Using Heuristic Techniques to Account for Engineering Aspects in Modularity-Based Water Distribution Network Partitioning Algorithm

E. Creaco[1]; M. Cunha[2]; and M. Franchini[3]

Abstract: This paper shows how heuristic techniques can be used to account for engineering aspects in the application of a water distribution network (WDN) partitioning algorithm. In fact, being based on graph-theory concepts, most WDN partitioning algorithms fail to consider explicitly such aspects as the number of boundary pipes and the similarity of district metered areas (DMAs) in terms of number of nodes, total demand, and total pipe length, which are often considered by water utility managers to make their decisions. The algorithm considered is the fast-greedy partitioning algorithm (FGPA), based on the original formulation of modularity as an indicator of the strength of WDN partitioning. This algorithm operates by merging the elementary parts of the WDN in sequential steps until the desired number of district metered areas is reached. Two heuristic optimization techniques were combined with FGPA to propose different merging combinations: the former reproduces some specific features of the simulated annealing algorithm while the latter is based on the multiobjective genetic algorithm. Applications were carried out on a real WDN considering the actual system of isolation valves. The partitioning solutions obtained by the traditional FGPA without heuristics and by a literature algorithm based on spectral clustering were taken as benchmark. The results proved that the former heuristic can help in obtaining numerous WDN partitioning solutions with high modularity. The performance of these solutions can be evaluated in terms of practical engineering aspects to help WDN managers make an informed choice about the ultimate solution. If the trade-off between engineering criteria needs to be thoroughly analyzed in the context of WDN partitioning, the latter heuristic, in which FGPA creates DMAs through information encoded in proper weights, can be effectively used. Compared to the benchmark solutions, the FGPA with the latter heuristic can yield solutions with fewer boundary pipes and better demand uniformity over the DMAs. **DOI:** 10.1061/(ASCE)WR.1943-5452.0001129.

Author keywords: Water distribution network; Graph theory; Modularity; Partitioning; District metered area (DMA); Heuristics; Simulated annealing; Genetic algorithm.

Introduction

The partitioning of a water distribution network (WDN) into district metered areas (DMAs) has become a very common practice. In fact, it is very beneficial, in that it facilitates demand management, leakage detection~~,~~ and abatement through service pressure control, model calibration, and so forth (Walski et al. 2003). The separation of each DMA from the rest of the WDN is carried out following the definition of boundaries. At each boundary pipe, the DMA can be physically or virtually separated from the remaining WDN, by closing an isolation valve or installing a flow meter, respectively. The final goal of WDN partitioning is the possibility of monitoring and controlling the exchange of flow between WDN DMAs, which is null in the case of physical separation. Examples of WDN partitioning into DMAs are available starting from the early 2000s (Farley 2001; Morrison 2004; Giugni et al. 2008).

Numerous algorithms have been proposed for WDN partitioning. Some of the algorithms were developed based on graph and spectral theories (e.g., Deuerlein 2008; Perelman and Ostfeld 2011; Zheng et al. 2013; Candelieri et al. 2014; Di Nardo et al. 2016; Galdiero et al. 2016; Hajebi et al. 2016; Herrera et al. 2016; Di Nardo et al. 2017; Zhang 2017; Liu and Han 2018). Others combine graph theory–based techniques and engineering principles, such as the algorithms proposed by Alvisi and Franchini (2013) and Ferrari et al. (2014). A further group of algorithms uses the concept of modularity (Diao et al. 2013; Giustolisi and Ridolfi 2014a, b; Perelman et al. 2015; Campbell et al. 2016; Ciaponi et al. 2016; Laucelli et al. 2017).

Modularity was first formulated for unspecific unweighted and weighted networks in the studies of Newman (2004a, b). It is a topological index that describes the possibility of identifying communities in a network. If the focus is just modularity, then the higher the modularity the better the identification of communities. The original formulation of modularity was used in some studies (e.g., Diao et al. 2013; Ciaponi et al. 2016) for WDN partitioning into DMAs. Specifically, Diao et al. (2013) and Ciaponi et al. (2016) made use of the fast-greedy partitioning algorithm (FGPA), which was based on modularity and developed through the graph

[1]Associate Professor, Dipartimento di Ingegneria Civile e Architettura, Univ. of Pavia, Via Ferrata 3, Pavia 27100, Italy; Honorary Senior Research Fellow, College of Engineering, Physical and Mathematical Sciences, Univ. of Exeter, Exeter EX4, UK; Adjunct Senior Lecturer in the School of Civil, Environmental and Mining Engineering, Univ. of Adelaide, Adelaide 5005, Australia (corresponding author). ORCID: https://orcid.org/0000-0003-4422-2417. Email: creaco@unipv.it

[2]Full Professor, ~~INESC Coimbra – Institute for Systems Engineering and Computers at Coimbra~~, Dept. of Civil Engineering, Univ. of Coimbra, Coimbra, Portugal. Email: mccunha@dec.uc.pt

[3]Full Professor, Dipartimento di Ingegneria, Univ. of Ferrara, Via Saragat 1, Ferrara 44122, Italy. Email: frnmrc@unife.it

Note. This manuscript was submitted on November 18, 2018; approved on April 15, 2019No Epub Date. Discussion period open until 0, 0; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Water Resources Planning and Management***, © ASCE, ISSN 0733-9496.

theory by Clauset et al. (2004). Starting from a configuration in which each node is a DMA of its own, this algorithm operates by aggregating nodes sequentially, while maximizing the increment of modularity at each step, until the target number of DMAs has been reached.

However, a limit of this formulation (FGPA applied to the original formulation of modularity) lies in the fact that it does not account directly for engineering aspects related to WDNs, such as the number of boundary pipes and the uniformity of DMAs in terms of demands and ground elevations. Furthermore, it neglects the fact that in WDNs, isolation valves are usually available at pipe ends rather than in the middle of pipes [an assumption considered by Diao et al. (2013) and by Ciaponi et al. (2016)]. The presence of these limits undermines the applicability of FGPA to real case studies. In fact, water utility managers always make their WDN partitioning decisions based on such engineering aspects as those mentioned previously.

Bearing these limits in mind, Giustolisi and Ridolfi (2014a, b) modified the original formulation of modularity to obtain a WDN-oriented modularity index. A further contribution of Giustolisi and Ridolfi (2014a, b) was to present a modularity-based multiobjective approach for WDN partitioning. The modified modularity index by Giustolisi and Ridolfi (2014a, b) is expressed as the sum of two contributions: the former is a decreasing function of the number of boundary pipes separating DMAs while the latter is a growing function of the similarity of DMAs in terms of a preassigned criterion, such as demand or pipe length distribution across DMAs. As Laucelli et al. (2017) showed, the modified modularity can be inserted in a multiobjective context, where it is maximized while the number of boundary pipes between DMAs is minimized, thus yielding a Pareto front of optimal trade-off solutions among which WDN managers can choose the ultimate partitioning solution.

An alternative approach is presented in this paper to account for engineering aspects in the application of WDN partitioning algorithms based on modularity. Unlike the approach proposed by Giustolisi and Ridolfi (2014a, b), based on a modified formulation of modularity, the novel approach presented in this paper is based on the application of heuristic techniques to the FGPA developed by Clauset et al. (2004) starting from the original formulation of modularity (Newman 2004a, b). To this end, two heuristic techniques were used, the former inspired by the simulated annealing optimization and the latter made up of a multiobjective genetic algorithm. By adding some randomness to the DMA merging in FGPA, Heuristic 1 obtains numerous WDN partitioning solutions featuring high modularity values, some of which are even larger than those obtained through the traditional FGPA. Besides proving the suboptimality of the solutions yielded by the traditional FGPA, Heuristic 1 offers the possibility of accounting for engineering aspects in the postprocessing. In fact, the solutions generated by FGPA modified with Heuristic 1 can be evaluated in terms of various engineering aspects (e.g., number of boundary pipes, demand and pipe length uniformity across DMAs, and so forth), thus enabling an informed choice of the ultimate partitioning solution. Subsequently, Heuristic 2 was developed to show that, if DMA merging in FGPA is driven by proper weights encoded in the individual genes of a multiobjective genetic algorithm, the trade-off between various engineering aspects to be simultaneously optimized can be easily considered directly in the optimization phase.

In the following sections, first the methodology, made up of FGPA and of the two heuristic techniques, is described. The applications to a real WDN follow. Finally, the primary findings of the study are summarized in the conclusions.

