Methods for preserving duration-intensity correlation on synthetically generated water demand pulses

Enrico Creaco[1], Stefano Alvisi[2], Raziyeh Farmani[3], Lydia Vamvakeridou-Lyroudia[4], Marco Franchini[5], Zoran Kapelan[6], Dragan Savic[7]

Abstract

This paper proposes the application of three different methods for preserving the correlation between duration and intensity of synthetically generated water demand pulses. The first two methods, i.e., the Iman and Canover method and the Gaussian copula respectively, are derived from known statistical approaches, though they had never been applied to the context of demand pulse generation. The third is a novel methodology developed in this work and is a variation in the Gaussian copula approach. Poisson models fitted with the methods are applied to reproduce the measured pulses in one household, with parameters being obtained with the method of moments. Comparisons are made with another method previously proposed in the scientific literature, showing that the three methods have similar effectiveness and are applicable under more general conditions.

Keywords: Water demand, demand pulses, intensity, duration, correlation, Iman-Canover, Gaussian copula

[1] College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF – UK. Email: E.F.Creaco@exeter.ac.uk
[2] Dipartimento di Ingegneria, Università degli Studi di Ferrara, Via Saragat 1, 44100 Ferrara – Italy. Email: stefano.alvisi@unife.it
[3] College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF – UK. Email: R.Farmani@exeter.ac.uk
[4] College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF – UK. Email: L.S.Vamvakeridou-Lyroudia@exeter.ac.uk
[5] Dipartimento di Ingegneria, Università degli Studi di Ferrara, Via Saragat 1, 44100 Ferrara – Italy. Email: marco.franchini@unife.it
[6] College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF – UK. Email: Z.Kapelan@exeter.ac.uk
[7] College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF - UK. Email: D.Savic@exeter.ac.uk

Introduction

In the last two decades, residential water demand generation has been extensively investigated. Various models (Buchberger and Wu, 1995; Buchberger and Wells, 1996; Guercio et al., 2001; Alvisi et al., 2003; Garcia et al., 2004; Buchberger et al., 2003; Alcocer et al., 2006; Blokker et al., 2010; Alcocer-Yamanaka and Tzatchkov, 2012; Alvisi et al., 2014; Creaco et al., 2015) have been proposed to generate water demand pulses at the scale of individual user with fine temporal resolution (down to 1 sec). In fact, this modelling is useful in the framework of the "bottom-up" approach (Walski et al., 2003) for network demand definition, since the generated pulses can be aggregated temporally and spatially to yield nodal demands inside water distribution models. Unlike other demand generation models, which produce demand values at prefixed time steps, pulse generation models generate the time arrival, duration and intensity of each pulse. Therefore, 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).

Some of these models (Buchberger and Wu, 1995; Buchberger and Wells, 1996; Guercio et al., 2001; Buchberger et al., 2003; Garcia et al., 2004; Alcocer et al., 2006; Creaco et al., 2015) use the Poisson pulse model for the generation of pulse time arrivals, whereas pulse durations and intensities are generated using suitable probability distributions. In most cases (Buchberger and Wu, 1995; Buchberger and Wells, 1996; Guercio et al., 2001; Buchberger et al., 2003; Garcia et al., 2004; Alcocer et al., 2006), pulse duration and intensity were considered to be independent random variables. However, Creaco et al. (2015) have recently shown that a non-negligible positive correlation exists between the two variables. The same authors then postulated that this has to be considered in order to obtain synthetic water demand pulses that are more consistent in terms of overall daily water demand volumes, while respecting statistical properties of measured demand pulses. In particular, the Authors' method is based on the use of a bivariate probability distribution (in particular, the bivariate normal distribution). However, researchers may choose to represent pulse durations and intensities through marginal probability distributions that do not provide for any bivariate distribution modelling. For example, this was previously done by Guercio et al. (2001) and Garcia et al. (2004), who used the normal and exponential distributions and the Weibull and exponential distributions respectively. Therefore, the issue of how correlation can be preserved in a more general context, i.e. when a bivariate probability distribution does not exist, needs to be dealt with. In this paper, three methods are described that can be applied to obtain correlated pulse intensities and durations for any marginal distribution used to represent the two variables as independent random variables.

