Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 154 (2016) 123 – 131

Procedia Engineering

www.elsevier.com/locate/procedia

12th International Conference on Hydroinformatics, HIC 2016

Assessing the applicability of the Bartlett-Lewis model in simulating residential water demands

Panagiotis Kossieris[a,]*, Christos Makropoulos[a], Enrico Creaco[b], Lydia Vamvakeridou-Lyroudia[b], Dragan A. Savic[b]

[a]*Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, 5, Iroon Polythechniou, Zografou 15773 Athens, Greece*

[b]*College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF – UK.*

Abstract

This paper presents the set-up and application of the Bartlett-Lewis clustering mechanism to simulate residential water demand at fine, i.e. sub-hourly, time scales. Two different variants of the model, i.e., the original and the random-parameter model, are examined. The models are assessed in terms of preserving the main statistical characteristics and temporal properties of demand series at a range of fine time scales, i.e., from 1-min up to 15-min. The comparison against the typical Poisson rectangular pulse model showed that clustering mechanism enables a better reproduction of demand characteristics at levels of aggregation other than those used in the fitting procedure.

Peer-review under responsibility of the organizing committee of HIC 2016

Keywords: residential water demand; Bartlett-Lewis model; stochastic simulation; Poisson models; smart metering; PRP model;

1. Introduction

The high temporal and spatial variability of water demand poses extra difficulties in the efficient planning and management of water distribution and sewerage systems [1,2]. To cope with this uncertainty stochastic simulation techniques are usually employed to generate a large number of possible realizations of demand events across a wide

* Corresponding author. Tel.: +30+210+7722828; fax: +30+210+7722828.
E-mail address: pkossier@central.ntua.gr

Peer-review under responsibility of the organizing committee of HIC 2016
doi:10.1016/j.proeng.2016.07.429

range of time and spatial scales, from minute to days and from the household to the DMA. Recently, the rising deployment of smart metering technologies, by providing high resolution data from households, enables the implementation of a "bottom-up" approach [3] that allow the demand series at higher temporal and spatial scales to be obtained by aggregating the demands from individual users. In this context, a special focus is given to the development of models that reproduce the statistical properties and the stochastic structure of water demand at individual user level. In this paper our focus is on fine time scales (i.e., sub-hourly) which are the temporal resolution of interest in water quality modelling of the peripheral parts of a network. In these parts the qualitative characteristics of water and the flows in the pipes are directly affected by the high variability of residential water demand [1,4].

At fine time scales, the stochastic modelling of residential water demand has been studied with the use of the Poisson [5–10] and Poisson-cluster processes [11–13]. The main representative of the first category is the Poisson rectangular pulse (PRP) model [6]. It represents the demand mechanism via rectangular pulses that occur in continuous time according to the Poisson point process. Each rectangular pulse is associated with a random duration and intensity that follow specific probability distributions. In the initial formulation of the model, pulse durations and intensities are assumed mutually independent variables and independent of the arrival process. Recently, a model variant that introduces correlation between the intensity and duration was proposed by Creaco et al. [10].

In the Poisson-cluster processes, the events are formulated via clustered rectangular pulses. In this case, two different Poisson processes are used to specify the time origins of clusters and the origins of pulses within clusters, respectively. Rodriguez-Iturbe et al. [14,15] first studied two alternative types of clustering, the Bartlett-Lewis and Neyman-Scott process, after observing the weakness of simple PRP-type models to reproduce the statistical characteristics of rainfall at multiple time scales. The main difference between the Bartlett-Lewis and Neyman-Scott mechanism lies in the way that pulse origins are distributed within a cluster. Transferring the clustering concept into demand modelling, the assumption of representing the consumption events through sequences of pulses, rather via individual pulses, seems more consistent with the actual demand mechanism, i.e., sequences of different water use activities are performed sequentially or simultaneously in the house [11]. Alvisi et al. [11] first examined the applicability of the original Neyman-Scott rectangular pulse (NSRP) model to simulate the total water demand at different fine time scales (i.e, from 1-min up to 15-min) of a small group of households. The Neyman-Scott model was later used by Alcocer-Yamanaka et al. [13,16]. On the contrary, the Bartlett-Lewis clustering process has never been studied for the residential demand modelling. Furthermore, a direct comparison between the Poisson and the Poisson-cluster models is not available in the literature.