Methodology

Fast-Greedy Partitioning Algorithm

To express modularity, reference is made hereinafter to a WDN with nn nodes and n_p pipes, including fictitious pipes representative of the n_{valve} present isolation valves (e.g., Creaco et al. 2010; Giustolisi and Savic 2010). First the weight ω of the WDN pipes must be set, in such a way as to have $\Sigma\omega = 1$. In the case of unweighted network, the weight ω of the generic pipe can be set at $1/n_p$ (leading to an identical weight for all pipes). If the pipes are weighted as a function of supplied demands, the weight ω of the generic pipe can be set at *Dem*/*Demt*, where *Dem* and *Demt* are the demand supplied along the pipe and the overall demand of the WDN, respectively. Otherwise, if the pipes are weighted as a function of pipe lengths, ω can be set at Lp/Lt, where Lp and Lt are the pipe length and the overall length of the WDN, respectively. This can be extended to whatever kind of weight. Then, the incidence topological matrix **A**, with size n_pxnn, can be constructed. In the generic row of **A**, associated with the generic network pipe, the generic element can take on the values 0, $-\sqrt{\omega}$ or $\sqrt{\omega}$, whether the node corresponding to the matrix element is not at the ends of the pipe, it is the initial node of the pipe, or the final node of the pipe, respectively. Starting from **A**, the vector **K** (nnx1) and matrix **D** (nnxnn) can be calculated through the following expressions:

$$\mathbf{K} = \text{diagonal}(\mathbf{A}^T \cdot \mathbf{A}) \tag{1}$$

$$\mathbf{K}_{\text{diag}} = \text{diag}(\mathbf{K}) \tag{2}$$

$$\mathbf{D} = |\mathbf{A}^T \cdot \mathbf{A} - \mathbf{K}_{\text{diag}}| \tag{3}$$

where diagonal( ), diag( ), and | | indicate the vector extracted from the diagonal of a square matrix, the diagonal square matrix constructed starting from a vector, and the absolute value, respectively. The **K** and **D** have an important topological meaning. In fact, the element k_i of vector **K** represents the total weight associated with the pipes connected to the ith node. The generic element D_{ij} of **D** represents the pipe weight connecting the ith and jth node. Following the definition of **K** and **D**, the WDN modularity M can be formulated as (Newman 2004a, b)

$$M = \frac{1}{2}\sum_{i=1}^{nn}\sum_{j=1}^{nn}\left(D_{ij} - \frac{k_i k_j}{2}\right)\delta(c_i, c_j) \tag{4}$$

where c_i and c_j are the DMAs to which the ith and jth nodes belong, respectively; and $\delta(c_i, c_j)$ is equal to 1, whether the ith and jth nodes belong to the same DMA (that is $c_i = c_j$). Otherwise, $\delta(c_i, c_j) = 0$.

Modularity M represents the strength of network partitioning. In fact, a high value of M means that the WDN subdivision is even. In other words, the sum of the weights $\Sigma\omega$ is quite uniformly distributed over the DMAs and only a small part of $\Sigma\omega$ is at the boundary pipes, which do not belong to any DMAs. Conversely, low values of M are associated with poor WDN subdivisions. In fact, the two terms inside the round bracket in Eq. (4) have two effects: the presence of the former, i.e., D_{ij}, guarantees that most of the total weight $\Sigma\omega$ is inside WDN DMAs, entailing that there are few boundary pipes. The subtraction of term $k_i k_j/2$ contributes to the uniform distribution of $\Sigma\omega$ over the DMAs. While the original formulation of Newman (2004a, b) gives the same relevance to the two terms inside the bracket, Giustolisi and Ridolfi (2014b) argued that the effects of the two terms can be modulated by introducing a multiplying factor in the second term.

The objective of the FGPA lies in obtaining a WDN partitioning featuring a high value of M. After the target number N_{dis} of DMAs has been set, the algorithm considers a starting partitioning of the WDN into $N_{disstart}$ DMAs. A suitable starting partitioning is made up of the segments, i.e., the smallest WDN pieces that can be disconnected, through closure of present isolation valves, while avoiding service disruptions throughout the whole network or in large portions (Creaco et al. 2010). For segment identification, suitable algorithms can be used, such as those proposed by Jun and Loganathan (2007), Giustolisi and Savic (2010), and Creaco et al. (2010), based on the real positions of the isolation valves in the WDN. Therefore, at the initial step, the number of DMAs in the WDN is equal to $N_{disstart}$. At the second step, two DMAs are merged, and the number of DMAs becomes $N_{disstart} - 1$. The aggregation process is repeated in the following steps until the network merges to N_{dis} DMAs. At the generic step, the choice of the two DMAs to merge is made to obtain the highest ΔM, where ΔM is a variation in M.

An explicative example of the FGPA is shown in Fig. 1 for a simple WDN, in which $n_{\text{valve}} = 6$ isolation valves are present [Fig. 1(a)]. After replacing the valves with fictitious pipes, the layout in Fig. 1(b) is obtained, made up of $nn = 13$ nodes and $n_p = 14$ pipes. The application of the algorithm of Creaco et al. (2010) for segment identification detects four segments in the WDN. The application of FGPA starting from $N_{disstart} = 4$ to obtain $N_{dis} = 2$ DMAs produces the merging of Segment 1 with Segment 3 and of Segment 2 with Segment 4, in two sequential steps [Fig. 1(c)]. As an explicative example, Fig. 1 shows that the merging of the DMAs takes place while keeping some fictitious pipes representative of isolation valves at DMA boundaries. Others, instead, are incorporated into DMAs. However, as Fig. 1 shows, note that the use of the configurations of $N_{disstart}$ segments as the starting condition for the propagation of FGPA guarantees that, in the aggregation of DMAs, there are always valve-fitted pipes at the boundaries, without any artificial tuning of pipe weights in the modularity function.

Because FGPA is modularity M driven, a remark must be made about how the presence of the fictitious pipes representative of isolation valves impacts on M. In the case of the unweighted graph, $\omega = 1/n_p$, for both the fictitious pipes and the other pipes of the WDN. Therefore, the value of M is influenced both by the uniform distribution of the total number of pipes over the DMAs and by the number of boundary pipes that are left outside DMAs. In the case of weighted graph (e.g., $\omega = Dem/Demt$ or $\omega = Lp/Lt$), instead, the fictitious pipes representative of the isolation valves have weight $\omega \approx 0$. In fact, Dem and Lp are close to 0 for these pipes. Therefore, the number of fictitious pipes at the boundaries has reduced impact on M, which is then affected only by the uniform distribution of the sum $\Sigma\omega$ over the DMAs.

In this implementation where FGPA starts propagating from the $N_{disstart}$ segments present in the WDN, the computational complexity of the algorithm (number of logical operations) is $O[n_{\text{valve}} \cdot d \cdot \log(N_{disstart})]$, where d is the depth of the dendrogram describing the community structure of the WDN. This means that the running time grows linearly with n_{valve}, d, and $\log(N_{disstart})$. The logic structure of FGPA can be summarized in the pseudocode in Fig. 2(a).

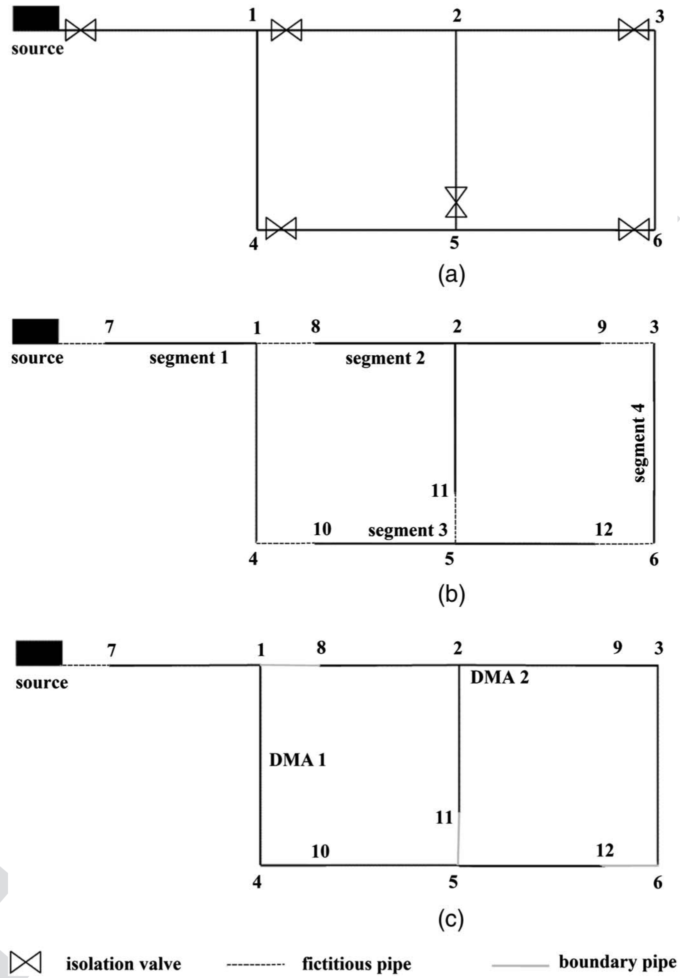

Fig. 1. (a) Network with isolation valves installed; (b) network with fictitious pipes installed instead of isolation valves to enable segment identification; and (c) joining of segments for the construction of two DMAs.