In the following sections, first the methodologies are described, then they are applied to a literature case study and a comparison with the method of Creaco et al. (2015) is also provided. Finally, results are analysed and conclusions are drawn.

Methodology

Hereinafter, first the typical Poisson model with no correlation between pulse intensity and duration is described. Then, the methods used to preserve correlation are presented, followed by the model parameter estimation.

Poisson model

Time axis is sampled with a certain time resolution Δt. The probability $P(z)$ 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):

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

where rate parameter λ represents the expected number of "events" or "arrivals" that occur per unit time. For each pulse generated, the associate duration T and intensity I are generated using suitable probability distributions. As an example, the density functions of the beta and gamma distributions (Johnson and Bhattacharyya, 1992) are provided in eqs. (2) and (3) respectively:

$$f(x) = \frac{1}{B} \frac{\left(\dfrac{x - x_{\min}}{x_{\max} - x_{\min}} \right)^{\alpha - 1}}{\left(1 - \dfrac{x - x_{\min}}{x_{\max} - x_{\min}} \right)^{\beta - 1}} \tag{2}$$

$$f(x)=\frac{x^{k-1}e^{-\frac{x}{\theta}}}{\theta^{k}\Gamma} \quad (3)$$

where x is the random variable, equal to T or I, depending on which variable has to be generated; α and β are the parameters of the beta distribution and $B=B(\alpha, \beta)$ is the beta function; k and θ are the parameters of the gamma distribution and $\Gamma=\Gamma(k)$ is the gamma function. Whereas the gamma distribution is defined on the interval [0,+∞[, the beta distribution is defined on the interval $[x_{min}, x_{max}]$. Therefore, in order for the latter to be used for the generation of either the duration or the intensity, the interval $[x_{min}, x_{max}]$ has to be defined. In any case, the cumulative distribution function F that ranges from 0 to 1 can be obtained as (Johnson and Bhattacharyya, 1992):

$$F(x)=\int_{0}^{x} f(x)\,dx \quad (4)$$

After distribution parameters values have been fixed, values of the generic random variable can be sampled by generating for F random numbers in the range [0,1] and then deriving the corresponding elements of x by inverting eq. (4).

If a Poisson model is used with constant parameter values to generate water demand pulses for a certain time duration, a sequence of n pulses, each of which featuring its own time arrival τ_i, duration T_i and intensity I_i, would be obtained, as shown in Table 1.

Variables T and I, as they appear in columns 2 and 3 of Table 1, are independent random variables; the corresponding correlation matrix **C** (see eq. 5) should thus feature an expected value of ρ, Pearson correlation coefficient out of the diagonal, equal to 0.

$$\mathbf{C}=\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \quad (5)$$

Preserving correlation

Method 1: Iman and Canover (IC) (1982)

The Iman and Canover (1982) procedure is made up of two steps. In step 1, variables T and I are generated as independent random variables, as is described above. This results in matrix **X**, which is in fact made up of columns 2 and 3 of Table 1; the corresponding correlation matrix **C** is given by eq. (5). Step 2 is then applied to find a new pairing for these variables, which enables the desired/observed correlation ρ_{ep} value to be preserved between the variables.

In step 2, the following matrix operations are performed, which first entail constructing matrix $\mathbf{C_p}$ related to the desired/observed correlation ρ_{ep} to be preserved:

$$\mathbf{C_p} = \begin{pmatrix} 1 & \rho_{ep} \\ \rho_{ep} & 1 \end{pmatrix} \qquad (6)$$

Then, matrices **L** and $\mathbf{L_p}$ can be obtained as lower triangular matrices from the Cholesky decomposition (Press et al., 1990) of matrices **C** and $\mathbf{C_p}$ respectively.

The new matrix $\mathbf{X_1}$, which features a correlation matrix equal to $\mathbf{C_p}$ in eq. (6), has to be calculated as:

$$\mathbf{X_1} = \mathbf{X} \cdot \left(\mathbf{L_p} \cdot \mathbf{L}^{-1}\right)^T \qquad (7)$$

The elements of each column of matrix **X** have to be reordered in order to have the same sorting as the elements of the corresponding column of matrix $\mathbf{X_1}$, thus producing the matrix $\mathbf{X_2}$. In this manner, the matrices $\mathbf{X_2}$ and $\mathbf{X_1}$ will have the same rank correlation matrix, and, consequently, similar (Pearson) correlation matrices. Since the application of step 2 simply modifies the sorting of the *T* and *I* values and does not change the values themselves, it preserves the exact form of the marginal distributions on these variables, as it comes from step 1.