This study presents the set-up and application of the Bartlett-Lewis rectangular pulse model to simulate the water demand of a single-household, as it is recorded via the smart meter. Specifically, two different variants of the model, i.e., the original and the random parameter model, are examined. To fully evaluate the clustering mechanism against individual pulse processes, the Bartlett-Lewis models were compared to the original PRP model with independent pulse intensities and durations. The assessment of the models was conducted in terms of preserving the main statistical characteristics and temporal properties of a demand series from a household in Athens, at a range of fine time scale, varying from 1-min up to 15-min.

2. Model description

2.1. The Poisson rectangular pulse model

The Poisson Rectangular Pulse model was initially proposed to simulate rainfall events [14], while the model used later in the modelling of water demand [6]. In the PRP model, event origins occur according to a Poisson process with rate λ. Each pulse i has a random intensity x_i and a random duration w_i that follow a specific probability distribution. Different probability distributions have been used to describe pulse characteristics depending on what is best fitted to the observed data. Regarding pulse intensity, the two-parameter Weibull distribution was assumed by Garcia et al. [8], while the normal [7], log-normal [17] or exponential [1,12] distributions have been also used. The exponential distribution has been mainly used for pulse duration [1,7,8,12] while log-normal distribution seems also valid [9,17].

The PRP model parameters can be obtained directly from individual pulse characteristics in the case that instantaneous demand measurements (i.e., 1-sec time step) are available. This process has been followed in most applications of the model in water demand modelling. However, such data is available only at a small number of pilot households, while the typical lower time resolution of commercially available smart metering devices is the 1 minute, and subsequently, the decomposition of total demand into equivalent pulses is not possible. In that case, model fitting is based on the theoretical equations of the model that relates the statistical properties of the discrete process to model parameters [7].

Recently, a new variant of the PRP model parameter that captures the positive correlation observed between water use intensity and duration was proposed by Creaco et al. [10]. More specifically, in the modified model, hereinafter notated as PRP_X, pulse intensities and durations are distributed according to a bivariate log-normal distribution with their marginal distribution being also the log-normal. The PRP_X model was further generalised to establish correlation in the case that other distributions are assumed for the two variables [9]. The analysis showed that PRP_X outperforms the original model with respect to the daily demand volumes. PRP_X can be fitted directly on the pulse series, while more recently a new parameterization approach was developed in the case that coarser meter readings are available [18].

2.2. The Bartlett-Lewis rectangular pulse models

The basic assumptions of the Bartlett-Lewis clustering mechanism [14], as shown in Fig. 1, are: (a) water consumption events occur according to a Poisson process with rate λ and each event is associated with a random number of pulses; (b) within each event, pulse origins t_{ij} occur from a second Poisson process with rate β; (c) within each event, the generation of pulses terminates after a time span v_i that follows the exponential distribution with rate γ; (d) each pulse has a duration w_{ij} that follows the exponential distribution with rate η; and (e) an intensity x_{ij} that, in the simplest case, follows the exponential distribution with mean μ_X. Model's structure implies that the number of pulses per consumption event has a geometric distribution of mean $\mu_c = 1 + \beta/\gamma$. The model allows both for event and pulse overlapping, while the total consumption intensity, $Y(t)$, at every instant t, is obtained by summing all active pulses at time t.