Heuristic 1

Unlike the original FGPA, the possibility of merging two DMAs with a lower ΔM than the highest value mentioned previously is considered in Heuristic 1 [pseudocode in Fig. 2(b)]. This was done to insert some randomness in DMA merging, which obtains various WDN partitioning solutions rather than the single deterministic solution of the traditional FGPA. Furthermore, it is not guaranteed that the DMA merging that produces the highest positive ΔM at the generic step is the most effective choice to obtain the best WDN partitioning into N_{dis} DMAs, i.e., the solution with the highest value of M. At the generic step of the partitioning algorithm, let us assume N_{comb} possible combinations of DMAs for the merging, each of which features its value of ΔM. These values can then be sorted in descending order and then associated with an *index*. A probability function F can then be calculated as

$$F = base + (1 - base)\left(\frac{index}{N_{comb}}\right)^{expo} \tag{5}$$

where *base* and *expo* are two parameters, to be set within the range [0, 1] and $[0, +\infty]$. Basically, F is a monotonic growing function of *index*, ranging from 0 to 1 and yielding the probability of nonexceedance of the generic value of ΔM. The generic combination of merging can be easily sampled from F. In fact, if a random number is generated between 0 and 1, the closest among the values of F larger than the random number, and its associated *index* through Eq. (5), can be easily identified.

(a)

- Let $N_{distart}$= number of segments (initial number of DMAs)
- Set pipe weights ω according to a prefixed criterion
- Evaluate traditional modularity M for the WDN configuration with $N_{distart}$ DMAs
- For k= $N_{distart}$ through N_{dis} (final number of DMAs)
  - Consider the DMA merging that produces the maximum positive ΔM
  - $M \leftarrow M + \Delta M$
- Output: Final configuration of N_{dis} DMAs

(b)

- Let $N_{distart}$= number of segments (initial number of DMAs)
- Set pipe weights ω according to a prefixed criterion
- Evaluate traditional modularity M for the WDN configuration with $N_{distart}$ DMAs
- For k= $N_{distart}$ through N_{dis} (final number of DMAs)
  - Randomly choose the DMA merging with an associated ΔM
  - $M \leftarrow M + \Delta M$
- Output: Final configuration of N_{dis} DMAs

(c)

- Let $N_{distart}$= number of segments (initial number of DMAs)
- Set pipe weights ω according to the first $N_{distart}$ genes of the generic individual in NSGAII
- Modulate M according to the ($N_{distart}$+1)*th* gene of the generic individual in NSGAII
- Evaluate modulated modularity M for the WDN configuration with $N_{distart}$ DMAs
- For k= $N_{distart}$ through N_{dis} (final number of DMAs)
  - Consider the DMA merging that produces the maximum positive ΔM
  - $M \leftarrow M + \Delta M$
- Output: Final configuration of N_{dis} DMAs

Fig. 2. Pseudocodes of (a) FGPA; (b) FGPA with Heuristic 1; and (c) FGPA with Heuristic 2.

An example is provided hereinafter to clarify this concept, considering $N_{comb} = 20$ possible combination of DMAs for the merging of N, producing ΔM ranging from 0.0001 to 0.01. These values are sorted in descending order and associated with *index* [Fig. 3(a)]. Then, function F is calculated as a function of *index* for three pairs of values of *base* and *expo*. Fig. 3(b) shows $F(index)$, from which the sampling of the merging combination is carried out. Fig. 3(b) shows that the pair $base = 0 - expo = 1$ gives an even probability to all the indexes, and therefore to all the values of ΔM. The growth of *base* and the drop of *expo* increase the probability of selection for the lower indexes, and therefore for the higher values of ΔM. Obviously, when a random number sufficiently close to 0 is generated, $index = 1$ is sampled from $F(index)$ in Eq. (5), corresponding to the DMA merging combination with the highest value of ΔM, which is the same merging combination that would be given by the traditional FGPA.

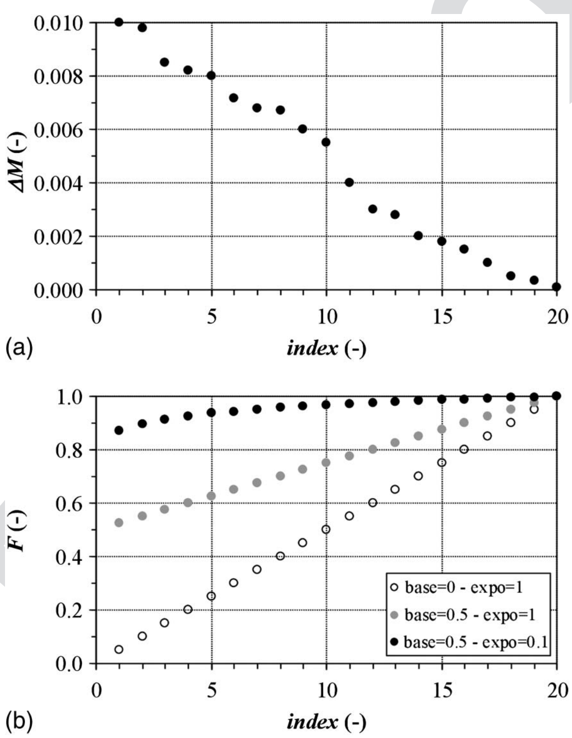

Fig. 3. Preparatory steps for the heuristic merging of DMAs: (a) association of each value of ΔM with an index; and (b) association of *index* with the probability of nonexceedance of ΔM.

Heuristic 1 is embedded in the traditional sequence of FGPA steps from N_{disstart} to N_{dis}. At the generic step, a random number is generated from 0 to 1 to sample the merging combination of DMAs available at that step from $F(index)$ in Eq. (5). Heuristic 1 can be repeated using different sequences of random numbers, producing different values of M for a number of DMAs ranging from $N_{disstart}$ to N_{dis}. Some of these values may result larger than those produced by the traditional FGPA with no heuristic.

As for *base* in Eq. (5), preliminary calculations were done to understand which values to assign to this variable as the steps of the FGPA proceed. Specifically, three options were explored:

1. An even value of *base* within the range [0, 1];
2. Growing values of *base*; and
3. Decreasing values of *base*.

Finally, Option 2 proved successful and the following expression was adopted, which yields a value equal to 0 at the initial step of FGPA and gradually larger values at the following steps:

$$base = \frac{N_{disstart} - N_{dis}}{N_{disstart}} \tag{6}$$

This enables DMA merging combinations with lower values of ΔM than the maximum possible value to be selected especially at the initial steps. When N is far from $N_{disstart}$, that is, at the final steps of FGPA, the merging combinations associated with very high M increment are privileged instead. This brings Heuristic 1 close to the simulated annealing technique (Kirkpatrick and Gelatt 1983) where directions different from that where the objective function experiences the steepest ascent are facilitated at the initial steps, in an attempt to find a global optimum.

The application of Heuristic 1 involves running FGPA for a certain number of times (N_{times}). The computational complexity of FGPA with Heuristic 1 is then N_{times} larger than that of the traditional FGPA.

Heuristic 2

In the framework of WDN partitioning, different objectives from the maximization of M are usually pursued, which include maximization of the uniformity of supplied demands over DMAs, service pressure inside DMAs or of other variables. A further practical objective is the minimization of the number of inter-DMA boundary pipes, at each of which either an isolation valve will be closed or a flow-meter will be installed, thus causing undesirably the loss of reliability or the disbursement of funds, respectively.

