1   *Article*

2   Generating scenarios of cross-correlated demands for
3   modelling water distribution networks

4   **Roberto Magini [1], Maria Antonietta Boniforti[2], Roberto Guercio [3]**

5   [1] Dept. of Civil, Building, and Environmental Engineering, Sapienza University of Rome, Via Eudossiana 18,
6   00184 Rome, Italy; roberto.magini@uniroma1.it
7   [2] Dept. of Civil, Building, and Environmental Engineering, Sapienza University of Rome, Via Eudossiana 18,
8   00184 Rome, Italy; antonietta.boniforti@uniroma1.it
9   [3] Dept. of Civil, Building, and Environmental Engineering, Sapienza University of Rome, Via Eudossiana 18,
10   00184 Rome, Italy; roberto.guercio@uniroma1.it
11   *   Correspondence: roberto.magini@uniroma1.it ; Tel.: +39-06-44585-057

12   **Abstract:** A numerical approach for generating a limited number of water demand scenarios and
13   estimating their occurrence probabilities in a Water Distribution Network (WDN) is proposed. This
14   approach makes use of the demand scaling laws in order to consider the natural variability and
15   spatial correlation of nodal consumptions. The scaling laws are employed to determine the statistics
16   of nodal consumption as a function of the number of users and the main statistical features of the
17   unitary user's demand. Besides, consumption at each node is considered to follow a Gamma
18   probability distribution. A high number of groups of cross-correlated demands, i.e. scenarios, for
19   the entire network were generated using Latin Hypercube Sampling (LHS) and the numerical
20   procedure proposed by Iman and Conover. The Kantorovich distance is used to reduce the number
21   of scenarios and estimate their corresponding probabilities, while keeping the statistical information
22   on nodal consumptions. By hydraulic simulation, the whole number of generated demand scenarios
23   was used to obtain a corresponding number of pressure scenarios on which the same reduction
24   procedure was applied. The probabilities of the reduced scenarios of pressure were compared with
25   the corresponding probabilities of demand.

26   **Keywords:** uncertain water demand; scaling laws; scenario generation; scenario reduction; water
27   distribution networks; hydraulic simulation
28

29   1. Introduction

30   The conventional modelling of WDNs normally is based on a deterministic approach, not
31   merely with regard to the geometrical and hydraulic features, but also with respect to the demand
32   loadings [1].   However, water demand, being influenced by many factors, i.e. type of users, socio-
33   economic conditions, geographic location with its climate, seasonal fluctuation of weather, water
34   fixtures technology, policies in water management, tariff, is subject to a natural variability. Surely,
35   the variability of the demand represents the major source of uncertainty, which affects the overall
36   reliability of the model for the assessment of the spatial and temporal distribution of pressure heads
37   as well as for the evaluation of the water quality in the different pipes. Uncertainty produced by
38   randomicity of demand assumes a different importance in relation to the spatial and temporal scales
39   that are considered in modelling the network. Obviously, they become more and more relevant as
40   the finer scales are reached, that is small groups of users and instantaneous demands are considered.
41   As stated by Bargiela & Sterling [2], it is possible to obtain accurate predictions for the network as a
42   whole, but estimating nodal consumptions for nodes where the population is low is more difficult.

43   Thus, considering and quantifying uncertainty of water demand could allow to associate an
44   acceptable probability/level of risk to the results from hydraulic and optimization models of WDNs
45   at different temporal and spatial scales: a more robust design and control of these systems can be
46   realized, and obvious opportunities can arise in dimensioning pipes, formulating water balances,
47   controlling system's components, identifying and quantifying leakages.

48   An approach to dealing with uncertainty of demand consists in explicitly considering different
49   possible realizations of its value at the nodes of a WDN, i.e. different loading scenarios and
50   associating to each of them a measure of their probability.   So, it is possible to find a feasible solution
51   which is also as close as possible to the optimum for all the scenarios: this scenario-based approach
52   is known as robust optimization [3]. In the same way it is possible also to derive WDN reliability [4]
53   or to localize leakages [5] or to map pressure-heads for a real-time control [6], all under uncertain
54   demand conditions. Different possible scenarios, which include various aspects, such as peak flows,
55   fire conditions at certain nodes, or pipe breakage, and the corresponding probabilities of occurrence
56   could be obtained by consulting a panel of experts. However, this solution can have strong limitations
57   in such a mathematically sensitive problem and can lead to arbitrary solutions.

58   At this aim this paper focuses on defining an objective methodology for the generation of
59   demand scenarios. In the literature the issue of generating demand scenarios has been faced in [7]
60   where uncertain future water consumptions are modeled using probability density functions (PDFs)
61   assigned in the problem formulation phase. Scenarios with correlated nodal demands are generated
62   using LHS and the procedure suggested by Iman and Conover [8]. All nodal demands follow a
63   Gaussian PDF with coefficient of variation $Cv = 0.10$. The correlation coefficient between any two
64   nodal demands is assumed equal to 0.50, as done by Tolson et al. [9]. In [10] different demand
65   scenarios for a WDN are derived combining demand values with specific return period at each node.
66   The probability of each scenario is obtained considering a Multivariate Normal Distribution (MVN).
67   The correlation between demand was found to significantly affect the occurrence probability of the
68   demand scenarios. Also, Eck et al. [11] generated demand scenarios considering a MVN distribution.
69   They assumed a prior estimate of the mean values and covariance matrix of water demand from a
70   preliminary analysis. An original, but computationally expensive, approach for estimating the
71   probability of a given demand scenario is based on the use of the contingency tables [12]. This is a
72   non-parametric method in which the various random variables are divided into classes and their
73   marginal probabilities are calculated through the frequency of occurrence of each class. The joint
74   probabilities are evaluated by counting the occurrences of the simultaneous classes.

75   In this paper a numerical approach for generating a limited number of water demand scenarios
76   and estimating their occurrence probabilities in a WDN is presented. Scaling laws [1,13,14] are used
77   to evaluate the statistics of aggregated demand at each node of the network, starting from the
78   statistics of demand of the unitary user. The Apulian WDN [15] is considered as a case-study and
79   two different hypotheses are made for the number of users in the nodes. For each of these hypotheses
80   a large number of cross-correlated demands, i.e. scenarios, was generated using LHS from Gamma
81   PDFs and a procedure based on the approach proposed by Iman and Conover [8]. The Kantorovich
82   distance [16] is used to reduce the number of scenarios and estimate their corresponding
83   probabilities. In this way, the limited number of network demand scenarios maintain the statistical
84   information on the demand and the effects on the performance of WDN. In the design and control of
85   WDNs, demand scenarios are mostly functional to the evaluation of service conditions and
86   specifically pressure-heads in the nodes.   In order to evaluate how uncertainty of demand is
87   conveyed to pressure-head field a demand-driven hydraulic model was also applied to solve Apulian
88   network. Many pressure scenarios were obtained, and the same reduction procedure was employed.

89   The probabilities of the reduced scenarios of pressure were compared with the corresponding
90   probabilities of the demand.

91   **2. Description of methodology**

92   *2.1. Scaling laws*

93   In the first phase of the proposed methodology the statistical features of WDN nodal demand
94   are derived using the scaling laws [1,13]: i.e. the first and second statistical moments and cross-
95   correlation are evaluated, and the probability distributions defined. Input data are the nodal number
96   of consumers and the statistical features of water demand of the typical single user, unitary user. The
97   statistical parameters of the unitary user' demand can be derived either by monitoring, or simulating
98   by descriptive models, such as End-Uses models [17] or PRP models preserving correlation [18].

99   We consider a network with $i = 1,2,3, \dots, N$ nodes and $n_i = n_1, n_2, n_3, \dots, n_N$ users at each node.
100   The demand of a unitary user is described by an ergodic stationary stochastic process, with mean μ_1,
101   variance σ_1^2 and cross-correlation coefficient between each couple of single-user ρ_1. The whole
102   nodal demand, that is the sum of all users' consumption at each node, is a realization of the stochastic
103   process whose statistics are dependent on the number of nodal users. The expected value $E[\mu_{n_i}]$ for
104   the mean of the aggregated process at the *ith* node is given by:

$$E[\mu_{n_i}] = n_i \cdot E[\mu_i],\tag{1}$$

105   and the expected value for the variance $E[\sigma_{n_i}^2]$ at the same node, neglecting the bias that can be
106   caused when using small the demand series (short observation periods) [13], is given by:

$$E[\sigma_{n_i}^2] = n_i{}^2 \cdot E[\rho_1] \cdot E[\sigma_1^2] + n_i \cdot \left[1 - E[\rho_1]\right] \cdot E[\sigma_1^2]\tag{2}$$

107   Equation (2) shows that the expected value of the variance of the aggregated process depends
108   on $E[\rho_1]$: if demands are perfectly correlated in space, i.e. $E[\rho_1]$ is equal to one, equation (2) is
109   simplified into:

$$E[\sigma_{n_i}^2] = n_i{}^2 \cdot E[\sigma_1^2]\tag{2a}$$

110   if demands are uncorrelated in space, i.e. $E[\rho_1]$ is equal to zero, equation [2] is simplified into:

$$E[\sigma_{n_i}^2] = n_i \cdot E[\sigma_1^2]\tag{2b}$$

111   For partially correlated demands a power law $E[\sigma_{n_i}^2] = n_i^\alpha E[\sigma_1^2]$, with α scaling variance
112   exponent, has been also derived [1,14]. Therefore, a complete statistical characterization of demand
113   requires not only the definition of its mean and variance, but also the definition of the correlation
114   between demands of each couple of users and groups of users. The cross-correlation refers to the
115   similarity between demand patterns from different consumers or from different nodes. This
116   parameter was proved to be not negligible [19] and to affect the hydraulic performance of a WDN as
117   well as the cost to achieve a desired level of reliability: it was verified that higher cross correlations
118   lead to higher pressure fluctuations, which have negative impacts on the reliability of the WDN [20].
119   Following the same assumption and notation, the expected value for the cross-correlation between
120   all nodes of the network is represented by a N-by-N square matrix, whose elements are given by:

121

122

$$E\left[\rho_{n_i n_j}\right] = \frac{E\left[cov_{n_i n_j}\right]}{E[\sigma_{n_i}] \cdot E[\sigma_{n_j}]} = \frac{n_i n_j E[\rho_1]}{\left[n_i\left(1 + E[\rho_1](n_i - 1)\right)\right]^{1/2} \cdot \left[n_j\left(1 + E[\rho_1](n_j - 1)\right)\right]^{1/2}} \quad (3)$$

123

124

125   with $i, j = 1,2,3,\dots,N$, and where, for example, $\rho_{n_1 n_2}$ is the cross-correlation coefficient between the
126   demand of n_1 aggregated users at node 1, and the demand of n_2 aggregated users at node 2.
127   Through equations (1,2,3) the nodal demand statistics of the network and the correlation
128   structure between them are entirely defined.