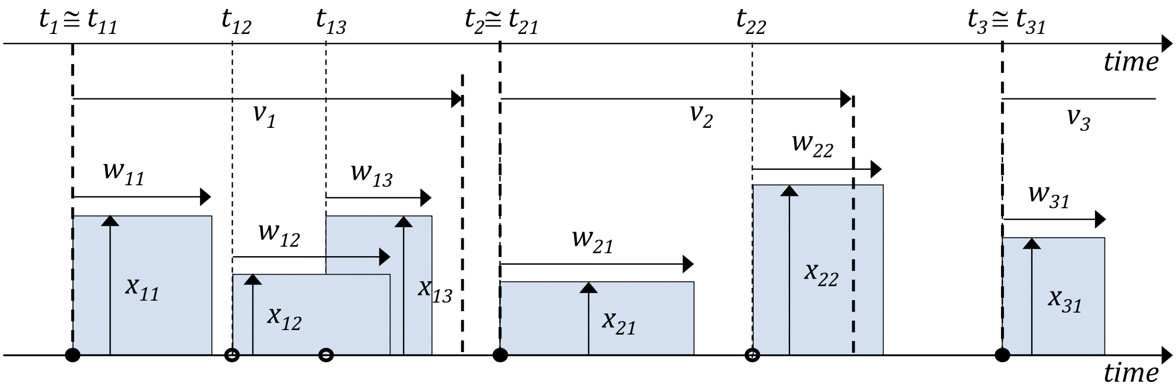

Fig. 1. Schematic representation of the Bartlett-Lewis clustering mechanism. Filled circles denote the origins of consumption events while open circles denotes pulse arrivals.

The basic version of the model, hereinafter referred to as the Rectangular Pulse Bartlett-Lewis (RPBL) model, has 5 constant independent parameters: $\{\lambda, \beta, \gamma, \eta, \mu_X\}$. This implies that all consumption events are structured by pulses whose characteristics are sampled from the same statistical population.

However, this hypothesis is not in full compliance with the actual demand mechanism where a consumption event may consist of short duration pulses (e.g., hand washing, cooking, dishwashing) or longer pulses (e.g., washing machine or bathing) or a combination of them. Such a variation between different consumption events is implied by the Random Parameter Bartlett-Lewis (BLRPR) model that assumes that parameter η of pulse duration varies randomly among events according to the gamma distribution with a shape parameter α and a rate parameter ν [15]. The parameterization of the BLRPR model assumes also that the pulse origin rate β and the event duration rate γ are also randomly varied parameters so that the ratios $\kappa = \beta/\eta$ and $\varphi = \gamma/\eta$ are kept constant across the process. Subsequently, in the BLRPR model, parameters β and γ are also random variables following gamma distribution with common shape parameter α and rate parameters ν/κ and ν/φ, accordingly. The BLRPR model has 6 independent parameters: $\{\lambda, \alpha, \nu, \kappa, \varphi, \mu_X\}$. The number of model parameters is increased by assuming longer-tailed distributions, such as gamma, Weibull or Pareto, for cell intensity variable [19].

The statistical properties, i.e., mean, variance, covariance, and probability of zero demand, of the RPBL and BLRPR processes in discrete time are given via analytical equations as a function of the model parameters [14,15]. The estimation of model parameters is conducted usually via the method of moments by equating theoretical moments with the observed statistics.

3. Application

3.1. Case study and parameter estimation

In the present paper, the performance of the Bartlett-Lewis and Poisson models was assessed using a time-series of total water consumption, with 1-min temporal resolution, from a single household in Athens, Greece. The data was drawn via the smart metering equipment installed at October 2014 under the framework of the iWIDGET project [20], and cover a period of 30 days during February-March 2015. The same dataset has been used also by Creaco et al. [18] and it was chosen due to the quality of data (i.e., small number of missing data) and the regular consumption of water.

As the time resolution of measurements does not allow fitting the models directly upon pulse characteristics, the analytical expressions of the first- and second-order moments and the probability of zero demand of the models were utilized. The parameter estimation for the PRP, RPBL and BLRPR models was conducted with the use of "MOMFIT" package [21], developed in R, that implements the generalized method of moments within a two-step optimization framework [22]. The form of the objective function Z used in this study is:

$$Z = \min\left[\sum_{i=1}^{N} W_i \left(\frac{F_i(X)}{F_i^{'}} - 1\right)^2\right] \tag{1}$$

where $F_i(X)$ is the model expression for statistic i, X is the parameter vector, F_i' the observed statistic i, Wi the weight assigned to statistic i and N is the total number of statistics used in parameter estimation. It is well known that parameter estimation process is highly sensitive to the combination of moments used in Eq. (1), resulting in equally good parameter sets in terms of minimum value of objective function [23]. However, a more comprehensive assessment of the overall performance of the model can be obtained by studying also the statistics that are not included in the fitting procedure.

