Environmental Modelling & Software 80 (2016) 33–40

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Parameterizing residential water demand pulse models through smart meter readings

Enrico Creaco [a, c, *], Panagiotis Kossieris [b], Lydia Vamvakeridou-Lyroudia [c], Christos Makropoulos [b], Zoran Kapelan [c], Dragan Savic [c]

[a] DICAr, University of Pavia, Via Ferrata 3, 27100 Pavia Italy
[b] Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, 5, Iroon Polythechniou, Zografou 15773 Athens, Greece
[c] College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF, UK

ARTICLE INFO

Article history:
Received 29 August 2015
Received in revised form
9 February 2016
Accepted 12 February 2016
Available online xxx

Keywords:
Water demand
Demand pulses
Intensity
Duration
Correlation
Parameterization

ABSTRACT

This paper proposes a method for parameterizing the Poisson models for residential water demand pulse generation, which consider the dependence of pulse duration and intensity. The method can be applied to consumption data collected in households through smart metering technologies. It is based on numerically searching for the model parameter values associated with pulse frequencies, durations and intensities, which lead to preservation of the mean demand volume and of the cumulative trend of demand volumes, at various time aggregation scales at the same time. The method is applied to various case studies, by using two time aggregation scales for demand volumes, i.e. fine aggregation scale (1 min or 15 min) and coarse aggregation scale (1 day). The fine scale coincides with the time resolution for reading acquisition through smart metering whereas the coarse scale is obtained by aggregating the consumption values recorded at the fine scale.

Results show that the parameterization method presented makes the Poisson model effective at reproducing the measured demand volumes aggregated at both time scales. Consistency of the pulses improves as the fine scale resolution increases.

1. Introduction

In recent decades, a substantial amount of research (e.g., Buchberger and Wu, 1995; Buchberger and Wells, 1996; Blokker et al., 2010; Guercio et al., 2001; Alvisi et al., 2003; Buchberger et al., 2003; Garcia et al., 2004; Alcocer-Yamanaka et al., 2006; Alcocer-Yamanaka and Tzatchkov, 2012; Creaco et al., 2015a, b, 2016) has been carried out to set up user demand models at high time resolution (down to 1 s). If suitably calibrated, these models can be used to generate consistent demand pulses coming from a single household or a group of households. Spatial and temporal aggregation of the pulses through the "bottom-up" approach (Walski et al., 2003) then enables reconstruction of nodal demand trends, to be used inside water distribution models. Furthermore, the local flow field given by these models can also be used as an input to water-quality models that require ultrafine temporal and spatial resolutions to predict the fate of contaminants moving through municipal distribution systems (Buchberger and Wu, 1995).

Unlike the model proposed by Blokker et al. (2010), which reproduces the demand from its respective micro components (i.e. by adding up the single water uses), most models reproduce the overall water demand, without distinguishing the contributions of the various appliances of the user's. To this end, they use stochastic processes such as the Poisson rectangular pulse process (Buchberger and Wu, 1995; Buchberger and Wells, 1996; Guercio et al., 2001; Buchberger et al., 2003; Garcia et al., 2004; Alcocer-Yamanaka et al., 2006; Creaco et al., 2015a, b, 2016) or the Neyman-Scott cluster process (Alvisi et al., 2003; Alcocer-Yamanaka and Tzatchkov, 2012).

A basic assumption considered in the demand pulse generation

* Corresponding author.DICAr, University of Pavia, Via Ferrata 3, 27100 Pavia, Italy.
E-mail addresses: creaco@unipv.it, E.F.Creaco@exeter.ac.uk (E. Creaco), pkossier@central.ntua.gr (P. Kossieris), L.S.Vamvakeridou-Lyroudia@exeter.ac.uk (L. Vamvakeridou-Lyroudia), cmakro@mail.ntua.gr (C. Makropoulos), z.kapelan@exeter.ac.uk (Z. Kapelan), D.Savic@exeter.ac.uk (D. Savic).

http://dx.doi.org/10.1016/j.envsoft.2016.02.019
1364-8152/

models is the independence of duration and intensity at the level of the single pulse (Buchberger and Wu, 1995; Buchberger and Wells, 1996; Guercio et al., 2001; Alvisi et al., 2003; Buchberger et al., 2003; Garcıa et al., 2004; Alcocer-Yamanaka et al., 2006; Alcocer-Yamanaka and Tzatchkov, 2012). On the basis of this assumption, after generating pulse arrivals, mono-variate probability distributions, such as the exponential, normal and lognormal, are used to generate the pulse durations and intensities in an independent way.

However, by working out calculations on the data collected on Milford households by Buchberger et al. (2003), Creaco et al. (2015a) have recently shown that a non-negligible positive correlation may exist between the two variables. Furthermore, Creaco et al. (2015a) proved that Poisson models with correlated pulse durations and intensities are advantageous compared to uncorrelated Poisson models. In fact, the use of suitably calibrated Poisson models with correlated pulse durations and intensities enables obtaining:

1 − statistically consistent pulse arrivals, durations and intensities;
2 − consistent values of the daily demand in the household.