129

130   *2.2. Generation of scenarios*
131   In simulation and optimization problems several methods exist to cope with uncertainty [21,22].
132   If we do not know exactly input data, in our case water demand at the nodes of a WDN, because they
133   can assume different values and then many combinations of them are possible, we are dealing with
134   scenarios. But, if the statistical features of uncertain data are known, numerical solutions can be
135   obtained by approximating the probability distribution function with discrete distributions having a
136   finite number of outputs, again referred to as scenarios.
137   Then, the second phase of the present approach consists in the generation of a large number of
138   water demand scenarios for a WDN, based on the statistics estimated by the scaling laws and making
139   the hypothesis of Gamma-distributed water demand at each node. The knowledge of the scaling laws
140   of the statistical moments and the type of the probability distributions of water demand in relation
141   to the number of users, prove to be a useful tool to face the inherent uncertainty of demand and in
142   particular to address the optimization problems. Using Gamma distribution for demand is
143   supported, when the number of aggregated users is high enough or time resolution is greater than
144   five minutes, by measurements in Latina case-study [1]. Furthermore, in a recent work Kossieris and
145   Makropoulos [23] investigated the performance of ten probabilistic models showing that both
146   Gamma and Weibull distributions can be used to adequately describe the nonzero water demand
147   recorded at different time scales.
148   Scenarios can be generated following different methods: by matching a small set of statistical
149   properties, e.g. moments [24,25] or simulating some defined mathematical process (e.g. Brownian
150   motion) or sampling from known distributions [26]. For scenarios with a large number of variables
151   correlated in between, sampling from the joint distribution is not usual for the difficulty in defining
152   the distribution itself. An alternative to specifying the joint distribution is to make use just of the
153   marginal distributions and the linear or rank correlation matrix.   If nothing is known about the form
154   of the joint distribution a coupling procedure can be used: in this case the generated scenarios will
155   respect an arbitrary dependency structure based on the procedure followed.
156   Each demand scenario D_u is defined here as a set or combination of nodal demand values
157   occurring simultaneously in the WDS. It can be represented by the N dimensional vector:

158

$$D_u = \left[d_{1,u}, d_{2,u}, \dots d_{N,u}\right], \qquad \forall\, i = 1, 2, \dots, N \ \ and \ \ u = 1, 2, \dots, S \quad (4)$$

159   where $u = 1,2,\dots,S$ is the index identifying the different scenarios, and $d_{i,u}$ is the demand at
160   node i for the uth scenario and D_u depicts a one-dimensional stochastic data process.

161

162   A sampling method from the marginal distribution based on the LHS and the Iman Conover
163 approach [8] is followed. The restricted pairing technique by Iman and Conover induces rank
164 correlation between the given marginals by shuffling finite-size samples from each of them. The
165 appropriate shuffle is determined by ranking the input samples the same as in a reference sample
166 with the desired rank correlation. The complete procedure, that will be described in the following, is
167 quite straightforward because it requires only the Cholesky decomposition, some matrix algebra, and
168 the final rearrangement of the original uncorrelated sample
169
170   *2.2.1. Procedure for generating scenarios*

171 **Step 1.** Create a random (*S, N*) dimensional matrix **Z*,** containing S Latin Hypercube Samples of
172   size *N* from a standardized normal distribution, where *S* is the number of scenarios and *N* the
173   number of the demand nodes in WDN. For this purpose the Matlab function *lhsnorm* was
174   used. Owing to the finite size of the samples their correlation matrix **R*** does not coincide with
175   the identity matrix **I**, that is they are not independent. Then, the **lower triangular** Cholesky
176   decomposition is applied to induce the desired correlation. Specifically:

177    $\mathbf{I} = \mathbf{C} \cdot \mathbf{C^T}$
178    $\mathbf{E} = \mathbf{R^*} \cdot \mathbf{R^{*T}}$
179    $\mathbf{Z} = \mathbf{Z^*} \cdot \mathbf{C} \cdot \mathbf{E^{-1}}$

180   and the (*S, N*) dimensional matrix **Z** of perfectly independent *S* samples of size *N* from a
181   standardized normal distribution is obtained. In order to obtain the Cholesky root the Matlab
182   function *chol* was used.

183 **Step 2.** Create a random (*S, N*) dimensional matrix containing *S* standardized normal samples with
184   the correlation matrix **Corr** from the scaling laws for nodal demand. To this aim the desired
185   correlation is induced in **Z** also applying the **lower triangular** Cholesky decomposition, that is:
186    $\mathbf{Corr} = \mathbf{P} \cdot \mathbf{P^T}$
187    $\mathbf{E} = \mathbf{R^*} \cdot \mathbf{R^{*T}}$
188    $\mathbf{G} = \mathbf{Z} \cdot \mathbf{P} \cdot \mathbf{C^{-1}}$

189 **Step 3.** Transform matrix **G** in the (*S, N*) dimensional matrix **D*** complying with the desired
190   marginal distributions at each demand node. Transformation is based on the inverse
191   Cumulative Distribution Function, CDF, of the desired marginals, F_i. Specifically, for N non-
192   normal random samples **D*** = [**D*₁, D*₂ ,…,D*ᵢ**] , $1 \le i \le N$, with desired CDF, the following
193   equation holds:

$$D_i^* = F_i^{-1}\big(\Phi(G_i)\big) \tag{5}$$

195   where $\Phi(G_i)$ is the CDF of the i*th* samples of G and is uniformly distributed. $\Phi(G_i)$ can also
196   be interpreted as a realization from the Gaussian copula. Applying the inverse CDF F_i^{-1} to
197   the uniform random variable $\Phi(G_i)$ ensures that D_i^* is distributed according to Φ_i .
198   Unfortunately, transformation in Equation 1 is non-linear, therefore the correlation matrix
199   **Corr*** of **D*** is no equal to the desired correlation matrix **Corr**.

200 **Step 4.** Apply the Iman-Conover algorithm proposed by Ekström [27] in order to get a better
201   approximation of the desired correlation matrix **Corr** for the (*S, N*) matrix of nodal demand
202   scenarios **D***. The algorithm is described in the following steps:

203

204   4.1 Calculate lower triangular Cholesky decomposition **V** of **Corr**, i.e. **Corr = V·V**^{T} .

205   4.2 Calculate lower triangular Cholesky decomposition **Q** of **Corr***, i.e. **Corr* = Q·Q**^{T}.

206   4.3 Obtain **T** such that **Corr = T·Corr·T**^{T} , can be calculated as **T = V·Q**$^{-1}$.

207   4.4 Obtain the matrix **ScoreD*** by rank-transforming **D** and convert to van der Waerden scores

208   defined as $F_i^{-1}\big(\Phi(i/(N+1)\big)$ where ϕ is the CDF of the standard normal distribution, i is

209   the assigned rank and N is the total number of samples.

210   4.5 Calculate the target scores matrix **ScoreD = ScoreD*·T**^{T}.

211   4.6 Match up the rank pairing in **D*** according to **ScoreD,** obtaining the new (*S,N*) dimensional

212   matrix **D** containing the *S* scenarios of the *N* nodal demand in the WDS. The N samples are

213   distributed according to the desired marginals and their correlation matrix is close to the

214   correlation matrix derived from the scaling laws.

215   *2.3. Scenario reduction*

216   With the scenario generation procedure, we obtain a great number of pictures each of which

217 represents a single snapshot of the whole water demand in the network. The higher the number of

218 scenarios generated the better is the description of the variability of water demand in the WDS. But

219 it is not possible to manage such a large number of scenarios to deal with stochastic or robust

220 optimization problems. Moreover, the probability associated with each of them is not very significant.

221 We have to reduce the scenarios and at the same time determine, for the reduced scenarios, a

222 significant weight representative of their possibility to be realized. Then, the goal of scenario

223 reduction is to approximate the discrete distribution of the generated scenarios with another discrete

224 distribution having fewer elements. At this point, the choice of the number of scenarios becomes a

225 critical step in obtaining meaningful solutions taking into account the system performance and the

226 robustness of the solution to variations in the uncertain data.

227 It is assumed that the probability distribution $\mathbb{P}$ of the N-dimensional stochastic data process is

228 approximately given by many scenarios

229

230 $$D_u = [d_{1,u}, d_{2,u}, \dots d_{N,u}], \ \forall \ i = 1, 2, \dots, N \ and \ u = 1, 2, \dots, S \tag{6}$$

231

232 to which the probabilities p_u are associated and $\sum_{u=1}^{S} p_u = 1$.