Regarding the choice of moments included in fitting procedure, the preliminary analysis showed that the use of statistics in Eq. (1) from only one aggregation level results in good fitting at this level, but not in adequate model performance at other time scales. On the other hand, by using statistics from finer (e.g., 1-min) and coarser (e.g., 15-min) levels, the reproduction of process at a wider range of time scales can be achieved. The statistical characteristics that were used to estimate the parameters of PRP, RPBL and BLRPR models are given in Table 1, while in all optimizations the weight vector was set equal to 1. The optimization routine requires the lower and upper bounds for each parameter. The identification of the bounds, shown in Table 2, was conducted through a preliminary analysis along with values proposed in the literature [7,11] as reasonable for the various water uses. In the case of the BLRPR model, parameters κ, φ, α and ν do not have a direct link to the actual demand mechanism. To define the feasible search space of the parameters the temporal properties of the process were used, i.e., the mean duration of water use activity, $E(\gamma) = \nu/((\alpha\text{-}1)\varphi)$, the mean pulse inter-arrival time, $E(\beta) = \nu/((\alpha\text{-}1)\kappa)$, and the mean pulse duration, $E(\eta) = \alpha/\nu$.

Table 1. Statistics used for models' fitting per level of aggregation

Model	Mean	Coef. of Variation	Lag-1 Autocorrelation	Probability of zero demand
PRP	1-min	1-min, 5-min	1-min, 5-min	-
BLRP	1-min	1-min, 5-min	1-min, 5-min	-
BLRPR	1-min	1-min, 5-min	1-min, 5-min	1-min, 5-min

Table 2. Upper and lower bounds for models' fitting

PRP			BLRP			BLRPR		
Parameters	Lower	Upper	Parameters	Lower	Upper	Parameters	Lower	Upper
λ^{-1} (min)	5	200	λ^{-1} (min)	5	1000	λ^{-1} (min)	5	1000
η^{-1} (min)	0.1	10	η^{-1} (min)	0.01	5	ν (min)	0.00001	2500
μ_X (L/min)	0.5	8	μ_X (L/min)	0.5	8	μ_X (L/min)	0.5	8
			β^{-1} (min)	0.016	20	α (-)	1.00001	500
			γ^{-1} (min)	1	30	κ (-)	0.001	315
			μ_c (n° of pulses)	1.5	62.5	φ (-)	0.00033	5

3.2. Results

The Poisson and the Bartlett-Lewis models were assessed in terms of preserving certain statistical characteristics of water demand at different time scales, varying from 1-min up to 15-min. To take into consideration the variability that demand characteristics exhibit within the day, the analysis was conducted by dividing the day in 12 two-hour intervals, following the approach of Alvisi et al. [11]. To this end, models' parameters were estimated, assuming that the process is stationary within each of the time intervals, while all parameters were left free to vary among different time intervals.

The estimated parameters are not provided in detail in this paper due to the lack of space. In general, all models exhibit different behaviour between the time intervals of higher (i.e., morning, afternoon and evening) and lower consumption (i.e., night hours). The optimized parameters are similar within these distinct periods. In the higher-consumption cases, the pulses occur with higher frequency, having greater mean intensity and duration than in the case of night (low consumption) hours. By comparing the parameters of the PRP model with those of the Bartlett-Lewis model, we can infer that during the periods of higher consumption, the former generates pulses with longer duration (e.g., 3.5 min) while the latter consist of pulses that occur more frequently but have shorter duration (e.g., 1 min).

The results of the present analysis are shown graphically in

Fig. 2 through Fig. 5. Each graph displays a specific statistical property for the 12 time intervals, at time scales of 1, 5, 10 and 15 minutes. Further to the statistics of the observed data, the statistics of the PRP, RPBL and BLRPR models are those obtained from theoretical equations with the use of the optimised parameters.

