Environmental Modelling & Software 102 (2018) 199–212

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Implications of data sampling resolution on water use simulation, end-use disaggregation, and demand management

A. Cominola [a], M. Giuliani [a], A. Castelletti [a, c, *], D.E. Rosenberg [b], A.M. Abdallah [b]

[a] Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milan, Italy
[b] Department of Civil and Environmental Engineering and Utah Water Research Laboratory, Utah State University, Old Main Hill, 4110, UT 84322-4110, Logan, USA
[c] Institute of Environmental Engineering, ETH, Zurich, Switzerland

A R T I C L E I N F O

Article history:
Received 22 May 2017
Received in revised form
22 October 2017
Accepted 9 November 2017

Keywords:
Smart meter
Sampling resolution
Water demand management
STREaM
Synthetic end-use model

A B S T R A C T

Understanding the tradeoff between the information of high-resolution water use data and the costs of smart meters to collect data with sub-minute resolution is crucial to inform smart meter networks. To explore this tradeoff, we first present STREaM, a STochastic Residential water End-use Model that generates synthetic water end-use time series with 10-s and progressively coarser sampling resolutions. Second, we apply a comparative framework to STREaM output and assess the impact of data sampling resolution on end-use disaggregation, post meter leak detection, peak demand estimation, data storage, and meter availability. Our findings show that increased sampling resolution allows more accurate end-use disaggregation, prompt water leakage detection, and accurate and timely estimates of peak demand. Simultaneously, data storage requirements and limited product availability mean most large-scale, commercial smart metering deployments sense data with hourly, daily, or coarser sampling frequencies. Overall, this work provides insights for further research and commercial deployment of smart water meters.

Software availability

Name of software: STREaM - STochastic Residential water End-use Model
Developers: Andrea Cominola, Matteo Giuliani, Andrea Castelletti, David Ezechiel Rosenberg, Adel Abdallah
Version: v1.0 - tested on Matlab 2016a
Contact email: andrea.cominola@polimi.it
Year first available: 2017
Available from: GitHub repository - https://github.com/acominola/STREaM
All other software used for the experiments in this paper available at: https://github.com/acominola/STREaM_Multi-Resolution-Assessment

1. Introduction

Over the last two decades, technological advances in the field of urban water demand metering have fostered the development of smart metering technologies that can sense water use with fine sub-daily sampling resolutions, down to a few seconds (Mayer and DeOreo, 1999). Scientific literature on water demand modelling and management reports an increasing number of successful studies and use cases (for a review, see Cominola et al., 2015, and references therein) demonstrating the benefits of smart metering technologies to support demand-side management strategies that can complement traditional water supply development (Gleick et al., 2003). Recent applications showed that effective demand management strategies are a result of understanding users' typical behaviours and the associated consumption patterns at different spatial and temporal resolutions (Jorgensen et al., 2009, 2013). Yet, the adoption of smart metering technologies is still limited in utility and commercial applications because utilities are conservative, reluctant to change (Stewart et al., 2010), and the costs, benefits, and tradeoffs for investing in smart meters are unclear.

At coarse temporal resolutions, water use data are usually collected on a quarterly or monthly basis focusing on the urban or

* Corresponding author. Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milan, Italy.
E-mail address: andrea.castelletti@polimi.it (A. Castelletti).

https://doi.org/10.1016/j.envsoft.2017.11.022
1364-8152/

suburban scale to inform strategic regional planning with predictions of the aggregated water demand at the municipal or district level (House-Peters and Chang, 2011). Moving towards higher temporal resolutions, the advent of smart meters in the late 1990s opened up a new potential to better characterize water demand patterns on the basis of water consumption data at very high spatial and temporal resolution, for instance enabling end-use disaggregation (Nguyen et al., 2013) and better estimates of demand peaks (Beal et al., 2016). Depending on the technology exploited in the meter, we can distinguish four types of sensors: (i) Accelerometers (e.g., Evans et al., 2004), which analyze vibrations in a pipe induced by the turbulence of the water flow; (ii) Ultrasonic sensors (e.g., Mori et al., 2004), which estimate the flow velocity by measuring the difference in time between ultrasonic beams generated by piezoelectric devices and transmitted within the water flow; (iii) Pressure sensors (e.g., Froehlich et al., 2011), which estimate the flow rate as a function of the pressure change generated by the opening/close of the water devices valves via Poiseuille's Law; (iv) Mechanical or magnetic flow meters (e.g., Mayer and DeOreo, 1999; Kowalski and Marshallsay, 2003), which correlate the number of revolutions or nutations of a piston, magnet, or disk to the water volume passing through the meter. These sensors offer theoretical resolutions finer than 0.02 L, but cost, staff time, privacy, and regulations strongly constrain the actual resolutions that can be guaranteed by large scale Advanced Metering Infrastructure (AMI; Boyle et al., 2013). Understanding the tradeoff between the value of the information provided by high-resolution data and metering economic and operational costs is crucial to inform the design of smart metering networks as well as to discover and guard against unintended consequences of deployment options.