233 In order to approximate the probability distribution $\mathbb{P}$ with another $\mathbb{Q}$ distribution, with a smaller

234 number of elements, so that $\mathbb{Q}$ will be as close as possible to $\mathbb{P}$ in terms of probabilistic distance,

235 we use the Kantorovich distance, K, which is the most common probability distance used in

236 stochastic optimization. It is defined between two probability distributions $\mathbb{P}$ and $\mathbb{Q}$ and represent

237 the optimal value to a linear transportation problem [28]. In the case where $\mathbb{P}$ and $\mathbb{Q}$ are finite

238 distributions, the Kantorovich distance is obtained by solving the following problem (see [26])

239 $$K(\mathbb{P}, \mathbb{Q}) = inf \left\{ \sum_{u=1}^{S} \sum_{w=1}^{\tilde{S}} \eta_{uw} c_N(d^u, \tilde{d}^w) : \ \eta_{uw} \geq 0, \ \sum_{u=1}^{S} \eta_{uw} = q_w, \ \sum_{w=1}^{\tilde{S}} \eta_{uw} = p_u \ \forall u, \forall w \right\} \tag{7}$$

240

241 where $c_N(d^u, \tilde{d}^w) := \sum_{\tau=1}^{i} |d_\tau^u - \tilde{d}_\tau^w|, \ i = 1, \dots, N$ with $|.|$ some norm on $\Re^n$, with w=1,..,$\tilde{S}$, and ,$\tilde{S}$

242 is the number of reduced scenarios. The cost function c_N measures the distance between scenarios,

243 p_u and q_w are respectively the probabilities of scenarios in $\mathbb{P}$ and in $\mathbb{Q}$.

244   The previous problem is not easy to solve and in order to overcome to these difficulty heuristic
245   algorithms have been developed, in particular fast backward and fast forward strategies have been
246   implemented [29].   In this paper, we make use of the forward selection algorithm. It defines an
247   iterative process which starts with an empty set. At each iteration, from the set of the non-selected
248   scenarios, the scenario minimizing the Kantorovich distance between the reduced and original sets
249   is selected and inserted in a reduced set . The optimal selection of a single scenario may be repeated
250   recursively until the prescribed number S of elements is reached.
251   Actually, the forward selection algorithm does not guarantee that the reduced set of scenarios is the
252   closest in the Kantorovich distance with respect to the original set and represents the optimal solution
253   of the original problem described by equation (7). However, the empirical results described in the
254   literature [30] indicate that the forward selection algorithm works well in practice.
255
256   *2.3.1. Procedure for reducing scenarios*
257   In this paper a scenarios reduction algorithm based on Kantorovich distance was used. In
258   particular, the fast forward selection algorithm as described in [28] was applied. Starting from an
259   empty set, an iterative process was followed until the required number of selected scenarios was
260   reached. Below is a brief description of this methodology.
261   First, the high number of generated demand scenarios at each node have been assumed to be
262   equiprobable. Thus, at each iteration, using the Euclidean norm ℓ^2, the distances between all the
263   possible pairs of scenarios were calculated for each node. Then, by summing the corresponding
264   distances for all the different nodes, a cost function matrix c_N was derived. This function allows to
265   evaluate the Kantorovich distances matrix between pairs of scenarios taking into account their
266   probability of occurrence.
267   The scenario corresponding to the minimum value of the Kantorovich distance is then selected and
268   the cost matrix is updated by replacing each element with the minimum value between the original
269   element and the one corresponding to the selected scenario. At this point, the procedure is repeated,
270   and a new scenario is added to the reduced scenario set until the number of requested scenarios is
271   reached. In the end, an optimal redistribution of probabilities was carried out by adding the
272   probabilities of non-selected scenarios to the probabilities of those in the reduced set, that is the
273   probability of each non-selected scenario was summed to the probability of the closest selected
274   scenario according to the cost function.
275   Therefore, according to equation, the new probability of a preserved scenario is equal to the sum of
276   its former probability and all the probabilities of the deleted scenarios that are closest to it according
277   to c_N. Obviously, the deleted scenarios have probability zero.

278

279

280

281

282

283

284
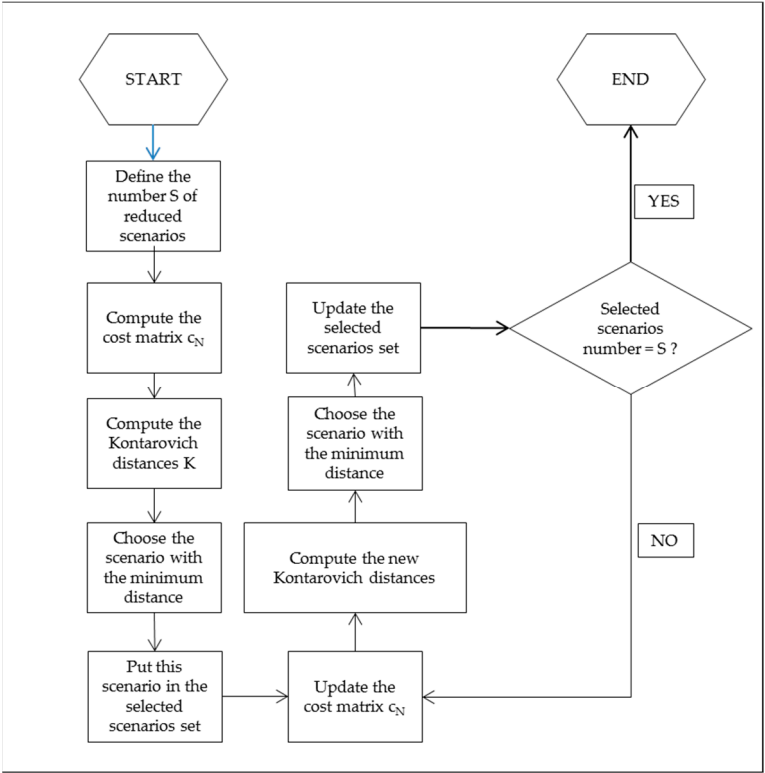
Figure 1. Flow-chart of the procedure for reducing scenarios.

3. Application example

3.1 Theoretical WDN: the Apulian network

For a better understanding of the developed approach an application using the Apulian network layout [15] is elaborated as an example. The Apulian network layout comprises 1 reservoir, 23 demand nodes and 34 pipes. (Figure 2). The network is used only with the objective of defining a scene and providing visual help to the application example.
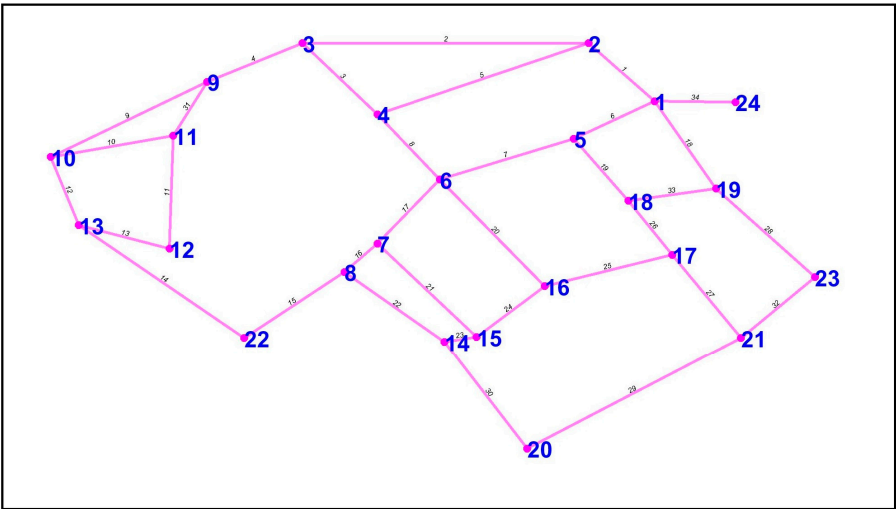
Figure 2. Apulian Network layout.

The geometrical features of the WDN are summarized in Table.1.

Table 1. WDN geometrical features, number of users and water demand at nodes

PIPES											
Pipe number	Start Node	End Node	Length (m)	C Hazen Williams	D (m)	Pipe number	Start Node	End Node	Length (m)	C Hazen Williams	D (m)
1	1	2	348.5	100	0.327	18	1	19	583.9	100	0.164
2	2	3	955.7	100	0.29	19	5	18	452	100	0.229
3	3	4	483	100	0.1	20	6	16	794.7	100	0.1
4	3	9	400.7	100	0.29	21	7	15	717.7	100	0.1
5	2	4	791.9	100	0.1	22	8	14	655.6	100	0.258
6	1	5	404.4	100	0.368	23	15	14	165.5	100	0.1
7	5	6	390.6	100	0.327	24	16	15	252.1	100	0.1
8	6	4	482.3	100	0.1	25	17	16	331.5	100	0.1
9	9	10	934.4	100	0.1	26	18	17	500	100	0.204
10	11	10	431.3	100	0.184	27	17	21	579.9	100	0.164
11	11	12	513.1	100	0.1	28	19	23	842.8	100	0.1
12	10	13	428.4	100	0.184	29	21	20	792.6	100	0.1
13	12	13	419	100	0.1	30	20	14	846.3	100	0.184
14	22	13	1023.1	100	0.1	31	9	11	164	100	0.258
15	8	22	455.1	100	0.164	32	23	21	427.9	100	0.1
16	7	8	182.6	100	0.29	33	19	18	379.2	100	0.1
17	6	7	221.3	100	0.29	34	24	1	158.2	100	0.368