Method 2: Gaussian copula

Unlike method 1 that is applied to pulse durations and intensities that have already been generated, method 2 precedes the generation. A further difference lies in the fact that method 2 does not entail matrix operations.

Method 2 is known as the method of the Gaussian copula (Nelsen, 1999). It is based on generating *n* couples of auxiliary random variables y_1 and y_2 with average values equal to 0 and standard deviations equal to 1 through the bivariate normal distribution (eqs. 8 and 9):

$$f(y_1, y_2) = \frac{1}{\sqrt{(2\pi)^2|\mathbf{\Sigma}|}} \mathrm{e}^{\left[-\frac{1}{2(1-\dot{\rho})}\left(y_1^2+y_2^2-2\dot{\rho}y_1y_2\right)\right]}, \qquad (8)$$

$$\Sigma = \begin{pmatrix} 1 & \dot{\rho} \\ \dot{\rho} & 1 \end{pmatrix} \quad (9)$$

where $\dot{\rho}$ represents correlation between y_1 and y_2.

For each of the n values generated for y_1, the value F_1 of the cumulative probability of the marginal distribution can be calculated. In a similar way, for each of the n values generated for y_2, the value F_2 of the cumulative probability of the marginal distribution can also be calculated.

The n values of F_1 and F_2 can be used to sample the probability distributions chosen for pulse durations and intensities (eq. 4) and then to obtain n couples of T and I. As a result of correlation $\dot{\rho}$ imposed between y_1 and y_2, a certain degree of correlation is also imposed on T and I. In particular, the resulting correlation between T and I is a monotonous function of $\dot{\rho}$. Iterative methods can then be applied in order to determine the suitable value of $\dot{\rho}$ that yields the expected correlation ρ_{ep} to be preserved between T and I.

Method 3

Method 3 is a modified version of method 2. Similar to the original method, it does not require matrix operations and is based on the use of the bivariate normal distribution (eqs. 8 and 9). However, following from method 1, it is applied after a preliminary first step, in which n uncorrelated couples of T and I are generated.

Then, n couples of y_1 and y_2 with correlation equal to $\dot{\rho}$ are also generated. As in method 2, the corresponding values of F_1 and F_2 can be obtained. T and I can be reordered using the same sorting as F_1 and F_2 respectively. As a result of correlation $\dot{\rho}$ imposed between y_1 and y_2, a certain degree of correlation is also imposed on T and I. Iterative methods can be applied in order to determine the suitable value of $\dot{\rho}$ that yields the expected correlation ρ_{ep} to be preserved between T and I.

Parameter estimation

The set of parameters of a Poisson model for pulse generation, in which T and I are generated as independent random variables, by using either the beta (eq. 2) or the gamma distribution (eq. 3), or using any kind of 2 parameter probability distribution, is 5: one parameter (λ) for pulse arrival and two parameters for either of T or I. If correlation between T and I needs to be accounted for, the number of parameter increases to 6 and correlation ρ_{ep} is the sixth parameter of the model.

As it was done by some authors (Alvisi et al., 2003; Buchberger et al., 2003; Creaco et al., 2015), the generic day of the month can be subdivided into a certain number of time slots (e.g., 12 bihourly) for parameter estimation. Robust models can then be obtained by allowing pulse arrival-related parameter λ to take on a different value in each daily time slot. Each of the other parameters (in this case, the parameters related to T and I and correlation ρ_{ep}), instead, is allowed to take on a single value valid for all the time slots.

For estimating the parameters of the Poisson model, the method of the moments (Hall, 2004) was used, which consists in setting the values of the parameters equal to the corresponding values in the measured pulses.

Applications

Case studies

By making calculations on data collected during an experimental campaign in some households in Milford, Buchberger et al. (2003) were able to reconstruct, with one second time step resolution, the water demand pulses which were taking place in these households in the period from April to October 1997. The data made available by the Authors concern pulse duration T, intensity I and volume $V=T \cdot I$.