In Heuristic 2, the possibility of considering some of the engineering aspects mentioned previously is accounted for by means of the multiobjective genetic algorithm NSGAII (Deb et al. 2002). Specifically, the objective functions considered include the coefficient of variation (ratio of the standard deviation to the mean value) of the total demands delivered to the DMAs, which is an inverse function of the uniformity of supplied demands, and the number of boundary pipes. Both objective functions are simultaneously minimized. These objective functions are in line with those considered by other authors in the scientific literature (e.g., Giustolisi and Ridolfi 2014a, b; Di Nardo et al. 2016; Liu and Han 2018). In fact, the minimization of the number of boundary pipes is considered in almost all the WDN partitioning algorithms, including those based on spectral clustering (Di Nardo et al. 2016; Liu and Han 2018), which aim to solve a relaxed version of the minimum cut problem for the graph. The issue of DMA uniformity, expressed in different forms including demand distribution, was also considered as design criterion by various authors (Giustolisi and Ridolfi 2014a, b; Di Nardo et al. 2016; Liu and Han 2018). The coefficient of variation of demands across DMAs can be related to the second term of the modified index of modularity of Giustolisi and Ridolfi (2014a, b), when pipe weights in the index are expressed as a function of allocated user demands. Furthermore, Liu and Han (2018) presented a design criterion based on a similar formulation to that used in this paper as the second objective function.

The genes of each individual in Heuristic 2 are used to drive the aggregation of DMAs in the traditional FGPA, applied to the initially unweighted graph, to obtain optimal solutions in the expected trade-off. To influence the sequential aggregation of DMAs with the aim to pursue this trade-off, the variation of the weights of the WDN pipes, initially all set at 1, and the modulation of the effects of the two terms present in the original formulation of modularity [Eq. (4)] are encoded in the genes. To obtain the modulation mentioned previously, coefficient α is added in Eq. (4) as multiplying factor of $k_i k_j/2$, yielding the following expression:

$$M = \frac{1}{2}\sum_{i=1}^{nn}\sum_{j=1}^{nn}\left(D_{ij} - \alpha\frac{k_i k_j}{2}\right)\delta(c_i, c_j) \tag{7}$$

To obtain the variation of the weights of the WDN pipes and the modulation of α, each individual is made up of $N_{disstart} + 1$ genes. The first $N_{disstart}$ genes are multiplicative factors $\omega_{s,j}$ of the weights ω of the WDN pipes (initially set at $1/n_p$), to be defined within the range $[0, +\infty]$. If a pipe belongs to the generic jth segment, its weight ω is multiplied by $\omega_{s,j}$. If a pipe is at the boundary between the jth and the kth segment, its weight ω is multiplied by $0.5(\omega_{s,j} + \omega_{s,k})$. Then, the weights ω of the WDN pipes can be rescaled to reobtain $\Sigma\omega = 1$, to be used to assess modularity. The last gene, ranging from 0 to 1, is used for α.

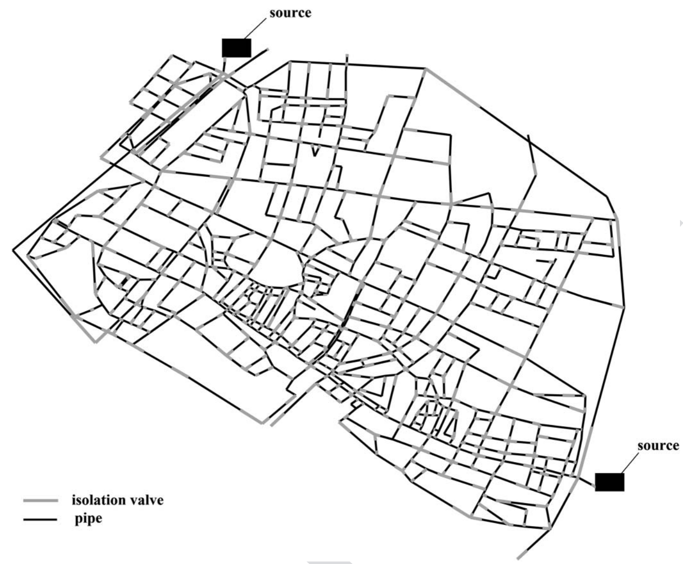

Fig. 4. Layout of the WDN. Division into $N_{distart} = 682$ segments by means of the system of isolation valves. (Adapted from Alvisi et al. 2011.) F4:1 F4:2 F4:3

If a population pop of individuals and a number n_{gen} of generations are considered in FGPA with Heuristic 2, a total number of fitness evaluations equal to $pop \cdot n_{gen}$ is carried out. This means that FGPA is run for $pop \cdot n_{gen}$ times. Therefore, the computational complexity of FGPA with Heuristic 2 is $pop \cdot n_{gen}$ times larger than that of FGPA alone. The sequence of instructions of FGPA with Heuristic 2 is summarized in the pseudocode in Fig. 2(c).

Applications

Case Study

The case study considered in the present work is the whole WDN serving a city in north-central Italy (Alvisi et al. 2011). The network considered, the layout of which is shown in Fig. 4, is made up of 538 nodes, including two reservoirs (Nodes 1 and 2) and 825 pipes. The network has a total length of around 87 km and diameters ranging from 25 to 600 mm. A fixed head equal to 30 m above sea level is assigned to the two source nodes and all the nodes are at the same ground elevation, equal to 0 m above sea level. The total user demand of the network in the peak hour is equal to around 367 L/s, including 20% of leakage. As reported by Alvisi et al. (2011), the network features 969 isolation valves, plus the two valves at the exit of the source nodes. This system of isolation valves subdivides the WDN into $N_{distart} = 682$ segments.

Results

All the algorithms used in this paper were implemented in the Matlab 2017b environment and run using one thread at a time in an Intel Core i7-7700 3.60 GHz CPU. The following subsections first report the results of the traditional FGPA, followed by the results of FGPA with Heuristics 1 and 2.

Traditional FGPA

The application of the traditional FGPA produced configurations of N_{dis} DMAs with N_{dis} ranging from $N_{distart} = 682$ (DMAs coincident with WDN segments) down to $N_{distart} = 1$ (whole WDN in

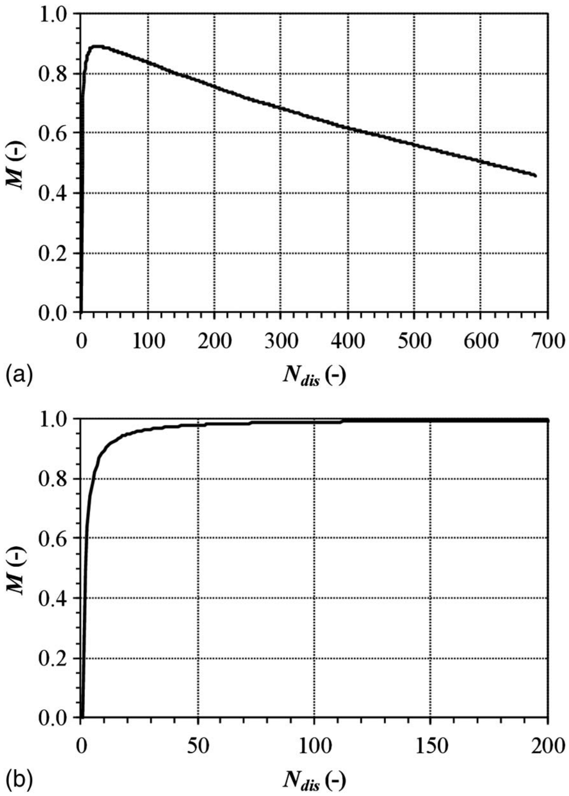

Fig. 5. Application of FGPA with modularity M expressed for (a) unweighted graph; and (b) graph weighted based on pipe demands.

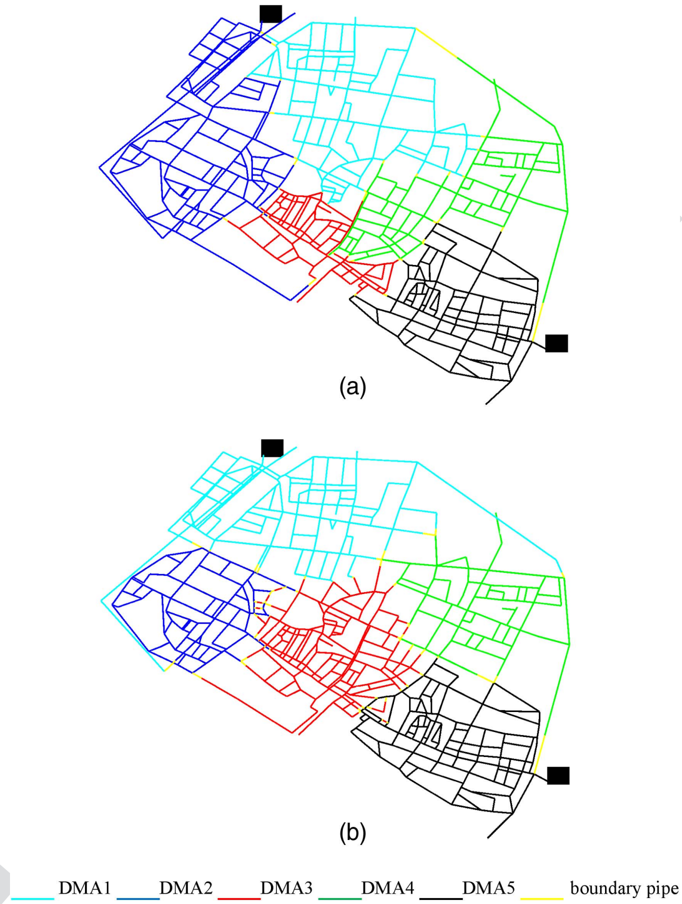

Fig. 6. (a) Application of FGPA with modularity M expressed for unweighted graph; and (b) benchmark solution obtained with the spectral clustering algorithm of Di Nardo et al. (2016).