NODES					
node ID	elevation (m)	users A	*DemandA* (l/s)	users B	*DemandB* (l/s)
1	6.4	932	10.86	155	1.8
2	7	1461	17.03	427	4.98
3	6	1282	14.95	48	0.56
4	8.4	1224	14.28	129	1.5
5	7.4	869	10.13	1270	14.81
6	9	1316	15.35	486	5.67
7	9.1	782	9.11	63	0.73
8	9.5	901	10.51	766	8.93
9	8.4	1045	12.18	63	0.73
10	10.5	1249	14.57	45	0.52
11	9.6	848	9.88	964	11.24
12	11.7	650	7.58	354	4.12
13	12.3	1303	15.2	122	1.42
14	10.6	1162	13.55	185	2.15
15	10.1	791	9.23	81	0.94
16	9.5	960	11.2	55	0.64
17	10.2	984	11.47	968	11.29
18	9.6	928	10.82	55	0.64
19	9.1	1258	14.68	45	0.52
20	13.9	1142	13.32	416	4.85
21	11.1	1255	14.63	567	6.61
22	11.4	1030	12.01	137	1.59
23	10	886	10.33	920	10.73
24(Reservoir)	36.4	24258	282.86	8321	96.97

349   Data from two-years long water demand measurements of 82 single household users in the town
350   of Latina, Italy [1] were considered for the definition of the statistical parameters of a typical
351   residential consumption and the calibration of the scaling laws. Table 2 summarizes the average
352   water demand of each typical user and its relevant statistics at peak hour (7-8 am), considering a five-
353   minute time step in data monitoring. The users were considered all the same type and for this reason
354   the same statistical parameters were employed. The number of users at each node of the WDN is also
355   listed in Table 1. In the following examples, residential water consumers are considered, but the
356   approach can be applied to any type of consumption, provided that the statistical characteristics of
357   the typical user are available through monitoring or numerical simulation.

358   Regarding the correlation between pairs of single households a very low value of the Pearson
359   coefficient, i.e. $E[\varrho_1] = 0.0043$, was considered, in agreement with most of the experimental data from
360   the case study of Latina. Table 1 also shows the number of consumption units for each water demand
361   node. In order to highlight how the number of users per node and their relationships affect the
362   generated demand scenarios two different frameworks were examined to which correspond
363   respectively *DemandA* and *DemandB* column.   The *DemandA* values are the same used by Giustolisi
364   et al. [31] for Apulian network and the users' number is consequent. Differently, *DemandB* values
365   were defined assuming a smaller total number of users and a greater variability of the number of
366   users per node.

367   **Table 2.** Relevant statistics and parameters derived from the consumption data of Latina.

Statistical parameter	value
$\mu_{average}$ (L/min)	0.365
$\mu_{peak\ hour}$ (L/min)	0.700
$\sigma_{peak\ hour}$ (L/min)	0.870
scaling law exponent α	1.230

368   *3.2 Generation of demand scenarios*

369   The methodology presented in this work was applied to generate scenarios of contemporary
370   water demands in the supply nodes of the Apulian distribution network. Two different hypotheses
371   have been made on the number of users at the demand nodes. In the first one, *DemandA*, the number
372   of users was determined in order to obtain the demand values assumed by Giustolisi et al. [31], which
373   are appropriate for the subsequent hydraulic simulation of the network. Instead, the second
374   hypothesis, *DemandB*, considers a lower number of total users, but, above all, considerably
375   differentiates the number of users at each node. This was done with the aim of highlighting how the
376   correlation matrix obtained from the scaling laws is influenced by the number of users in the nodes
377   and by their mutual relations, and the proposed methodology can manage complex scenarios. The
378   statistical parameters describing the unitary user' demand are obtained from the experimental data
379   of a group of users in the case study of Latina [1]. Data are referred to peak hour and their sampling
380   interval is equal to five minutes.

381   *3.2.1 DemandA*

382   Twelve-thousand demand scenarios were generated using the statistics estimated by the scaling laws
383   and making the hypothesis of Gamma-distribution at each node, Figure 3.   The large number of

scenarios makes it difficult to distinguish them. But above all, nothing can be said about their probability.
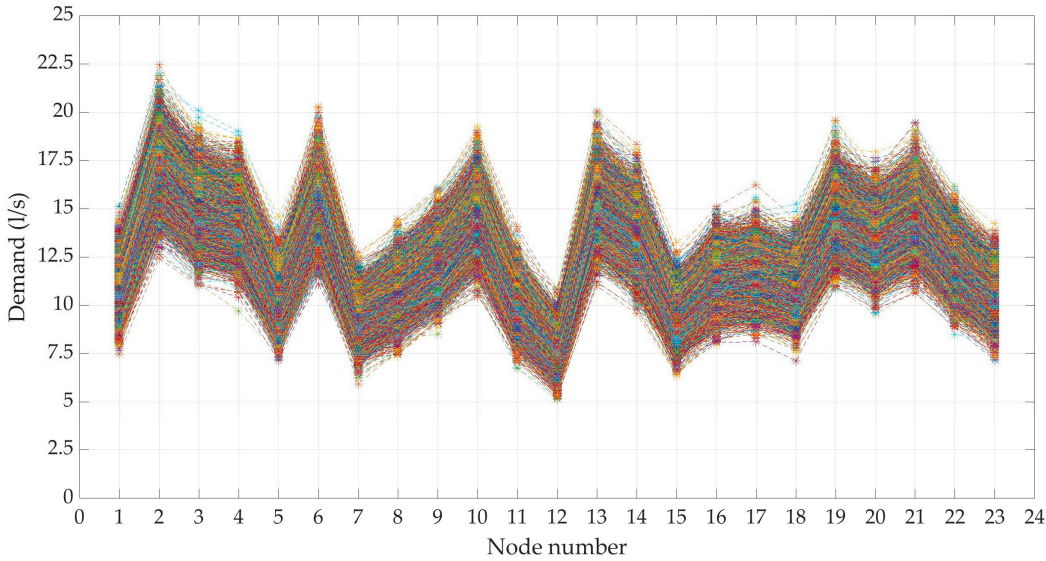
Figure 3. Generated demand scenarios, *DemandA*.

Gamma distribution proves to be the best in fitting the generated data in all nodes of the WDN, Figure 4. The scale and shape parameters of the input distributions $\Gamma(a,b)$, estimated by the scaling laws (INPUT), well agree with the corresponding parameters of the generated scenarios (OUTPUT), Table 3. Regarding the correlation matrix of the generated scenarios, it almost perfectly matches with the input correlation matrix obtained from the scaling laws. Table 4 compares the minimum, average and maximum value of the input and output correlation matrices.
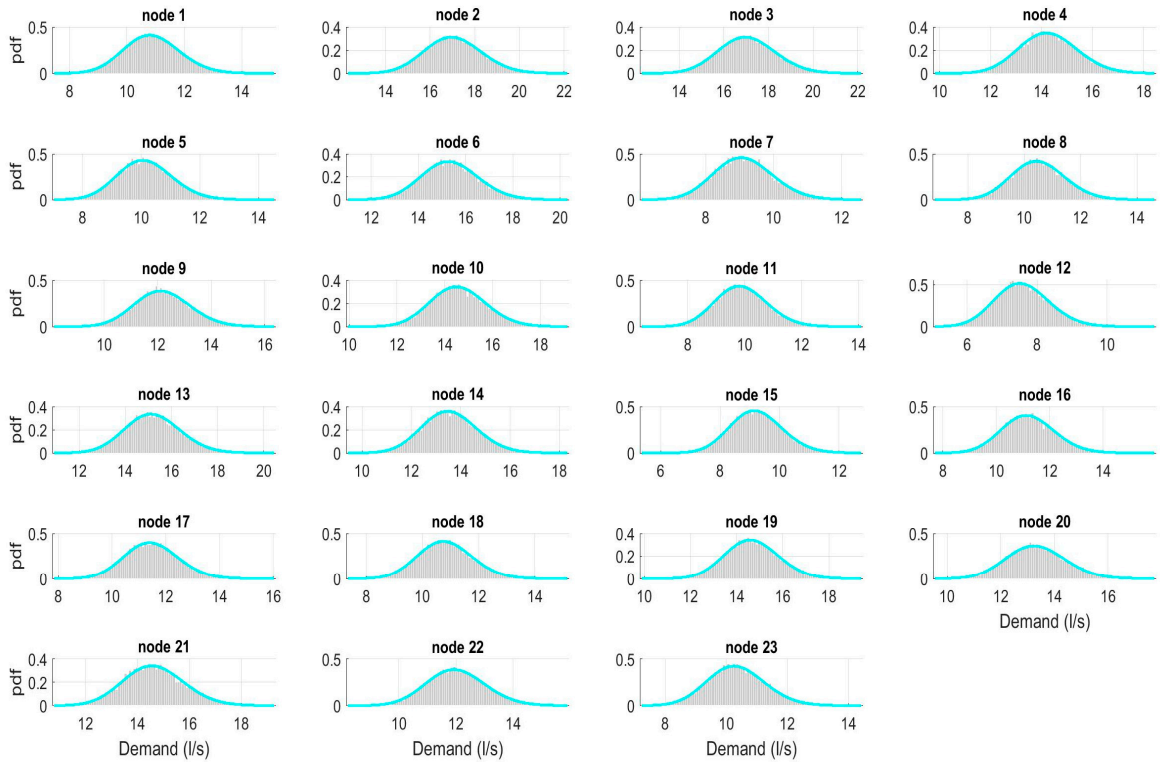
Figure 4. Gamma pdf distributions (cyan line) fitting demand data (grey bar), *DemandA*.