As case study in this work, the indoor water demand pulses recorded in one of the households, i.e. household 2, in the month of April were selected. This case study has already been chosen by Creaco et al. (2015) on the basis of the regularity of the daily water consumption. The basic statistical parameters of measured water consumption variables z, T, I, V, D and ρ are reported in Table 2.

The modelling framework of this paper is aimed at investigating the extent to which the methods described in this paper for preserving correlation lend themselves to being used inside Poisson models.

Overall, four Poisson models were constructed and compared with two benchmark Poisson models, hereinafter indicated as models A and B and drawn from the work of Creaco et al. (2015). In particular, model A features pulse durations and intensities being generated through the bivariate lognormal distribution. This enables correlation to be obtained between the two variables in this model. Model B differs from model A in that pulse durations and intensities are generated as independent (uncorrelated) variables by making use of the lognormal distribution. The four models constructed in this work are models C-1, C-2, C-3, and D. Models C-1, C-2 and C-3 differ from model A in the way correlation is preserved. In fact, unlike model A, these models feature adoption of one of the three methods described hereinabove (methods 1, 2 and 3 respectively). Model D differs from model C-3 in the generation of pulse durations and intensities, which take place through the beta (eq. 2) and gamma (eq. 3) distributions respectively, instead of the lognormal distribution.

Overall, applications consisted in 3 phases for each test:

phase 1 – parameter assessment;

phase 2 – generation of synthetic water demand pulses;

phase 3 – analysis of the results of the models and comparison with the observed data.

Results

Phase 1

The results of phase 1 for models C-1, C-2, C-3 and D are reported in Table 4 and 5. The results for models A and B, instead, can be found in the work of Creaco et al. (2015). For the analysis of these tables, it has to be recalled that only λ is parametrized in 12 daily time slots; the other parameters, instead, are assigned a single daily value. The data reported in Tables 4 and 5, related to models C-1, C-2, C-3 and D, were obtained by applying the method of the moments (Hall, 2004). As was expected, Table 4 shows that the λ values obtained in the various time slots are identical for models C-1, C-2, C-3 and D. This is due to the fact that all these models deal with pulse arrival generation in the same way and they only differ in the way pulse intensities and durations are generated. In model D, in which the beta distribution is used to generate pulse durations, the interval $[x_{min},x_{max}]$ was set to [0,850] following analysis of the experimental pulses.

Phase 2

The models calibrated in phase 1 were then applied in order to create synthetic demand pulses for one month (i.e. 30 days - April) for each test. In order to account for the influence of the random seed, each generation was repeated 100 times.

In each of the models which use one of the methods described hereinabove to preserve correlation (models C-1, C-2, C-3 and D), the method was applied once at the end of each of the 100 monthly generations of pulses.

As an example, Figure 1 shows the single realization of the simulated total demand for a typical day at the scale of 1 sec, obtained by using model C. As expected, the figure shows a higher concentration of the pulses in the morning and in the late afternoon, when the household occupants usually get up and get back home after work. Very few pulses are instead generated at nighttime. This is a direct consequence of the λ values in the bihourly time slots.

Phase 3

A first analysis was made concerning the basic statistical parameters of water consumption variables z, T, I, V and ρ derived from the pulses generated by models A-D, in comparison with those of the measured pulses (see Table 1). This table shows that all the models reproduce well mean(z), mean(T), var(T), mean(I) and var(I). This is a direct consequence of the goodness of the method of the moments for parameter calibration.

Correlation ρ is only preserved in those models which are constructed considering the correlation, i.e. models A, C-1, C-2, C-3 and D. As expected, model B leads to ρ=0, since the pulse duration and intensity are generated independently from each other in this model. Similar considerations can be made with regard to cov(T,I). With regard to ρ, the analysis of Table 1 proves that the methods adopted in models C-1, C-2, C-3 and D have similar effects to the use of the bivariate distribution in model A. The advantage for these methods of being also applicable in the cases (see model D) where no bivariate distribution is available, that is when T and I are represented by two different kinds of marginal probability distributions, must also be highlighted. As for pulse volume V, only the models that consider correlation, i.e. A, C-1, C-2, C-3 and D, are those that return consistent values of mean(*V*), i.e. values of mean(*V*) close to the value mean(*V*)=9.45 L of the measured pulses. As a consequence of this, the same models provide consistent values of mean (*D*), i.e. values of mean(*D*) close to the value of mean(*D*)=442 L/day associated with the measured pulses.