In

Fig. 2 and Fig. 3 we can see that all models reproduce the mean water demand exactly at all aggregation levels and perform well with respect to the variance. It is worth remembering that these two properties were included in the fitting of the three models. Regarding the lag-1 autocorrelation coefficient (Fig. 4), a very good fit was achieved with the two Bartlett-Lewis models, especially for the time scales of 1-min and 5-min, as well as the intermediate scales, i.e., 2-min, 3-min (not shown in this paper). The PRP model reproduces well that property at the 1-min time scale, producing, however, an autocorrelation structure that decays much faster as the level of aggregation increases. Furthermore, the analysis also showed that the Bartlett-Lewis mechanism, compared to the Poisson model, matches much better the autocorrelation structure of observed data for higher lag values. Finally, we examine the performance of the models in reproducing the probability of zero demand. As it was discussed in section 3.1, the probability of zero demand was used as fitting property only in the case of the BLRPR model, while its application to the fitting of the PRP and BLRP models not only does not improve the preservation of this property but also leads to the deterioration of models' performance upon rest statistics. As it is shown in Fig. 5, the BLRPR model outperforms the other models, enabling a better representation of the probability of zero demand, across all-time scales. This behaviour can be attributed to its structure that allows the generation of water events consisting of pulses with different characteristics. On the contrary, the RPBL model tends to overestimate that characteristic, especially in the periods of higher consumption where different water uses are performed at the same time. The PRP model also exhibit good performance, especially at the fine time scales.

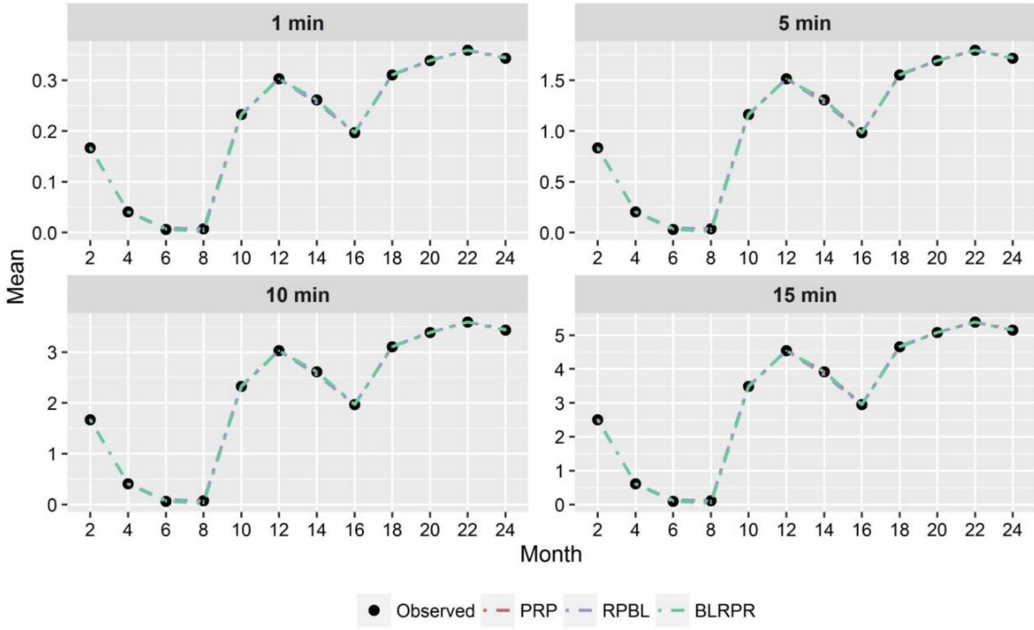

Fig. 2. Comparison of historical and modelled mean for each time interval and aggregation level