433　　　　　　　　　**Table 3.** Parameters *a* and *b* of Gamma distributions of input and output data, *DemandA*
434

	INPUT		OUTPUT		Node ID	INPUT		OUTPUT	
Node ID	*a*	*b*	*a*	*b*		*a*	*b*	*a*	*b*
1	125.20	0.0868	125.21	0.0868	13	162.06	0.0938	162.07	0.0938
2	176.99	0.0963	177.01	0.0963	14	148.38	0.0914	148.39	0.0914
3	160.04	0.0935	160.06	0.0934	15	110.35	0.0836	110.36	0.0836
4	154.44	0.0925	154.45	0.0925	16	128.09	0.0874	128.10	0.0874
5	118.63	0.0855	118.65	0.0855	17	130.55	0.0879	130.56	0.0879
6	163.30	0.0940	163.32	0.0940	18	124.79	0.0868	124.80	0.0868
7	109.38	0.0834	109.39	0.0834	19	157.73	0.0930	157.75	0.0930
8	121.98	0.0862	122.00	0.0862	20	146.41	0.0910	146.42	0.0910
9	136.73	0.0892	136.75	0.0892	21	157.44	0.0930	157.46	0.0930
10	156.86	0.0929	156.88	0.0929	22	135.22	0.0889	135.24	0.0889
11	116.42	0.0850	116.43	0.0850	23	120.42	0.0858	120.43	0.0858
12	94.86	0.0799	94.87	0.0799	-	-	-	-	-

435

436　　　　　　　**Table 4.** Min, average, max values of input and output cross-correlation matrix, *DemandA*
437

	$E[\varrho_1]=0.0043$	
	ϱ input correlation matrix (scaling laws)	ϱ output correlation matrix (scenarios)
min	0.7542	0.7542
average	0.8147	0.8147
max	0.8568	0.8567

438

439　*3.2.2 DemandB*

440　　　Also, in this case twelve-thousand scenarios were generated using the statistics estimated by the
441　scaling laws, considering Gamma-distributed demands at each node, Figure 5. Differently from
442　*DemandA* case, the value of generated data shows a great excursion between node and node due to
443　the large variability in the number of users.
444　　　Similarly, to the previous case, Gamma distribution proves to be the best in fitting the generated
445　data, Figure 6. Also, the INPUT distributions parameters, estimated by the scaling laws, well agree
446　with the corresponding parameters of the generated scenarios distributions (OUTPUT), Table 5. In
447　this case the correlation matrix of the generated scenarios, it almost perfectly matches with the input
448　correlation matrix obtained from the scaling laws. Table 6 compares the minimum, average and
449　maximum value of the input and output correlation matrices. The input and output correlation
450　matrices are almost identical, but it should be noted that the correlation coefficient has very low
451　values when considering node pairs with low number of users, intermediate values when one of the
452　two has many users, higher values when the number of users is high for both. This responds to the

453   fact that the correlation coefficient depends on the product of the number of users of the two
454   considered nodes, see eq.3.
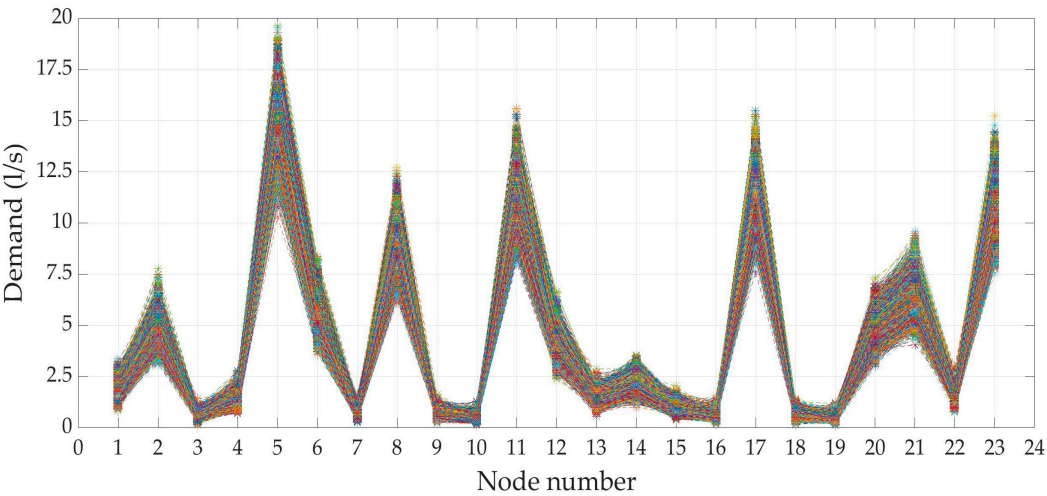
Figure 5. Generated demand scenarios, *DemandB*.
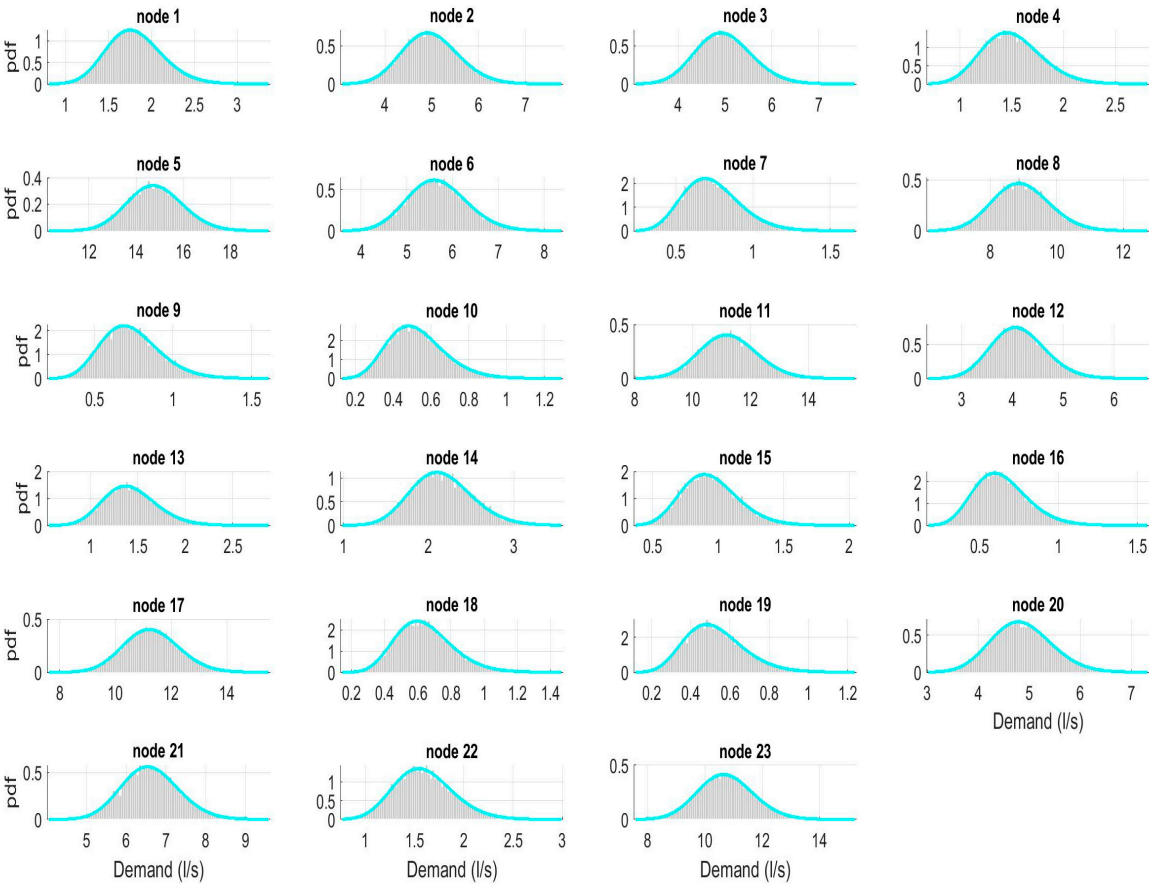
Figure 6. Gamma PDFs (cyan line) fitting demand data (grey bar) at each node, *DemandB*.

504

505 **Table 5.** Gamma PDFs parameters *a* and *b* input and output data at each node, *DemandB*.

506

	INPUT		OUTPUT			INPUT		OUTPUT	
Node ID	*a*	*b*	*a*	*b*	Node ID	*a*	*b*	*a*	*b*
1	31.46	0.0575	31.46	0.0575	13	26.16	0.0544	26.16	0.0544
2	68.64	0.0726	68.65	0.0726	14	36.05	0.0599	36.05	0.0599
3	12.76	0.0439	12.76	0.0439	15	19.08	0.0495	19.09	0.0495
4	27.31	0.0551	27.31	0.0551	16	14.17	0.0453	14.17	0.0453
5	158.89	0.0933	158.90	0.0932	17	128.91	0.0876	128.92	0.0876
6	75.83	0.0748	75.84	0.0748	18	14.17	0.0453	14.17	0.0453
7	15.73	0.0467	15.73	0.0467	19	12.14	0.0433	12.14	0.0432
8	107.65	0.0830	107.66	0.0830	20	67.28	0.0721	67.28	0.0721
9	15.73	0.0467	15.73	0.0467	21	85.39	0.0775	85.40	0.0775
10	12.14	0.0433	12.14	0.0432	22	28.60	0.0559	28.61	0.0559
11	128.50	0.0875	128.51	0.0875	23	123.96	0.0866	123.97	0.0866
12	59.41	0.0695	59.42	0.0695	-	-	-	-	-

507
508
509 **Table 6.** Min, average, max values of input and output cross-correlation matrix, *DemandB*
510

	$E[\varrho_1] = 0.0043$	
	ϱ input correlation matrix (scaling laws)	ϱ output correlation matrix (scenarios)
min	0.1627	0.1627
average	0.4329	0.4330
max	0.8261	0.8261

511

512 *3.3 Reduction of demand scenarios*

513   The final phase of the procedure consists in reducing the number of scenarios by aggregating
514 them in relation to the distance of Kantorovich. For the application of the method the Euclidean norm
515 was considered here. A reduced number of scenarios equal to 20 was chosen. In general, the choice
516 of the number of scenarios should be based on the requirements of robust optimization problems and
517 on the need for the reduced set to continue to describe the whole probability distribution of the
518 demand at each node of WDN.