Another test was then carried out to compare the synthetic water demand pulses generated by means of the models with the measured water demand pulses in terms of overall daily water demand volume *D*. In particular, the total synthetic water demand volume *D* was calculated for each day in the generic one month long pulse generation of each test. Then, the cumulative frequency curve was constructed reporting, for each value of *D*, the Weibull cumulative frequency *F* of days in the month that feature a value of the overall daily water demand volume lower than or equal to *D*. Since each model application comprises 100 one-month long pulse generations, a band of synthetic cumulative frequencies was then obtained for each test. For each test, the band upper envelope (BUE), lower envelope (BLE) and mean value (BMV) of the 100 cumulative frequency curves were determined for all the models. The cumulative frequency of the measured daily water demand volume (ECF) was also calculated.

The graphs in Figure 2 report BUE, BLE and BMV obtained using the various models as well as ECF. Analysis of the graphs shows that, as already highlighted by Creaco et al. (2015), the BMV obtained with model A (model that takes into account the mutual dependence of pulse intensity and duration by means of the bivariate distribution) follows ECF much more closely than that obtained with model B (model that neglects the mutual dependence of pulse intensity and duration). Furthermore, all the data points of ECF lie inside the band of cumulative frequency obtained with model A. Only a few ECF data points, instead, are found inside the band of cumulative frequency obtained with model B. This attests to the better capability of model A to generate water demand pulses that are consistent with the observed demand pulses in terms of overall daily water demand volume. Figure 2 also shows that models C-1, C-2 and C-3, which use the methods described in this paper to preserve correlation, have an almost identical performance to model A. The change in the distributions used to represent the pulse durations and intensities does not affect results significantly (see results of model D with method 3). This is due to the fact that, in the case study considered, the beta and gamma probability distributions fit the measured pulse durations and intensities in a similar way to the lognormal distributions used in method A.

Conclusions

This paper presented the application of three different methods that can be used to preserve correlation between duration and intensity of water demand pulses. Whereas the first two methods are derived from the known statistical approaches, the third was newly developed in this work. Applications showed that the three methods yield similar results to those previously reported by Creaco et al. (2015) with the advantage of being applicable with any marginal distributions to represent the duration and the intensity.

Subsequently, the following consideration can be made as far as the preservation of duration/intensity correlation in synthetically generated pulses is concerned. In order to preserve this correlation, a bivariate probability distribution has to be chosen to represent pulse durations and intensities, as was shown by Creaco et al. (2015), when the marginal distributions chosen to represent these two variables can be inserted in the framework of such modelling. If the latter condition does not hold, as it is the case with the probability distributions chosen by some authors (e.g., Guercio et al., 2001; Garcia et al., 2004), one of the three methods described in this paper can be profitably applied if correlation needs to be preserved.

Though the three methods have similar effectiveness, methods 1, based on the Iman-Canover (1982) method, and method 3, novel variation in the Gaussian Copula (Nelsen, 1999), may turn out to be more attractive for engineers. In fact, they can be easily implemented downstream of the standard methods for generating independent pulse durations and intensities, in order to impose the correlation by post-processing the results of the latter.

After the work of Creaco et al. (2015) showed that pulse generation models fitted with duration/intensity correlation have advantages in comparison with traditional models and after this work has shown how correlation can be obtained in a more general way, future work will be dedicated model parameter assessment on the basis of smart meter readings. Compared to the current parameterization, which is based on the method of the moments and stringently requires knowledge of real pulse features, the new parameterization will significantly extend the applicability of the pulse generation models fitted with duration/intensity correlation.