one district). First, the traditional FGPA was run considering the unweighted graph (pipe weight $\omega = 1/n_p$). Then, it was run considering the weighted graph (pipe weight $\omega = Dem/Demt$). In this case, the demands distributed along the pipes in the peak hour were considered. The results of these runs are reported in the graphs in Figs. 5(a and b) in terms of $M(N_{\text{dis}})$. The graph in Fig. 5(a) reports a value of M equal to about 0.46 in correspondence to $N_{distart}$. When N_{dis} decreases due to the merging of DMAs, M grows up to a maximum value of about 0.89 in correspondence to $N_{dis} = 23$. This means that at $N_{dis} = 23$ it is possible to obtain the most modular WDN partitioning into DMAs, with a uniform distribution of pipes over the DMAs and with a low number of boundary pipes left out of the DMAs. To the left of this value, M falls to 0 for $N_{dis} = 1$. In fact, at $N_{dis} = 1$ all the pipes belong to a single DMA and there is no WDN partitioning. The graph in Fig. 5(b) shows a different pattern $M(N_{dis})$. In fact, the maximum of M lies in correspondence to the highest value of $N_{dis} = N_{distart} = 682$. To the left of this value, M decreases toward 0 at $N_{dis} = 1$. The different behavior in the two graphs in Fig. 5 is because, as mentioned previously, in the case of weighted graph, M is affected only by the distribution of $Demt$ over the DMAs while the number of boundary pipes has no impact on M.

As an example, Fig. 6(a) reports, for $N_{dis} = 5$, the WDN partitioning results of FGPA-unweighted graph. Though being slightly small in light of the WDN total size and demand, a total number $N_{dis} = 5$ of DMAs was chosen in this context because it enables easy visualization of the results of WDN partitioning. In the graph, ~~to make distinction between the DMAs, a different color is used to characterize the pipes of each DMA. Boundary~~ pipes, the end nodes of which belong to two different DMAs, are plotted. The WDN partitioning in Fig. 6(a) features $N_{bp} = 39$ and $M = 0.774$ evaluated in the case of unweighted graph.

As a benchmark, Fig. 6(b) reports the WDN partitioning solution obtained with the spectral clustering algorithm of Di Nardo et al. (2016), applied with the constraint of having boundary pipes at valve-fitted pipes. Though being obtained without considering modularity explicitly, this solution features a quite high value of M, equal to 0.749 evaluated in the case of unweighted graph. In fact, the key ingredients of modularity, namely the balancing between DMAs and the low number of boundary pipes, are also the objectives of spectral clustering partitioning methods. However, the fact that the M value for Fig. 6(b) is smaller than that for Fig. 6(a) is due to the larger number of boundary pipes (57) provided by the algorithm of Di Nardo et al. (2016).

FGPA with Heuristic 1

This subsection aims to prove how the results of FGPA with Heursitic 1 can be evaluated in terms of engineering aspects. The FGPA with Heuristic 1 was run $N_{\text{times}} = 100$ in the unweighted graph. Each time, a pattern $M(N_{dis})$ similar to Fig. 5(a) was obtained. Due to the stochastic nature of this algorithm, the results were different from one run to the other. Then, the pattern $r_M(N_{dis})$ was calculated in each run, where r_M is the ratio of the M value obtained in the run for the generic value of N_{dis} to the corresponding M obtained in the traditional FGPA. The graph in Fig. 7(a) shows that $r_M(N_{dis})$ is always around 1, highlighting that the FGPA with Heuristic 1 is always able to yield WDN partitioning solutions with high modularity. Furthermore, it must be remarked that the

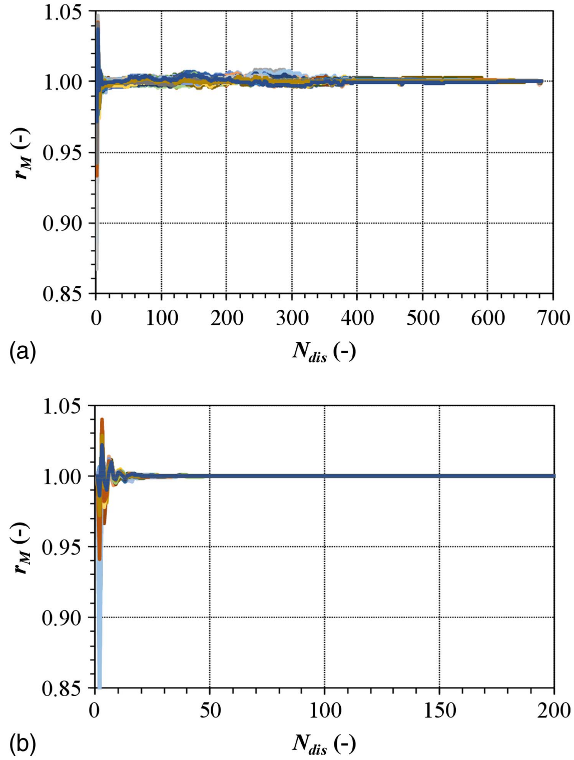

Fig. 7. Application of FGPA with Heuristic 1. Ratio r_M as a function of N_{dis} for (a) unweighted graph; and (b) weighted graph based on pipe demands. Each ~~color~~ indicates a different run.

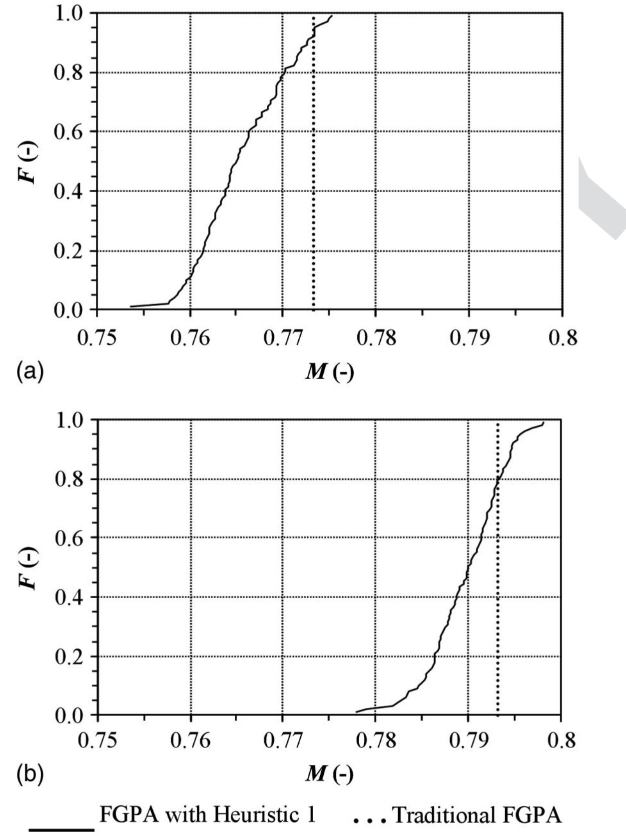

Fig. 8. ~~Weibul~~ frequency F of the modularity function M obtained through FGPA with Heuristic 1 for $N_{dis} = 5$ in the (a) unweighted graph; and (b) weighted graph based on pipe demands.

upper envelope of $r_M(N_{dis})$ is always larger than 1. This proves both the suboptimality of the traditional FGPA and the possibility to explore solutions with higher modularity thanks to adoption of Heuristic 1. The best gain obtainable with the FGPA with Heuristic 1 is for $N_{dis} < 10$, with a maximum value of r_M close to 1.05.

The FGPA with Heuristic 1 was run 100 times also in the weighted graph ($\omega = Dem/Demt$). Similar calculations to those of the weighted graph led to the graph reporting $r_M(N_{dis})$ in Fig. 7(b). This graph only reports the values of r_M for $N_{dis} < 200$ because the others are almost coincident with 1. Similar remarks to the application with the unweighted graph can be made also in this case.