519 *3.3.1 DemandA*

520   In Figure 7 the reduced demand scenarios are represented. The 'mean scenario', i.e. the one
521 defined by the average nodal values of nodal demand is plotted by the black dotted line. The red

dotted line indicates the most probable scenario, as reported in Figure 7. Mean and most probable scenarios, actually, do not coincide but are very close.
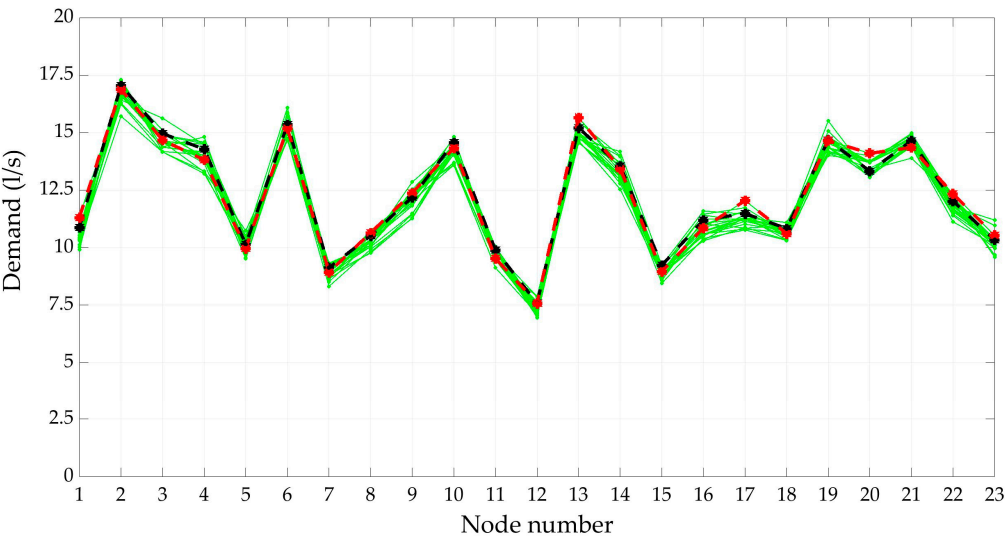
Figure 7. Reduced demand scenarios, *DemandA*.

- reduced nodal scenarios -> green line
- average nodal values -> black dotted line
- scenario with max probability -> red dotted line
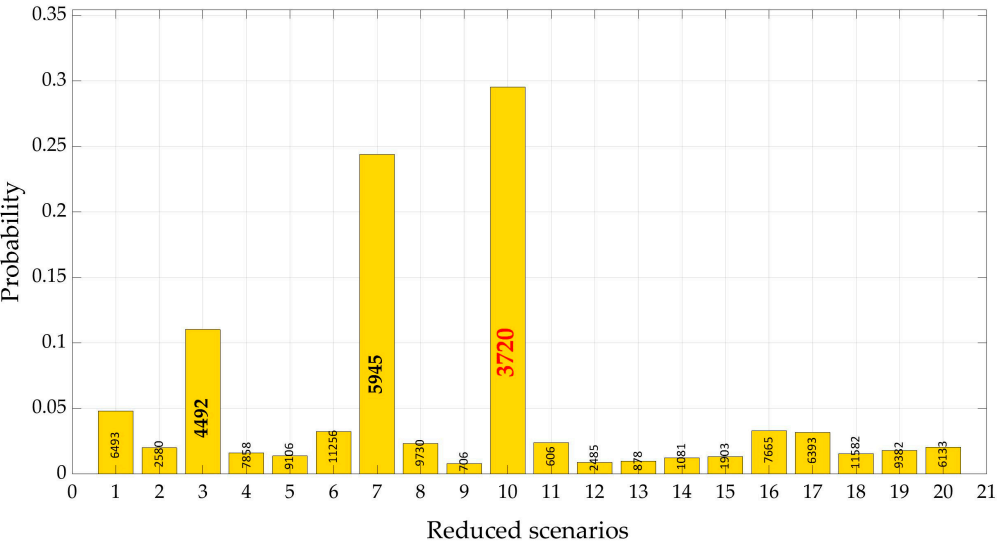
Figure 8. Reduced demand scenarios' probabilities, *DemandA*. Inside the bar is the number of the rank order of the generated scenario.

The most probable scenario in the set of generated ones, number 3720, shows a relevant weight, approximately equal to 0.3. Also, scenarios 5945 and 4492 exhibit significant weights, respectively 0.25 and 0.11. All the others have weight lower than 0.05, Figure 8. Nodal demands of the three most remarkable scenarios are reported in Table 7 and compared with the average value of demand in each node, 'mean scenario'. We can notice that all the selected scenarios are close to the mean one.

574
575
576   **Table 7.** Nodal demand (*l/s*) of most probable scenarios and mean scenario, *DemandA*.

Node ID	Scenario 3720	Scenario 5945	Scenario 4492	Mean scenario	Node ID	Scenario 3720	Scenario 5945	Scenario 4492	Mean scenario
1	10.87	11.31	10.29	10.73	13	15.20	15.66	14.98	14.94
2	17.04	16.88	15.71	17.27	14	13.56	13.43	13.09	13.53
3	14.96	14.67	14.15	14.66	15	9.23	8.96	8.44	8.90
4	14.28	13.82	13.32	14.35	16	11.20	10.85	10.40	11.01
5	10.14	9.97	10.73	10.40	17	11.48	12.04	10.77	11.59
6	15.35	15.19	14.71	16.08	18	10.83	10.61	10.31	10.64
7	9.12	8.94	8.31	8.71	19	14.68	14.65	14.11	14.71
8	10.51	10.64	9.92	10.24	20	13.32	14.10	13.45	13.40
9	12.19	12.38	11.26	12.85	21	14.64	14.35	14.39	13.89
10	14.57	14.33	14.19	14.41	22	12.02	12.34	11.12	12.19
11	9.89	9.53	9.53	9.96	23	10.34	10.53	10.12	10.37
12	7.58	7.58	7.33	7.25	-	-	-	-	-

577

578   *3.3.2 DemandB*

579   In Figure 9 the reduced demand scenarios are represented. In this case it is more difficult to
580 distinguish one scenario from the other because of the big difference of average demand in the nodes
581 Also, here there is no coincidence between the mean and the most probable scenario, number 6232 in
582 the set of generated scenarios, but they are very close, Figure 10. In this case only another scenario,
583 number 9392 has a weight greater than 0.1, but four scenarios exceed the probability threshold of
584 0.05. In Table 8 nodal demands of the three most remarkable scenarios are reported and compared
585 with the average value of demand in each node, 'mean scenario'.

602   **Figure 9.** Reduced demand scenarios, *DemandB*.
603   -   reduced nodal scenarios -> green line
604   -   average nodal values -> black dotted line
605   -   scenario with max probability -> red dotted line
606
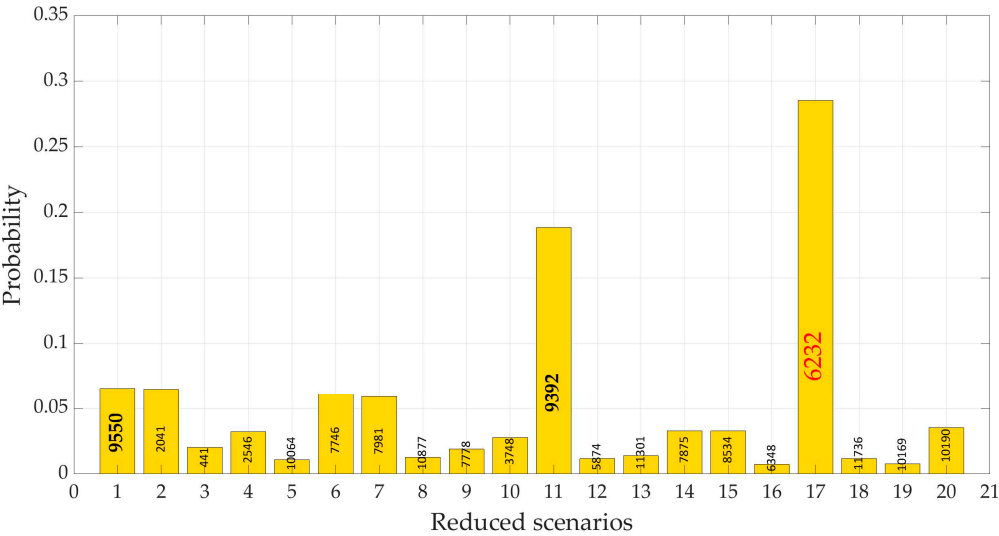
Figure 10. Reduced demand scenarios' probabilities, *DemandB*. Inside the bar is the number of the rank order of the generated scenario

Table 8. Nodal demand (*l/s*) of most probable scenarios and mean scenario, *DemandB*.