Acknowledgements

The authors thank Prof. S.G. Buchberger for providing the data demand for the Milford households. This study was carried out as part of the ongoing projects: (i) “iWIDGET” (Grant Agreement No 318272), which is funded by the European Commission within the 7th Framework Programme, (ii) the ongoing PRIN 2012 project “Tools and procedures for an advanced and sustainable management of water distribution systems”, n. 20127PKJ4X, funded by MIUR, and (iii) under the framework of Terra&Acqua Tech Laboratory, Axis I activity 1.1 of the POR FESR 2007–2013, project funded by the Emilia-Romagna Regional Council (Italy) (http://fesr.regione.emiliaromagna.it/allegati/comunicazione/la-brochure-dei-tecnopoli).

References

Alcocer-Yamanaka, V.H., Tzatchkov, V., and Buchberger, S.G., 2006. Instantaneous Water Demand Parameter Estimation from Coarse Meter Readings. In: Proceedings of the Water Distribution Systems Analysis Symposium 2006, pp. 1-14.

Alcocer-Yamanaka, V.H., and Tzatchkov, V.G., 2012. Modeling of Drinking Water Distribution Networks Using Stochastic Demand. Water Resources Management, 26, 1779-1792.

Alvisi, S., Franchini, M., and Marinelli, A., 2003. A stochastic model for representing drinking water demand at residential level. Water Resources Management, 17 (3), 197–222.

Alvisi, S., Ansaloni, N., Franchini, M. 2014. Generation of synthetic water demand time series at different temporal and spatial aggregation levels. Urban Water Journal, 11 (4), 297-310.

Blokker, E.J.M., Vreeburg, J.H.G., and van Dijk, J.C., 2010. Simulating residential water demand with a stochastic enduse model. Journal of Water Resources Planning and Management, 136 (1), 375–382.

Buchberger, S.G., Carter, J.T., Lee, Y.H., and Schade, T.G., 2003. Random Demands, Travel Times and Water Quality in Dead-ends, prepared for American Water Works Association Research Foundation, AwwaRF Report No. 294, 2003, 470 pages.

Buchberger, S.G., and Wells, G.J., 1996. Intensity, duration and frequency of residential water demands. Journal of Water Resources Planning and Management, 122 (1), 11–19.

Buchberger, S.G., and Wu, L., 1995. Model for instantaneous residential water demands. Journal of Hydraulic Engineering, 121 (3), 232–246.

Creaco, E. , Farmani, R., and Kapelan, Z., Vamvakeridou-Lyroudia, L., Savic, D., 2015. Considering the mutual dependence of pulse duration and intensity in models for generating residential water demand. Journal of Water Resources Planning and Management, doi: 10.1061/(ASCE)WR.1943-5452.0000557.

Garcıa, V.J., Garcıa-Bartual, R., Cabrera, E., Arregui, F., and Garcıa-Serra, J. 2004. Stochastic model to evaluate residential water demands. Journal of Water Resources Planning and Management, 130 (5), 386–394.

Guercio, R., Magini, R., and Pallavicini, I., 2001. Instantaneous residential water demand as stochastic point process. In: Water resources management, C. A. Brebbia et al. eds., WIT Press, Southampton, U.K., 129-138.

Hall, A.R., 2004. Generalized Method of Moments. Advanced Texts in Econometrics, OUP Oxford, ISBN: 0198775202, 9780198775201.

Iman, R.L., and Conover, W.J., 1982. A distribution-free approach to inducing rank correlation among input variables. Communication in Statistics, 11, 311–334.

Johnson, R.A. and Bhattacharyya, G.K., 1992. Statistics: Principles and Methods. 2nd Edition. Wiley, 1992.

Nelsen, R.B., 1999. An Introduction to Copulas. Springer, New York, ISBN 0-387-98623-5.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T., 1990. Numerical recipes: the art of scientific computing. New York: Cambridge University Press.

Walski M., Chase D., Savic D., Grayman W., Beckwith S. and Koelle E. (2003). Advanced water distribution modelling and management. Waterbury, CT: Haestad Press.

Tables

Table 1. Arrival time τ, duration T and intensity I of the pulses generated by the Poisson model.