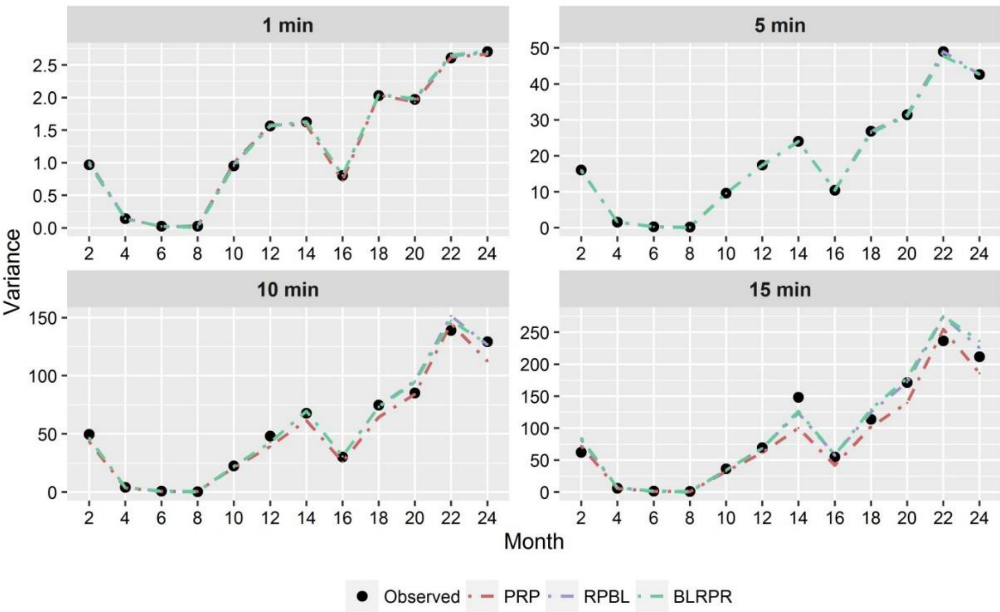

Fig. 3. Comparison of historical and modelled variance for each time interval and aggregation level

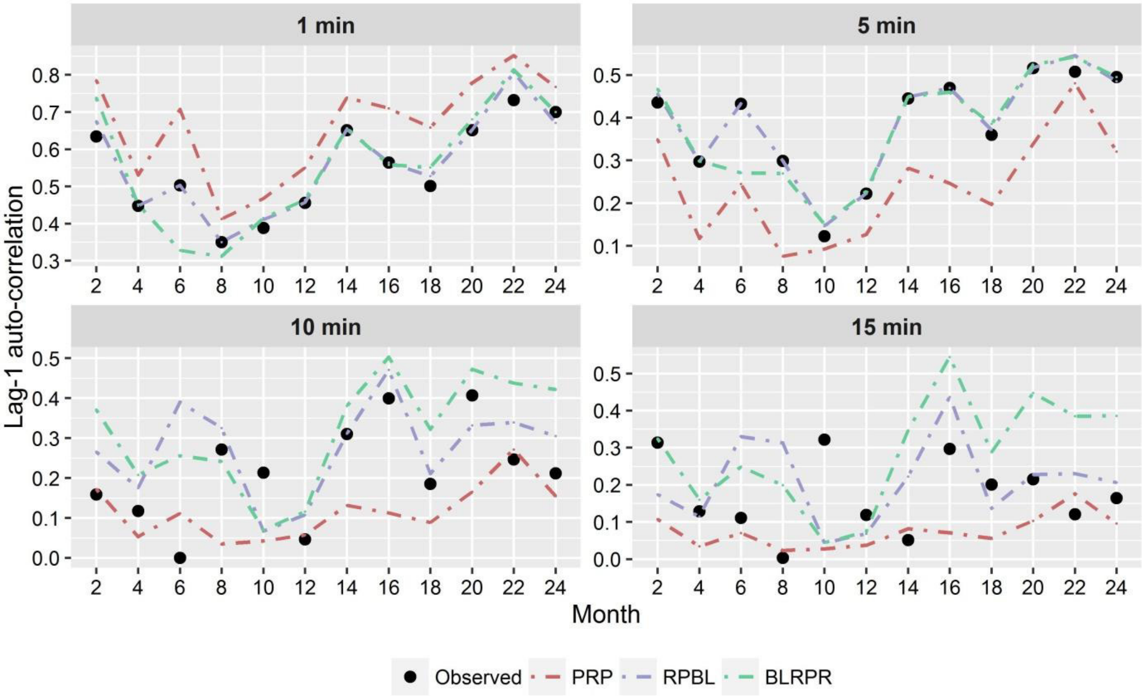

Fig. 4. Comparison of historical and modelled Lag-1 autocorrelation coefficient for each time interval and aggregation level