Node ID	Scenario 6232	Scenario 9392	Scenario 9550	Mean scenario	Node ID	Scenario 6232	Scenario 9392	Scenario 9550	Mean scenario
1	1.81	1.81	1.75	1.80	13	1.42	1.52	1.27	1.39
2	4.98	5.01	4.80	4.71	14	2.16	2.41	1.58	2.13
3	0.56	0.46	0.64	0.48	15	0.94	1.19	0.99	1.06
4	1.50	1.52	1.39	1.55	16	0.64	0.66	0.65	0.58
5	14.82	15.03	14.55	14.92	17	11.29	11.34	11.24	11.16
6	5.67	5.38	5.38	5.62	18	0.64	0.60	0.83	0.64
7	0.73	0.56	0.59	0.68	19	0.52	0.49	0.45	0.55
8	8.94	8.82	8.45	8.82	20	4.85	5.04	4.86	4.91
9	0.73	0.65	0.76	0.69	21	6.61	7.05	6.65	6.59
10	0.52	0.50	0.34	0.46	22	1.60	1.72	1.53	1.10
11	11.25	11.45	11.04	11.24	23	10.73	11.30	11.12	10.64
12	4.13	4.14	4.04	4.05	-	-	-	-	-

3.4 Hydraulic simulation with scenarios from demand A

Water demand is the forcing parameter of WDN and its natural variability is reflected on the variability of the quantities describing the hydraulic behavior of the whole system. The following question arises: how uncertainty of water demand determines uncertainty of nodal pressure-heads and pipe flow-rate in a WDN? Specifically: the most probable water demand scenario coincides with the most probable pressure-head scenario? At this aim a set of pressure scenarios was derived from the set of generated demand scenarios in the *DemandA* case using a demand-driven hydraulic model based on the Global-Gradient algorithm proposed by Todini and Pilati [32]. Twelve-thousand pressure scenario have been obtained for Apulian network, Figure 11. This example does not take

into account service constraints on each node. Actually, not all pressure scenarios deriving from demand scenarios can be significant in dealing with design and optimization problems. Using pressure-driven hydraulic models can allow automatic filtering of unrealistic scenarios. The reduction algorithm was also carried on and twenty reduced scenarios with the corresponding probabilities were derived, Figures 12 and 13.
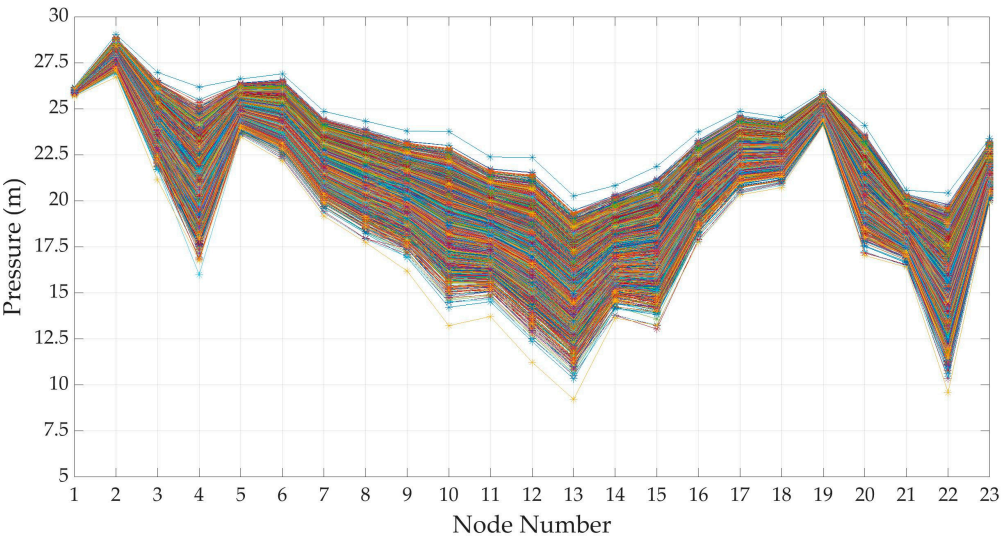
Figure 11. Pressure headlines from the generated demand scenarios (*DemandA*).
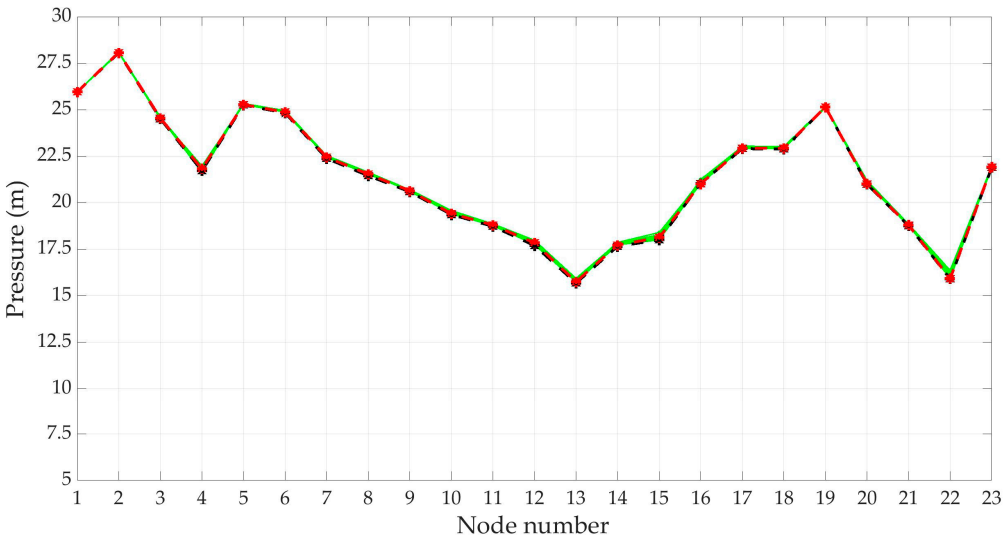
Figure 12. Reduced pressure headlines scenarios, demand A.
- total reduced scenarios -> green line
- average nodal values -> black dotted line
- scenario with max probability -> red dotted line
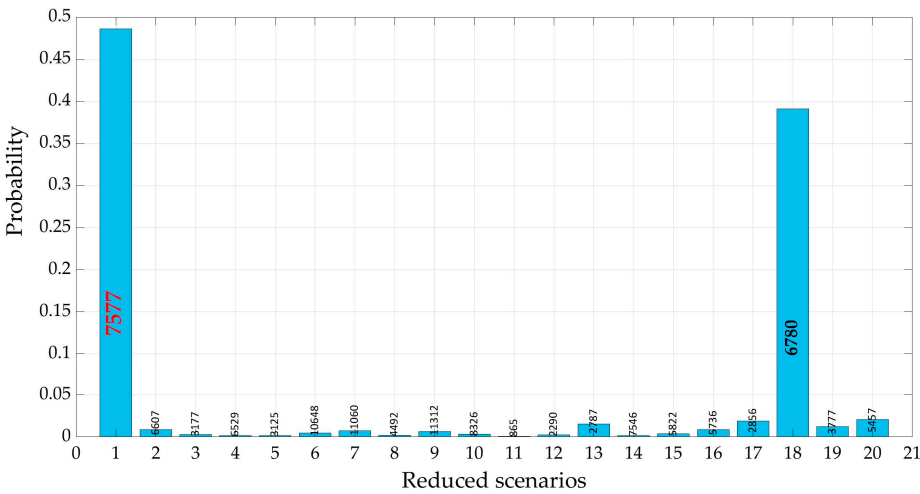
Figure 13. Reduced pressure scenarios' probabilities, demand A. Inside the bar is the number of the rank order of the generated demand scenario.

The results show how the twelve-thousand scenarios generated, unlike the corresponding ones of the water demand, have the same 'shape', which depends on the geometrical and hydraulic characteristics of the WDN. From the results it is also possible to identify the nodes in which the pressure value shows greater variability: this can provide an important help for positioning pressure gauges. Only two scenarios, number 7577, with weight approximately equal to 0.48, and number 6780, with weight about 0.39, remark a significant probability. All the others have probability lower than 0.025. In Table 8 pressure height in each node is reported for the two most probable pressure scenarios, for the scenario of mean pressure, for that of most probable demand (3720) and for the 'mean scenario'. There are no significant differences between the highlighted scenarios.

Table 9. Nodal pressure (*m*) for different scenarios, *DemandA*.

Node ID	Scenario 7577	Scenario 6780	Scenario mean pressure	Scenario 3720	Scenario mean demand	Node ID	Scenario 7577	Scenario 6780	Scenario mean pressure	Scenario 3720	Scenario mean demand
1	25.965	25.962	25.963	25.966	25.965	13	15.754	15.904	15.684	15.839	15.723
2	28.077	28.052	28.061	28.085	28.070	14	17.727	17.850	17.661	17.697	17.692
3	24.548	24.543	24.512	24.592	24.534	15	18.135	18.278	18.001	18.107	18.041
4	21.829	21.852	21.714	21.963	21.755	16	21.023	21.222	21.042	21.138	21.068
5	25.280	25.314	25.254	25.289	25.267	17	22.946	23.040	22.908	22.930	22.927
6	24.875	24.929	24.829	24.888	24.849	18	22.933	22.966	22.895	22.906	22.911
7	22.453	22.522	22.383	22.449	22.407	19	25.149	25.185	25.147	25.139	25.155
8	21.545	21.618	21.454	21.522	21.483	20	21.019	21.193	20.997	21.021	21.026
9	20.647	20.691	20.592	20.694	20.620	21	18.794	18.859	18.768	18.731	18.785
10	19.421	19.553	19.366	19.501	19.404	22	15.926	16.363	15.908	16.111	15.959
11	18.808	18.864	18.731	18.855	18.764	23	21.919	21.976	21.894	21.862	21.909
12	17.846	17.902	17.700	17.901	17.747	-	-	-	-	-	-

723

724

5. Conclusions

This paper proposes a complete procedure to generate and reduce the number of water consumption scenarios which is able to represent the uncertainty due to the variability of demand in a WDN. Moreover, with this procedure, an objective measure of their probability of occurrence is related to each of the reduced scenarios.