τ (s)	T (s)	I (L/s)
τ_1	T_1	I_1
τ_2	T_2	I_2
…	…	…
τ_i	T_i	I_i
…	…	…
τ_{n-1}	T_{n-1}	I_{n-1}
τ_n	T_n	I_n

Table 2. Basic statistical parameters of water consumption variables z, T, I, V, D and ρ derived from the measured pulses and from the pulses generated by the models.

pulse type	mean (z) [s^{-1}]	mean (T) [sec]	var (T) [sec^2]	mean (I) [L/s]	var (I) [L^2/s^2]	mean (V) [L]	mean (D) [L/day]	cov (I,T) [L^2]	ρ [-]
measured	0.00054	56	12743	0.106	0.00742	9.45	442	3.50	0.36
model A	0.00054	56	11822	0.106	0.00730	9.39	438	3.48	0.37
model B	0.00054	56	12321	0.106	0.00737	5.95	278	0.03	0.00
model C-1	0.00054	56	12321	0.106	0.00737	9.26	431	3.34	0.35
model C-2	0.00054	56	12049	0.105	0.00712	9.17	427	3.27	0.35
model C-3	0.00054	56	12321	0.106	0.00737	9.27	432	3.36	0.35
model D	0.00054	56	12723	0.106	0.00744	9.44	440	3.50	0.36

Table 3. Features of the models used in this work.

	distribution for duration	**distribution for intensity**	**parameter estimation method**	**correlation method**
model A	bivariate lognormal	bivariate lognormal	moments	Creaco et al. (2015)
model B	lognormal	lognormal	moments	no correlation
model C-1	lognormal	lognormal	moments	method 1
model C-2	lognormal	lognormal	moments	method 2
model C-3	lognormal	lognormal	moments	method 3
model D	beta	gamma	moments	method 3

Table 4. Calibrated values of λ for the various time slots, for models C-1, C-2 and C-3, D.

models	**slot 0h-2h**	**slot 2h-4h**	**slot 4h-6h**	**slot 6h-8h**	**slot 8h-10h**	**slot 10h-12h**	**slot 12h-14h**	**slot 14h-16h**	**slot 16h-18h**	**slot 18h-20h**	**slot 20h-22h**	**slot 22h-24h**
C-1,C-2, C-3	0.000065	0.000028	0.001056	0.000991	0.000579	0.000356	0.000130	0.000333	0.001023	0.001032	0.000745	0.000157
D	0.000065	0.000028	0.001056	0.000991	0.000579	0.000356	0.000130	0.000333	0.001023	0.001032	0.000745	0.000157

Table 5. Calibrated values of daily parameters for models C-1, C-2 and C-3, D.

parameters for C-1,C-2,C-3	**values**	**parameters for D**	**values**
μ_{lnT}	3.22	α	0.17
σ_{lnT}	1.27	β	2.33
μ_{lnI}	-2.50	θ	0.07
σ_{lnI}	0.71	κ	1.51
ρ_{ep}	0.36	ρ_{ep}	0.36

Figures

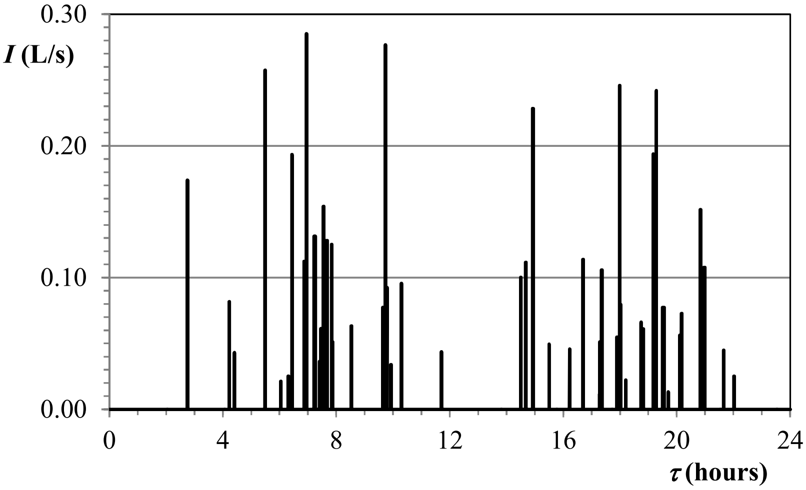

Figure 1. Model C - single realization of the simulated total demand for a typical day at the scale of 1 sec.