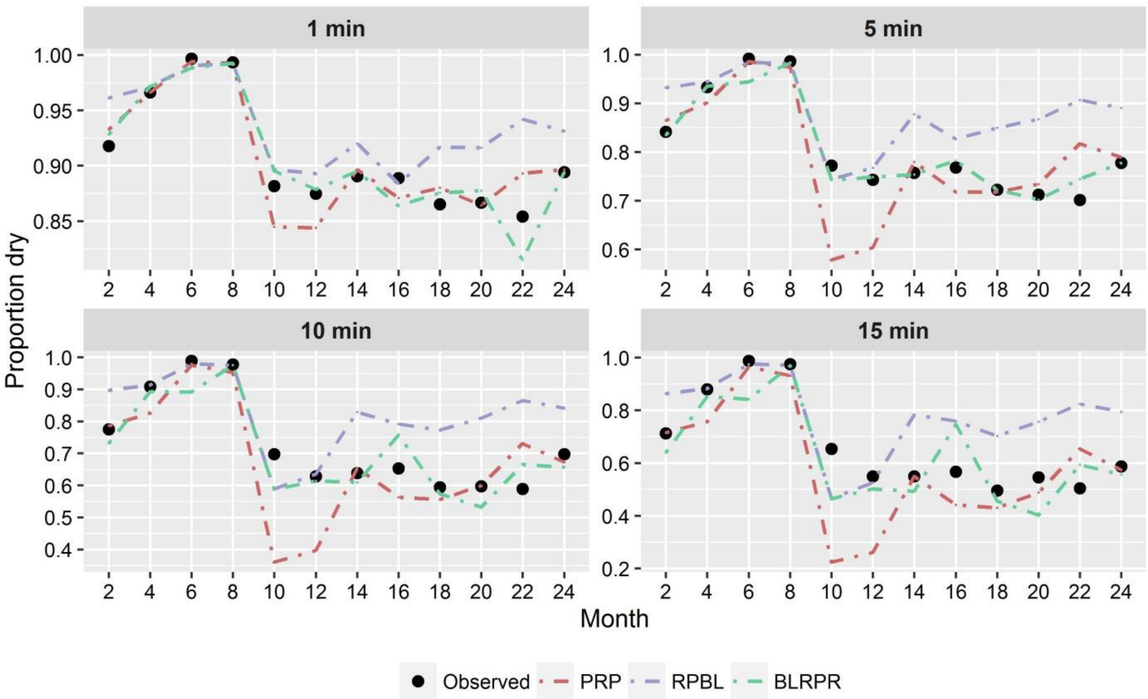

Fig. 5. Comparison of historical and modelled probability of zero demand for each time interval and aggregation level

4. Conclusions and future research

This paper examines the applicability of the Bartlett-Lewis clustering mechanism in simulating residential water demand at fine time scales. Two variants of this mechanism, the original and the random model, were assessed on the basis of preserving certain statistical characteristics and the temporal properties of demand at different time scales, varying the time step from 1-min up to 15-min. The models were compared against the widely used Poisson rectangular pulse model with independent pulse intensity-duration. The analysis was based on 1-min water demand data collected during the iWIDGET project and showed that all models perform well with respect to the mean and variance of water demand across time scales. Additionally, the clustering mechanism provides the flexibility of better preserving the autocorrelation structure and the probability of zero demand at time scales that are not directly involved in model fitting.

This paper is part of a wider study and presents some initial findings of the application of the Poisson and Poisson-cluster models to estimate model parameters in the case that only aggregated data from smart meters is available. Ongoing research is focused on a series of challenging issues such as the application of such models to simulate water demand of a small group of households, the performance of the models at higher temporal, i.e., daily or hourly water peaks and volumes, and spatial aggregation levels, i.e., large group of households. Further to that and in the case of Bartlett-Lewis models, the use of other probability distributions for pulse intensity and duration should be also considered. Finally, the use of recently developed variants of the Bartlett-Lewis mechanism to simulate residential water demand is under investigation.

Acknowledgements

This research was carried out in the framework of iWIDGET Project (2012-2015), received funding from the European Union Seventh Framework Programme (Grant Agreement No 318272).

References

[1] Magini R, Pallavicini I, Guercio R. Spatial and Temporal Scaling Properties of Water Demand. J Water Resour Plan Manag 2008;134:276–84. doi:10.1061/(ASCE)0733-9496(2008)134:3(276).

[2] Bao BY. M o d e l for w a t e r distribution system reliability 1991;116:1119–37.

[3] Walski T, Chase D, Savic D, Grayman W, Beckwith S, Koelle E. Advanced Water Distribution Modeling and Management. 2003.

[4] Blokker EJM, Vreeburg JHG, Buchberger SG, van Dijk JC. Importance of demand modelling in network water quality models: a review. Drink Water Eng Sci Discuss 2008;1:1–20. doi:10.5194/dwesd-1-1-2008.