The subsequent results in this subsection are shown for $N_{dis} = 5$, that is considering WDN partitioning into five DMAs. Fig. 8 reports the ~~Weibul~~ frequency of the M values obtained in the 100 runs of FGPA with Heuristic 1 in the unweighted graph [Fig. 8(a)] and in the weighted graph based on pipe demands [Fig. 8(b)]. This figure shows that the M values obtained are always very high, some of which resulting to be larger than that provided by the traditional FGPA. This attests to the high effectiveness of Heuristic 1.

Each configuration of WDN partitioning obtained in the unweighted graph was then re-evaluated as number N_{bp} of boundary pipes between DMAs and as uniformity of DMAs in terms of number of nodes N_{nd} in each DMA, pipe lengths L in each DMA and demands Dd in each DMA. As for the uniformity, the coefficients of variation $C_{v,Nnd}$, $C_{v,L}$, and $C_{v,Dd}$ were calculated. Graphs in Figs. 9(a), 9(b), 9(c), and 9(d) report the relationship between M, and N_{bp}, $C_{v,Nnd}$, $C_{v,L}$ and $C_{v,Dd}$, respectively.

The graphs in Fig. 9 show that though the solutions of FGPA run probabilistically with Heuristic 1 are very close in terms of M, which ranges from about 0.745 to 0.775, they feature very different values of N_{bp}, $C_{v,Nnd}$, $C_{v,L}$, and $C_{v,Dd}$. A negative correlation exists between M and N_{bp} ($\rho = -0.65$) and between M and $C_{v,Nnd}$ ($\rho = -0.84$). This was expected because, in the case of unweighted graph, M depends on the uniform subdivision of the number of pipes (and therefore of nodes) over the DMAs and on the number of boundary pipes. The re-evaluation of the configuration of WDN partitioning obtained in the weighted graph based on pipe demands is reported in the graphs in Fig. 10, for which similar considerations to Fig. 9 can be made. The only significant difference lies in the fact that, in the case of weighted graph, the correlation between M and N_{bp} disappears and that between M and $C_{v,Dd}$ emerges ($\rho = -1.0$). This is because, as mentioned previously, in the case of weighted graph, M is affected only by the distribution of $Demt$ over the DMAs while being totally unaffected by the number of boundary pipes.

Overall, the graphs in Figs. 9 and 10 offer practitioners an insight into the significant engineering performance of the various WDN partitioning solutions, in terms of N_{bp}, $C_{v,Nnd}$, $C_{v,L}$, and $C_{v,Dd}$, from which an informed choice of the final solution can be made. For the assessment of M, the use of the weighted graph instead of the unweighted one is preferable when a strong correlation between M and one specific variable must be searched for [Fig. 10(d)]. However, considering the weighted graph in the assessment of M tends to yield larger numbers N_{bp} of boundary pipes [compare Figs. 10(a) and 9(a)].

FGPA with Heuristic 2

If the trade-off between engineering aspects need to be explored more deeply, Heuristic 2 can be profitably used, as is proven in

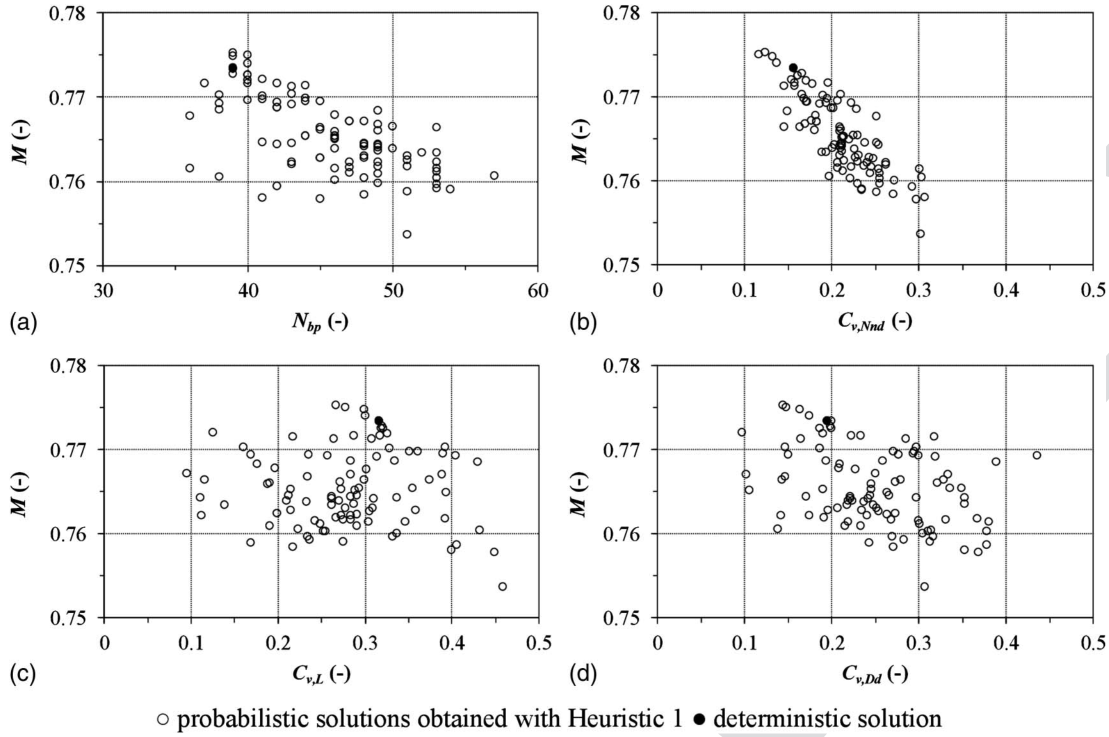

Fig. 9. Relationship between M and (a) N_{bp}; (b) $C_{v,Nnd}$; (c) $C_{v,L}$; and (d) $C_{v,Dd}$ for the solutions obtained through FGPA with Heuristic 1 in the unweighted graph. Comparison shown with the values of the traditional FGPA.

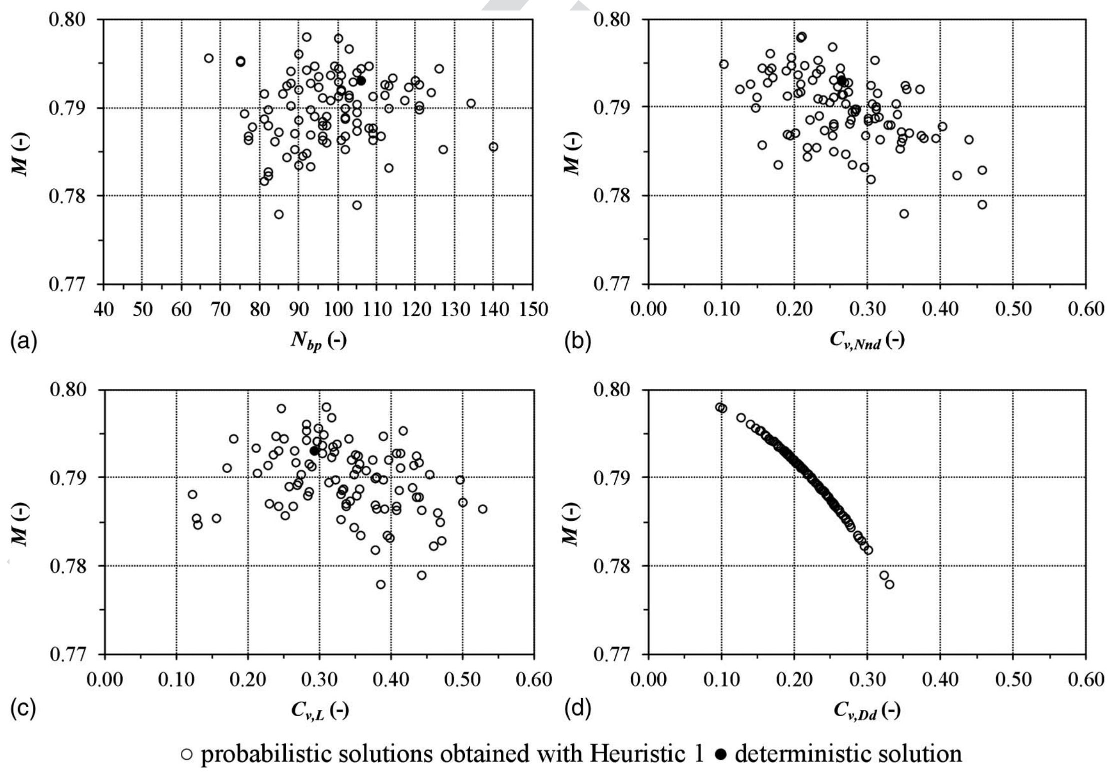

Fig. 10. Relationship between M and (a) N_{bp}; (b) $C_{v,Nnd}$; (c) $C_{v,L}$; and (d) $C_{v,Dd}$ for the solutions obtained through FGPA with Heuristic 1 in the weighted graph based on pipe demands. Comparison shown with the values of the traditional FGPA.