For this purpose, the scaling laws, proposed by Magini et al. [1], prove all their potential. In fact, they allow to define the main statistical features of consumption for different number of aggregated consumers, starting from the statistical characteristics of a typical user. Here, residential water consumers like the ones monitored in the case study of Latina [1] have been considered. However, it is possible to examine different types of water uses, different social categories and geographical locations of the same use if water consumers are adequately typified.

One of the main objectives of the proposed procedure is to correctly reproduce the statistical characteristics of water consumption. As derived from the experimental data the nodal water demand is supposed to follow a Gamma probability distribution with parameters dependent on the number of users. In particular, to respect cross-correlation between water demand in each node of the WDN by considering marginal non-normal probability distributions, a generation algorithm based on the approach of Iman-Conover is implemented [8]. The generated water demand scenarios show to respect the hypothesized probability distributions in each node and to reproduce the cross-correlations almost perfectly.

Regarding the reduced scenarios, the possibility of associating a probability/weight to each of them is verified. Obviously, the weight depends on the choice of the number of scenarios that are generated. The results of the simulations here reported show that the most likely scenarios are close to the scenario represented by the average demand in each node. This result may seem trivial, less trivial is the weight that is associated with it.

The reduction procedure has been applied also to the pressure-head scenarios related to the demand generated ones. Also, in this case the most probable scenario is close to that defined by the average pressures in each node of the WDN, but it does not coincide with that produced by the most probable demand scenario.

The proposed procedure is certainly of considerable importance not only in the problems of robust optimization, but more in general in the modeling of the WDNs, both for their design and for their control. This procedure could be further developed with the objective to derive, in addition to the most probable scenarios, that as showed in this work tend to be positioned around the 'mean demand' scenario, some extreme scenarios with low probability of occurrence but of interest in specific issues.

Acknowledgments: This research was supported by 'Progetti Ateneo Sapienza, small projects', protocol RP116154FD8D5BF3.

References

1. Magini, R., Pallavicini, I., Guercio, R., *Spatial and Temporal Scaling Properties of Water Demand*, Journal of Water Resources Planning and Management, 2008. **134**(5): 276–284, doi: 10.1061/(ASCE)0733-9496.

2. Bargiela, A., Sterling, M., *Adaptive forecasting of daily water demand*, in Comparative model for electrical load forecasting. D.W.Bunn and E.D.Farmer, eds.,1985. John Wiley and Sons, New York, N.Y.

3. Perelman, L., M. Housh, and A. Ostfeld, *Robust optimization for water distribution systems least cost design.* Water Resources Research, 2013. **49**(10): p. 6795-6809.

768   4.   Xu, C.C. and I.C. Goulter, *Reliability-based optimal design of water distribution networks.* Journal of Water
769   Resources Planning and Management 1999. **125**(6).

770   5.   Steffelbauer, D., Neumayer, M., Günther, M., Fuchs-Hanusch, D.,   *Sensor Placement and Leakage*
771   *Localization considering Demand Uncertainties,* 2014. Procedia Engineering **89**: p. 1160 – 1167.

772   *6.*   Pallavicini, I., Magini, R., Guercio, R., *Assessing the spatial distribution of pressure head in municipal water*
773   *networks,* 2005, CCWI 2005: Proceedings of the Eighth International Conference on Computing and
774   Control for the Water Industry, Exeter, UK

775   7.   Kapelan, Z.S.,Savic, D.A., and Walters, G.A., *Multiobjective design of water distribution systems under*
776   *uncertainty*, Water Resources Research, **41**(11), W11407, doi:10.1029/2004WR003787, 2005.

777   8.   Iman, R. L., and Conover, W. J., *A distribution-free approach to inducing rank correlation among input*
778   *variables*, Commununications in Statistics, 1982, **11**: p. 311-334.

779   9.   Tolson, B. A., H. R. Maier, A. R. Simpson, and B. J. Lence, *Genetic algorithms for reliability-based*
780   *optimisation of water distribution systems*, Journal of Water Resources Planning and Management, 2004.
781   **130**(1), 141(5): p. 63-72.

782   10.   Vertommen, I., R. Magini, and M.d.C. Cunha, *Generating Water Demand Scenarios Using Scaling Laws.*
783   Procedia Engineering, 2014. **70**: p. 1697-1706.

784   11.   Eck B., Fusco F., Taheri N., *Scenario Generation for Network Optimization with Uncertain Demands.*
785   Proceedings of World Environmental and Water Resources Congress, 2015. 17th Water Distribution
786   Systems Analysis Symposium(**3**): p. 844 - 852.

787   12.   Ridolfi, E., Vertommen, I., Magini, R., *Joint probabilities of demands on a water distribution network: A non-*
788   *parametric approach.* AIP Conference Proceedings, 2013. **1558**(1): p. 1681-1684.

789   13.   Vertommen, I., Magini R., Cunha M.C., Guercio R., *Water demand uncertainty: the scaling law approach*,
790   in Water supply systems analysis: selected topics, Ostfeld, D.A., Editor, 2012, ISBN: 978-953-51-0889-4,
791   InTech: p. 1-25.

792   14.   Vertommen, I., Magini, R., and Cunha M., *Scaling Water Consumption Statistics.* Journal of Water
793   Resources Planning and Management, 2015. **141**(5): p. 04014072.

794   15.   Giustolisi, O., Kapelan, Z.,Savic,D., *Algorithm for Automatic Detection of Topological Changes in*
795   *Water Distribution Networks.* Journal of Hydraulic Eng., 2008. **134**(4):435-446, doi: 10.1061/(ASCE)0733-
796   9429(2008)134:4(435).

797   16.   Vershik, A., *Kantorovich metric: Initial history and little-known applications*. Journal of Mathematical
798   Sciences, 2006. **133**(4):1410–1417.

799   17.   Blokker, E.J.M., J.H.G. Vreeburg, and J.C.v. Dijk, *Simulating Residential Water Demand with a Stochastic*
800   *End-Use Model.* Journal of Water Resources Planning and Management, 2010. **136**(1): p. 19-26.

801   18.   Creaco, E., Alvisi, S., Farmani, R., Vamvakeridou-Lyroudia, L., Franchini, M., Kapelan, Z.,Savic, D.,
802   *Preserving Duration-intensity Correlation on Synthetically Generated Water Demand Pulses.* Procedia
803   Engineering, 2015, **119**, p. 1463-1472

804   19.   Filion, Y.R., et al., *Cross Correlation Analysis of Residential Demand in the City of Milford, Ohio*, in *Water*
805   *Distribution Systems Analysis Symposium 2006*. 2008, ASCE.

806   20.   Filion, Y., B. Adams, and B. Karney, *Cross Correlation of Demands in Water Distribution Network Design.*
807   Journal of Water Resources Planning and Management, 2007. **133**(2): p. 137-144.

808   21.   Savic, D. *Coping with risk and uncertainty in urban water infrastructure rehabilitation planning.* Acqua e
809   Città-1° Convegno Nazionale di Idraulica Urbana, 2005.

810   22.   Hutton, C. J., Kapelan Z., Vamvakeridou-Lyroudia L, Savic D.,*Dealing with Uncertainty in Water*

811   *Distribution System Models: A Framework for Real-Time Modeling and Data Assimilation*. Journal of Water
812   Resources Planning and Management, 2014, **140** (2): p.169-183.

813   23.   Kossieris, P., Makropoulos, C., *Exploring the Statistical and Distributional Properties of Residential Water*
814   *Demand at Fine Time Scales*, Water, 2018, **10**, 1481, doi:10.3390/w10101481.

815   24.   Ponomareva, K., Roman, D., Date, P., *An algorithm for moment-matching scenario generation with*
816   *application to financial portfolio optimization,* European Journal of Operational Research, 2015, **240** (3):
817   p.678-687.

818   25.   Magini, R., Capannolo, F., Ridolfi, E., and Guercio R., *Demand uncertainty in modelling WDS: scaling laws*
819   *and scenario generation*, WIT Transactions on Ecology and The Environment, 2016, **210**, WIT Press,
820   doi:10.2495/SDP160621.

821   26.   Mitra, S., *Scenario Generation for Stochastic Programming*, 2006, White Paper Series, Optirisk Systems,
822   Domain: Finance Reference Number: OPT 004.

823   27.   Ekstrom, P. A., *A simulation toolbox for sensitivity analysis*, Masters Degree Project, 2005, Faculty of
824   Science and Technology, Uppsala Universitet.

825   28.   Dupacová, J., N. Gröwe-kuska, and W. Römisch, *Scenario reduction in stochastic programming, An*
826   *approach using probability metrics.* Math. Program. , 2002. **Ser. A 95**: p. 493-511.

827   29.   Morales J.M., Pineda, S., Conejo, A.J., , Carrión, M., *Scenario Reduction for Future Market Trading in*
828   *Electricity Markets*. IEEE Trasactions on Power Systems, 2009, **24**(2): p. 878-888.

829   30.   Heitsch, H., Romisch W., M., *Scenario reduction in stochastic programming*, Computational Optimization
830   and Applications, 2003, **24**: p.187-206.

831   31.   Giustolisi, O., D. Laucelli, and A.F. Colombo, *Deterministic versus stochastic design of water distribution*
832   *networks.* Journal of Water Resources Planning and Management, 2009. **135**(2).

833   32.   Todini, E., Pilati, S. , *A gradient algorithm for the analysis of pipe networks.* Computer Applications in Water
834   Supply, Research Studies Press, Letchworth, Hertfordshire, UK, 1988.