[5] Buchberger SG, Wells GJ. Intensity, Duration, and Frequency of Residential Water Demands. J Water Resour Plan Manag 1996;122:11–9. doi:10.1061/(ASCE)0733-9496(1996)122:1(11).

[6] Buchberger SG, Wu L. Model for instantaneous residential water demands. J Hydraul Eng 1995;121 (3):232–46.

[7] Guercio R, Magini R, Pallavicini I. Instantaneous residential water demand as stochastic point process 2001;48.

[8] Garcia VJ, Garcia-Bartual R, Cabrera E, Arregui F, Garcia-Serra J. Stochastic Model to Evaluate Residential Water Demands. J Water Resour Plan Manag 2004;130:386–94. doi:10.1061/(ASCE)0733-9496(2004)130:5(386).

[9] Creaco E, Alvisi S, Farmani R, Vamvakeridou-Lyroudia L, Franchini M, Kapelan Z, et al. Preserving duration-intensity correlation on synthetically generated water demand pulses. Procedia Eng 2015;119:1463–72. doi:10.1016/j.proeng.2015.08.1007.

[10] Creaco, Farmani R, Kapelan Z, Vamvakeridou-Lyroudia L, Savic D. Considering the Mutual Dependence of Pulse Duration and Intensity in Models for Generating Residential Water Demand. J Water Resour Plan Manag 2015;141:04015031. doi:10.1061/(ASCE)WR.1943-5452.0000557.

[11] Alvisi S, Franchini M, Marinelli a. A stochastic model for representing drinking water demand at residential level. Water Resour Manag 2003;17:197–222.

[12] Alcocer-Yamanaka VH, Tzatchkov VG, Buchberger S. Instantaneous Water Demand Parameter Estimation from Coarse Meter Readings. Proc. Water Distrib. Syst. Anal. Symp., 2006, p. 1–14.

[13] Alcocer-Yamanaka VH, Tzatchkov VG, Arreguin-Cortes FI. Modeling of drinking water distribution networks using stochastic demand. Water Resour Manag 2012;26:1779–92. doi:10.1007/s11269-012-9979-2.

[14] Rodriguez-Iturbe I, Cox DR, Isham V. Some Models for Rainfall Based on Stochastic Point Processes. Proc R Soc Lond A Math Phys Sci 1987;410:269–88.

[15] Rodriguez-Iturbe I, Cox DR, Isham V. A Point Process Model for Rainfall: Further Developments. Proc R Soc London A Math Phys Eng Sci 1988;417:283–98.

[16] Alcocer-yamanaka VH, Buchberger S. FROM COARSE METER READINGS 2006.

[17] Buchberger SG, Carter JT, Lee YH, Schade TG. Random demands, travel times, and water quality in dead ends. Awwarf 2003.

[18] Creaco E, Kossieris P, Vamvakeridou-Lyroudia L, Makropoulos C, Kapelan Z, Savic D. Parameterizing residential water demand pulse models through smart meter readings. Environ Model Softw 2016;80:33–40. doi:10.1016/j.envsoft.2016.02.019.

[19] Onof C, Wheater HS. Improvements to the modelling of British rainfall using a modified Random Parameter Bartlett-Lewis Rectangular Pulse Model. J Hydrol 1994;157:177–95. doi:10.1016/0022-1694(94)90104-X.

[20] Savic D, Vamvakeridou-Lyroudia L, Kapelan Z. Smart meters, smart water, smart societies: The iWIDGET project. Procedia Eng 2014;89:1105–12. doi:10.1016/j.proeng.2014.11.231.

[21] Chandler R, Lourmas G, Jesus J. Momfit: software for moment-based fitting of single-site stochastic rainfall models. Department of Statistical Science, University College London. Available at http://www.ucl.ac.uk/~ucakarc/work/momfit.html. 2010.

[22] Jesus J, Chandler RE. Estimating functions and the generalized method of moments. Interface Focus 2011;1:871–85. doi:10.1098/rsfs.2011.0057.

[23] Onof C, Chandler RE, Kakou A, Northrop P, Wheater HS, Isham V. Rainfall modelling using Poisson-cluster processes: a review of developments. Stoch Environ Res Risk Assess 2000;14:384–411. doi:10.1007/s004770000043.