the present subsection. NSGAII was run with a population $pop = 100$ individuals and for $n_{gen} = 100$ generations to search for optimal WDN partitioning into five DMAs, in the trade-off between N_{bp} and $C_{v,Dd}$, to be simultaneously minimized. Coefficient $C_{v,Dd}$, was calculated starting from peak hour demands distributed along pipes. The gene $\omega_{s,1}$ was set to a fixed value (i.e., 1) whereas $\omega_{s,j}$, with $j = 2, \ldots, N_{disstart}$, were allowed to range. This was done to prevent the gene rescaling (performed to guarantee

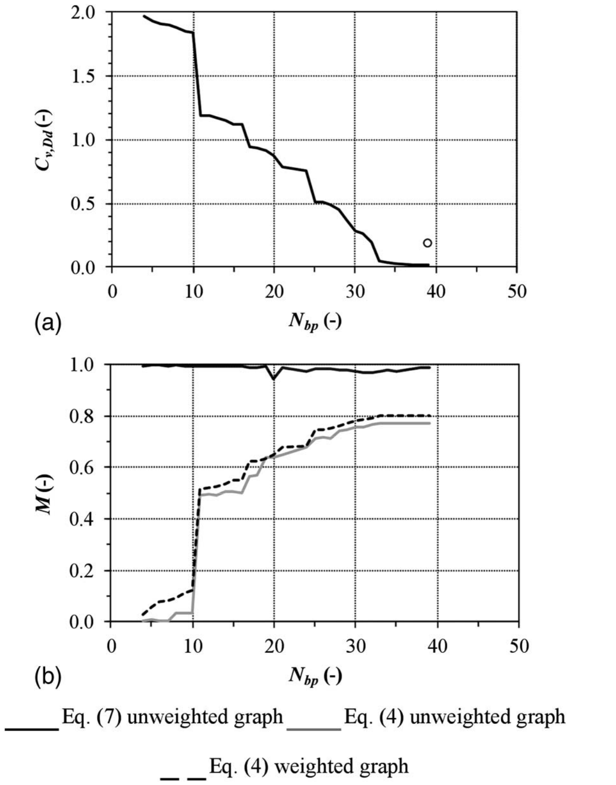

Fig. 11. Pareto front of optimal solutions in the trade-off between N_{bp} and $C_{v,Dd}$; comparison with the deterministic solution obtained with the traditional FGPA~~: (a) deterministic solution FGPA based on unweighted graph (Pareto front obtained with Heuristic 2)~~; and (b) re-evaluations of the optimal solutions in terms of modularity M.

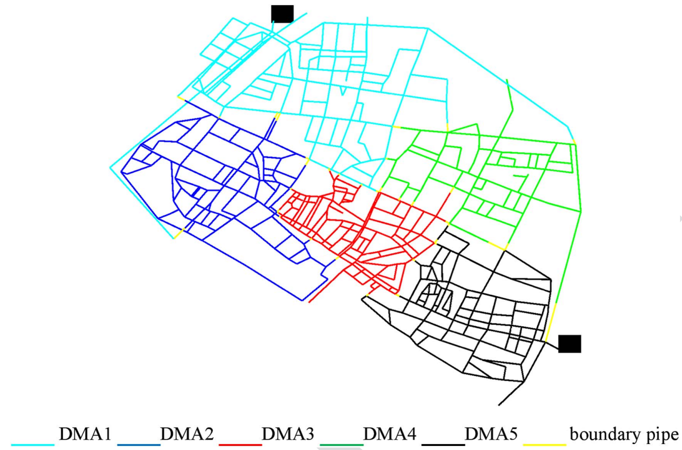

Fig. 12. Graphical representation of the selected solution of WDN partitioning into five DMAs for Heuristic 2.

$\Sigma\omega = 1$) from generating identical values of ω starting from different $N_{disstart}$-tuple $\omega_{s,j}$. The range for $\omega_{s,j}$, with $j = 2, \ldots, N_{disstart}$, was set at [0, 5]. The previous choices for the number of individuals and generations and for the range for $\omega_{s,j}$ were made because preliminary calculations proved these values helped NSGAII in reaching good efficiency in the results. As for the choice of pop and n_{gen}, the selected values enabled a good trade-off between computational overhead and convergence of the Pareto front. Because the range of possible values for $\omega_{s,j}$ is $[0, +\infty]$, various values in between 1 and 10 were tried for the upper boundary. Then, the best Pareto front was obtained when considering an upper boundary equal to 5.

The results of the optimization are reported in Fig. 11(a) as a Pareto front in the $N_{bp} - C_{v,Dd}$ space. Due to the conflicting nature of the two objectives, the front features lower values of $C_{v,Dd}$ as N_{bp} increases. In other words, a more uniform distribution of demand leads to a larger number of boundary pipes to be selected for WDN partitioning into DMAs. A first comparison between the graph in Fig. 11(a) and the graphs in Figs. 9 and 10 highlights that Heuristic 2 enables a wider range to be explored for the variables of interest (e.g., see $C_{v,Dd}$ ranging from about 0.1 to about 0.45 in Fig. 9 and from about 0.017 to about 1.97 in Fig. 11). Furthermore, the trade-off between these variables is investigated thanks to Heuristic 2.

The graph in Fig. 11(a) also shows that the deterministic solution obtained with the traditional FGPA, which features $N_{bp} = 39$ and $C_{v,Dd} = 0.195$, is dominated by numerous solutions of the Pareto front. The graph in Fig. 11(b) presents the re-evaluation of the optimal solutions belonging to the front in terms of modularity M, evaluated according to Eq. (4) considering both unweighted and weighted graph. It also reports the values of M according to Eq. (7), used inside Heuristic 2 to drive DMA merging in FGPA.

The analysis of Fig. 11(b) shows that the values of M according to Eq. (7) are very close to 1, for all the values of N_{bp}. This attests to the maximization potential of FGPA in the context of the modulated modularity in Eq. (7), when suitable values are adopted for the pipe weighting coefficients and for α. The re-evaluated values of M according to Eq. (4), not used in FGPA with Heuristic 2, are always lower. In fact, the values of M evaluated on the unweighted graph grow with N_{bp} increasing from about 0 to around 0.77, close to the value of M obtained with the traditional FGPA for $N_{dis} = 5$. Instead, the values of M evaluated on the weighted graph grow from about 0 to around 0.80, which is even larger than the value of $M = 0.793$ obtained with the traditional FGPA for $N_{dis} = 5$. This represents further evidence of the suboptimality of the solutions obtained with the traditional FGPA.

However, besides featuring high values of M according to Eq. (4), the solutions in the Pareto front in Fig. 11(a) with $N_{pb} \geq 33$ are more interesting from the viewpoint of practical engineering, in that they are associated with WDN partitioning solutions with very even distribution of demand over the DMAs. As an example of these solutions, the optimal WDN partitioning with $N_{pb} = 35$ and $C_{v,Dd} = 0.026$ is shown in Fig. 12. Compared to the solution obtained with the traditional FGPA [Fig. 6(a)], the solution shown in Fig. 12 has different sizes and shapes for the DMAs. Indeed, these different sizes and shapes, which lead to a lower number of boundary pipes (35 versus 39) and to a lower value of $C_{v,Dd}$ (0.026 versus 0.195), is obtained thanks to the tuning of pipe-weights performed by Heuristic 2.

In terms of the considered objective functions, the solution shown in Fig. 12 is also better than the benchmark solution shown in Fig. ~~5(e)~~ [obtained through the spectral clustering algorithm of Di Nardo et al. (2016)]. In fact, this latter solution features $N_{pb} = 57$ and $C_{v,Dd} = 0.16$.