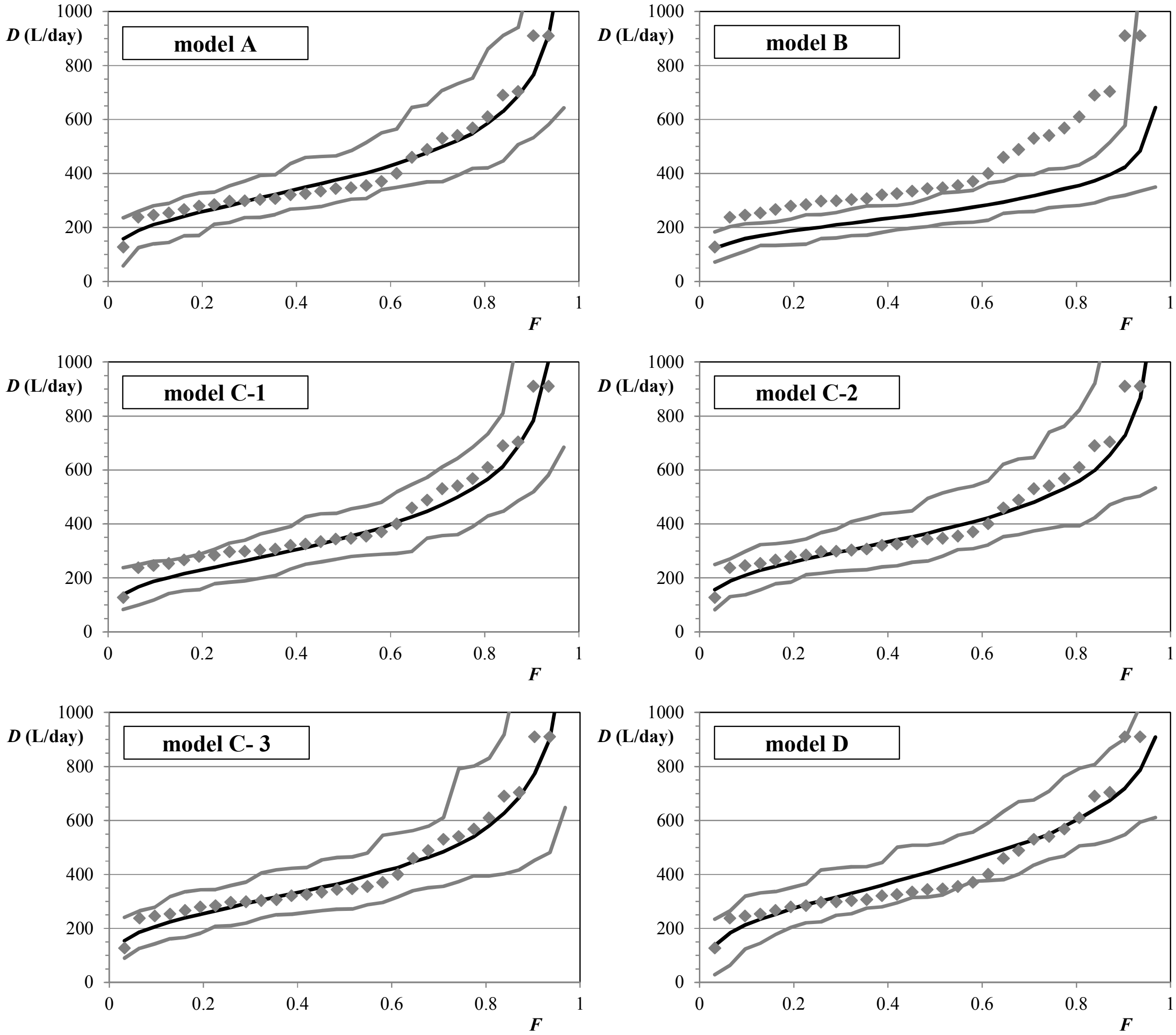

Figure. 2. Upper (BUE) and lower (BLE) envelopes (grey lines) and mean value (BMV) (black line) of the band of Weibull cumulative frequencies *F* of daily water demand *D* produced by models A, B, C-1, C-2, C-3 and D, in comparison with the daily water demand cumulative frequency calculated starting from the measured data (ECF) (dots).