At one extreme of this tradeoff curve, the availability of high-resolution smart metered data generates several opportunities for advancing water demand management. Sub-minute sampling resolution is needed to run most water end-use disaggregation algorithms and provide a reliable breakdown of household level water use into different categories (e.g., shower, toilet, clothes washing machine; Nguyen et al., 2013b, 2015). The knowledge of timings, peak-hours, and frequencies of use of the different consumption devices is key to understand consumer behaviours, identify consumption anomalies, and, ultimately, design targeted personalized demand management strategies, including economic incentives to upgrade inefficient appliances (e.g., Mayer et al., 2004; Suero et al., 2012) or awareness campaigns targeting specific end uses (e.g., Willis et al., 2010; Abdallah and Rosenberg, 2014).

Yet, this metering strategy inevitably increases the amount of data the water utility must collect and handle. Sampling at one-minute resolution, for instance, implies replacing the four annual readings per user with 525,600 data readings. This increase may challenge business hardware and software performance due to existing issues with respect to power source, battery life, telemetry network capacity, black spots, i.e., data gaps, and billing software (Stewart et al., 2010). In addition, there is still no consensus about the best architecture to store consumption data. A centralized system facilitates checking the accuracy of the collected data, while a distributed one would significantly reduce transmission costs (Oracle, 2009).

Intermediate metering strategies attempt to balance these competing interests by sampling at resolutions of a few minutes to 1 h. Although this choice prevents an accurate characterization of end-use consumption profiles from aggregate signals with time spacing larger than a minute (e.g., toilet flushing or tap usage usually last a few seconds, showering a few minutes, thus it is hard to unpack end-use information from aggregate signals at coarser resolutions), these data still provide valuable information to water utilities and agencies for designing and managing the water supply system. In fact, sub-daily sampling resolutions allow extracting consumption patterns and accurately estimating the total water demand that the water supply system should be able to deliver to a group of users (e.g., Cardell-Oliver, 2013; Cominola et al., 2018). This can be seen by looking at the sample water use data reported in Fig. 1, which shows how the variability of water use patterns is gradually masked as data are sampled at progressively longer time intervals. Moreover, medium-resolution data can also support the identification of anomalous events occurring in the network or downstream of the household meter (e.g., post meter leakage, empty houses, or frauds). This is a major interest for water utilities because post meter leakages account for up to 10% of total residential water use. Reducing the amount of water wasted through leakages also generates secondary benefits in terms of reduced water-related energy consumption and treatment costs (see, for instance, Britton et al., 2013 study in Australia).

This tradeoff between metering cost and accuracy can influence the type of demand management operations and strategies available to utility managers, program costs, and corresponding benefits for water consumers and utilities. In this paper, we quantitatively assess how different temporal resolutions to read residential water meters impact information retrieval and demand management by answering the following research questions: which aspects of water demand modelling and management can be accurately, feasibly, and cost-effectively informed by different data resolutions? Are there resolution thresholds discriminating on these aspects?

To answer these questions, we contribute a comparative framework to explore the tradeoffs between data sampling resolution and accuracy in end-use disaggregation, time to detect post meter leaks, errors in estimating the volume and timing of peak flows, data storage requirements, and commercial availability. Given the low availability of residential water use data at different resolutions, we first developed a stochastic simulation model named STochastic Residential water End-use Model (STREaM). STREaM relies on a large dataset including observed and disaggregated water end uses from over 300 single-family households in nine U.S. cities (DeOreo, 2011). STREaM generates synthetic time series of water end use with diverse sampling resolutions. Second, we applied the comparative framework on STREaM output. STREaM allows the generation of residential water demand traces at the end-use level up to a 10-s resolution. Each water end-use fixture in our model is characterized by its signature (i.e., typical consumption pattern), as well as its probability distributions of number of uses per day, single use durations, single use water volumes, and time of use during the day. STREaM was used to generate a set of annual consumption traces for 500 heterogeneous households in terms of both number of occupants and efficiency of the end-use fixtures. The implications of adopting different data sampling resolutions are then explored by aggregating the generated 10-s water consumption trajectories up to the 1-d resolution and by evaluating a set of performance metrics including end-use disaggregation accuracy, costs due to leakage detection delay, precision in reproducing volume and timing of water demand peaks, data storage requirements, and commercial availability of metering systems. We use the framework to explore which temporal data resolutions might enable water demand management actions, utilities operations, and communication of customized information to water consumers. Findings from our multi-resolution assessment can support further research and commercial development in water meters and deployment of AMI, as well as assist utilities in trading off benefits from second-to-minute data sampling resolution and cost of adopting and maintaining high-resolution metering infrastructures.

The paper is organized as follows: the next section introduces the proposed comparative framework for multi-resolution assessment and formalises the set of performance metrics used in this study. Section 3 illustrates the synthetic generation of residential water