis, E. M. 2013 Predicting household water consumption with individual-level variables. *Environ. Behav.* **46** (7), 872–897. https://doi.org/10.1177/0013916513482462.

Kowalski, M. & Marshallsay, D. 2003 A system for improved assessment of domestic water use components. In: *II International Conference Efficient Use and Management of Urban Water Supply*, Tenerife, Spain. International Water Association, Spain.

Loh, M. & Coghlan, P. 2003 *Domestic Water Use Study*. Water Corporation, Perth.

Mayer, P. W., DeOreo, W. B., Opitz, E. M., Kiefer, J. C., Davis, W. Y., Dziegielewsji, B. & Nelson, J. O. 1999 *Residential End Uses of Water*. AWWA Research Foundation and American Water Works Association, Denver, CO, USA. http://www.waterrf.org/PublicReportLibrary/RFR90781_1999_241A.pdf.

Nguyen, K. A., Zhang, H. & Stewart, A. R. 2013 Development of an intelligent model to categorise residential water end use events. *J. Hydro-Environ. Res.* **7** (2013), 182–201. https://dx.doi.org/10.1016/j.jher.2013.02.004.

Nguyen, K. A., Stewart, R. A., Zhang, H. & Sahin, O. 2018 An adaptive model for the autonomous monitoring and management of water end use. *Smart Water J.* **3**, 5. doi:10.1186/s40713-018-0012-7.

Otaki, Y., Otaki, M., Sugihara, H., Mathurasa, L., Pengchai, P. & Aramaki, T. 2011 Comparison of residential indoor water consumption patterns in Chiang mai and Khon Kaen, Thailand. *J. AWWA* **103**, 104–110.

Pastor-Jaboloyes, L., Arregui, F. J. & Cobacho, R. 2018 Water end use disaggregation based on soft computing techniques. *Water* **10**, 46.

Pretorius, D., Crouch, M. L. & Jacobs, H. E. 2019 Diurnal water use patterns for low-cost houses with indigent water allocation: a South African case study. *J. Water Sanitation Hyg. for Develop.* **9** (3), 513–521. doi:10.2166/washdev.2019.165.

Roberts, P. 2005 Yarra Valley Water, Residential end use measurement study. Yarra Valley Water, Melbourne, 1-77. June 2005.

Sensus 2018 iPERL™ (International) Water Meters. Available from: https://sensus.com/products/iperl-international/ (accessed 7 June 2018).

Stewart, R. A., Willis, R. & Capati, B. 2009 Closing the loop on water planning: an integrated smart metering and web-based knowledge management system approach. In: *10th IWA International Conference on Instrumentation Control and Automation*, 14–17 June 2009, Cairns, Australia.

First received 11 December 2019; accepted in revised form 24 March 2020. Available online 30 April 2020