On the other hand, the same Authors showed that demand pulses generated by uncorrelated Poisson models may fail to comply with either condition. Other evidence of the advantages of the correlated models was provided by Creaco et al. (2015b, 2016).

For model parameterization, Creaco et al. (2015a) proposed using the methods of the moments (Hall, 2004), consisting in setting the values of the parameters equal to the corresponding values in the measured pulses. However, this method can only be applied when data are available concerning the measured pulses. Unluckily, this seldom happens. Actually, the only example of such data in the scientific literature comes from the work of Buchberger et al. (2003). On the other hand, the use of smart metering technologies is currently on the rise (Boyle et al., 2013). These technologies enable acquisition and storage of household water consumption volumes with a very high temporal resolution (down to 1 min), which is only lowered under conditions of limited logging capacity (Mayer et al., 2004; Arregui et al., 2006; Kim et al., 2008; Cominola et al., 2015). However, to the best of the authors' knowledge, no methodology for parameterizing, through smart meter readings, Poisson models with correlated pulse durations and intensities has been proposed in the scientific literature so far. A specific methodology for such a situation is proposed in this paper, where the model parameters associated with pulse frequencies, durations and intensities are searched for numerically, in such a way as to preserve the mean demand volume and the cumulative trend of demand volumes, as obtained through smart metering, at various aggregation scales at the same time.

In the following sections, first the methodology is presented, followed by applications, where data derived from iWidget project were also used.

2. Methodology

Hereinafter, an overview of the Poisson model that considers correlation between pulse durations and intensities is provided, followed by the calibration procedure proposed.

2.1. Poisson model with correlated pulse durations and intensities

Inside the model, time axis is sampled with a certain time resolution Δt. The probability P of having z generated pulses in the time interval Δt that follows the generic time τ is described by the Poisson distribution (Buchberger and Wu, 1995):

$$P(z) = \frac{e^{-\lambda\Delta t}(\lambda\Delta t)^z}{z!} \quad \text{with } z = 0, 1, \ldots \tag{1}$$

where λ represents the expected number of "events" or "arrivals" that occur per unit time.

For each pulse arrival, the corresponding duration T and intensity I have to be generated. In order to preserve the correlation between the two variables, a bivariate model can be used (Creaco et al., 2015a). As an alternative, resorting algorithms can be used to obtain correlation between the variables, when these are generated independently (Creaco et al., 2016). In particular, the lognormal bivariate distribution model, whose probability density function f takes on the following form, was used by Creaco et al. (2015a) to generate demand pulses:

$$f(y_1, y_2) = \frac{1}{\sqrt{(2\pi)^2|\boldsymbol{\Sigma}|}} e^{\left\{-\frac{1}{2(1-\dot{\rho}^2)}\left[\left(\frac{y_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y_2-\mu_2}{\sigma_2}\right)^2 - \frac{2\dot{\rho}(y_1-\mu_1)(y_2-\mu_2)}{\sigma_1\sigma_2}\right]\right\}}, \tag{2}$$

where y_1 is equal to $ln(T)$, y_2 is equal to $ln(I)$, μ_i $(i = 1,2)$ is equal to the mean of y_i and $\boldsymbol{\Sigma}$ is the covariance matrix, which takes on the following expression:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \dot{\rho}\sigma_1\sigma_2 \\ \dot{\rho}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \tag{3}$$

where σ_i and $\dot{\rho}$ indicate the standard deviation of variable y_i $(i = 1,2)$ and the Pearson correlation between y_1 and y_2 respectively.

Once generated, the pulses can be composed to obtain the water demand time series of the individual user, as is shown in Fig. 1. In this context, it has to be noted that the constant pulse intensity, which causes the single synthetic pulse to take on a rectangular shape (Fig. 1), is an assumption that was considered by various other authors (Buchberger and Wu, 1995; Buchberger and Wells, 1996; Guercio et al., 2001; Alvisi et al., 2003; Buchberger et al., 2003; Creaco et al., 2015a, b, 2016) in the scientific literature. Indeed, it has to be interpreted as the average intensity of a real pulse, since DeOreo (2011) showed that the pulse intensity of some real appliances can be quite variable in time.

2.2. Parameterization

A Poisson model with correlated pulse durations and intensities, such as that described in the previous section, features 6 parameters. The first parameter is rate parameter λ, related to pulse arrivals. Then, μ_1 and σ_1 relate to pulse durations T whereas μ_2 and σ_2 relate to pulse intensities I. Finally, $\dot{\rho}$ is representative of the correlation between T and I. When the pulse generation model is parameterized in order to reproduce the demand pulses of a household in a generic month, each of parameters μ_1, σ_1, μ_2, σ_2 and $\dot{\rho}$ is assigned a single value, valid for the whole period. Parameter λ, instead, is generally assumed to take on time-varying daily values (Creaco et al., 2015a). To this end, the day can be conveniently subdivided into a certain number n_{slot} of time slots (e.g., 12 time slots according to Alvisi et al., 2003 and Creaco et al., 2015a). Inside the generic mth slot, the value λ_m of rate parameter λ is considered constant.

Since data concerning real demand pulses are generally unavailable, the model parameters can be assessed as a function of the demand volumes at a given aggregation scale. Though some analytical methods are present in the scientific literature (e.g.,