The last calculations were carried out to test, for the WDN partitioning solution shown in Fig. 12, the hydraulic performance and therefore the ultimate feasibility. To this end, the methodology of Creaco et al. (2017) was applied to optimally select the isolation valves to close and the flow meter to install at the boundary pipes. The methodology was applied to maximize the number N_{civ} of closed isolation valves while maximizing the hydraulic performance of the WDN, expressed through the generalized resilience/failure index GRF of Creaco et al. (2016) under peak

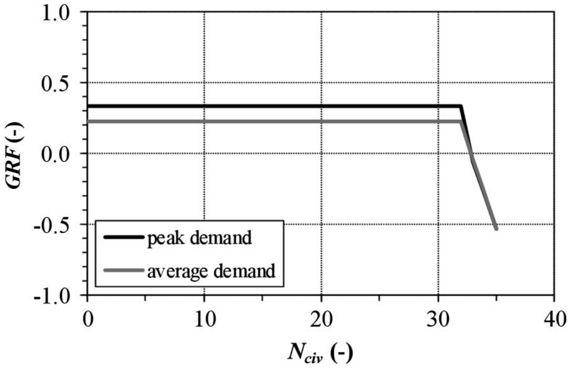

Fig. 13. Pareto front of optimal trade-off solution between number N_{civ} of closed isolation valves and generalized resilience/failure GRF index under peak demand conditions for the configuration of five DMAs shown in Fig. 12. Postprocessing of the solutions under average demand conditions.

demand conditions. This index ranges from −1 to 1 and is associated with the ratio of the power supplied to WDN users to the power leaving the sources. The higher the GRF, the higher the service pressure in the WDN. The Pareto front obtained in this optimization is shown in Fig. 13 (see black line). The value of $GRF = 0.34$ for $N_{civ} = 0$ relates to the virtual WDN partitioning, in which DMAs are not physically disconnected. In fact, for $N_{civ} = 0$ flow meters are installed at all the boundary pipes. The virtual partitioning features the same hydraulic performance as the unpartitioned WDN. Fig. 13 highlights that the hydraulic performance of the WDN stays almost unchanged up to a very large number of N_{civ}. If $N_{civ} = 32$ is chosen, only $N_{bp} - N_{civ} = 3$ flow meters must be installed to monitor the flux exchanges between the DMAs. The reason why the hydraulic performance is not affected by the increase in N_{civ} is the high redundancy of the WDN, which is overly looped and interconnected. However, the sudden decrease in *GRF* to the right of $N_{civ} = 32$ is because at least three boundary pipes must be kept open to guarantee the supply of water to all the [19] five DMAs. In fact, ~~the~~ two DMAs are directly connected to the source nodes. Each of the three other DMAs, instead, needs at least one open boundary pipe to receive water from the rest of the WDN. Therefore, the WDN partitioning solutions for N_{civ} ranging from 33 to 35 are infeasible due to the absence of service to large parts of the WDN. Though the choice between closed isolation valve and installed flow meter at each boundary pipe was made based on peak demand hydraulic analysis, the validity of the results was tested against average demand conditions (about 60% lower than peak demand). To this end, for the solutions of the Pareto front in Fig. 13, the *GRF* was re-evaluated under average demand conditions. As expected, the *GRF* values under average demand conditions (gray line in Fig. 13) are slightly lower than those under peak demand conditions. This is because the power delivered to the users is lower when average demands are considered instead of peak demands. However, similar hydraulic considerations can be made as previously mentioned, pointing out that the hydraulic performance of the WDN does not change up to $N_{civ} = 32$.

Conclusions

In this paper two heuristics were combined with the fast-greedy algorithm for the partitioning of WDNs into DMAs. FGPA operates by assembling small parts of the network in sequence until the desired number of DMAs is reached. In the traditional version of FGPA, each merging is performed to maximize the increment in modularity, a variable representing the strength of WDN partitioning. The former heuristic technique is implemented inside FGPA and enables lower increments in modularity than the maximum to be probabilistically accounted for, above all at the initial steps of FGPA, in the fashion of the simulated annealing optimization. In the latter heuristic technique, FGPA is embedded inside a multiobjective genetic algorithm, which modulates the weights of WDN pipes and the terms inside the modularity function. Applications to a real WDN in northern Italy proved the effectiveness of both the heuristic techniques. In fact, the former enables obtaining numerous high modularity partitioning solutions, some of which are even more modular than that obtained with the traditional FGPA. The solutions can be postprocessed in terms of various variables, such as the number of boundary pipes and the uniformity of DMAs, to enable practitioners to make an informed decision about the final solution. If the trade-off between two or more variables of interest needs to be explored more deeply, the latter heuristic technique can be profitably adopted, though requiring a larger computational burden than the former. In an explicative example of WDN partitioning into five DMAs, the partitioned WDN configuration proved to keep a very similar hydraulic performance to the unpartitioned WDN, even when isolation valves are closed at most boundary pipes to separate the DMAs.

The calculations of this work were carried out on a real WDN with 682 segments to be aggregated into five DMAs, taking overall about 2 days of computation time. Similar calculations could be repeated on a more complex real WDN. To get results in an acceptable time frame in this case, the two following options could be chosen: either to make use of multicore processing or to set as initial configuration for DMA aggregations larger WDN portions than the segments. The former option would enable various WDN partitioning solutions to be analyzed in parallel by the optimizer whereas the latter would reduce the number of steps in FGPA in each evaluation.

Further research will be dedicated to the analysis of the effectiveness of different heuristic techniques. Furthermore, the possibility of implementing heuristics in other kinds of algorithm for WDN partitioning will be explored. While the former heuristic presented in this paper is primarily tailored to FGPA, the latter can be easily generalized to numerous WDN partitioning algorithms. Because most algorithms are developed based on the graph theory, the use of different pipe weights is expected to affect the results of WDN partitioning. Therefore, the optimization of pipe weights to obtain optimal trade-off solutions between engineering aspects is an option that could be fruitful also in the case of other WDN partitioning algorithms.

Data Availability Statement

The readers can access the data upon request to the corresponding author. The software used in this study is made available upon request by the authors of the paper. The data related to the real water distribution network are confidential and can be provided with restrictions.

References

Alvisi, S., E. Creaco, and M. Franchini. 2011. "Segment identification in water distribution systems." *Urban Water J.* 8 (4): 203–217. https://doi.org/10.1080/1573062X.2011.595803.

Alvisi, S., and M. Franchini. 2013. "A heuristic procedure for the automatic creation of district metered areas in water distribution systems."

Urban Water J. 11 (2): 137–159. https://doi.org/10.1080/1573062X.2013.768681.

Campbell, E., J. Izquierdo, I. Montalvo, A. Ilaya-Ayza, R. Perez-Garcia, and M. Tavera. 2016. "A flexible methodology to sectorize water supply networks based on social network theory concepts and multiobjective optimization." *J. Hydroinf.* 18 (1): 62–76. https://doi.org/10.2166/hydro.2015.146.

Candelieri, A., D. Conti, and F. Archetti. 2014. "A graph based analysis of leak localization in urban water networks." *Procedia Eng.* 70 (Jan): 228–237. https://doi.org/10.1016/j.proeng.2014.02.026.

Ciaponi, C., E. Murari, and S. Todeschini. 2016. "Modularity-based procedure for partitioning water distribution systems into independent districts." *Water Resour. Manage.* 30 (6): 2021–2036. https://doi.org/10.1007/s11269-016-1266-1.

Clauset, A., M. E. J. Newman, and C. Moore. 2004. "Finding community structure in very large networks." *Phys. Rev. E* 70 (6): 066111. https://doi.org/10.1103/PhysRevE.70.066111.

Creaco, E., A. Di Nardo, M. Di Natale, C. Giudicianni, and G. F. Santonastaso. 2017. "Multi-object approach for WSN partitioning in the framework of pressure driven analysis." https://figshare.com/articles/CCWI2017_F19_Multi-object_approach_for_WSN_Partitioning_in_the_framework_of_Pressure_Driven_Analysis_/5364121/1.

Creaco, E., M. Franchini, and S. Alvisi. 2010. "Optimal placement of isolation valves in water distribution systems based on valve cost and weighted average demand shortfall." *Water Resour. Manage* 24 (15): 4317–4338. https://doi.org/10.1007/s11269-010-9661-5.

Creaco, E., M. Franchini, and E. Todini. 2016. "Generalized resilience and failure indices for use with pressure driven modeling and leakage." *J. Water Resour. Plann. Manage. (ISI)* 142 (8): 04016019. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000656.

Deb, K., S. Agrawal, A. Pratapm, and T. Meyarivan. 2002. "A fast and elitist multi-objective genetic algorithm: NSGA-II." *IEEE Trans. Evol. Comput.* 6 (2): 182–197. https://doi.org/10.1109/4235.996017.

Deuerlein, J. 2008. "Decomposition model of a general water supply network 357 graph." *J. Hydraul. Eng.* 136 (6): 822–832. https://doi.org/10.1061/(ASCE)0733-9429(2008)134:6(822).

Di Nardo, A., M. Di Natale, C. Giudicianni, R. Greco, and G. F. Santonastaso. 2016. "Weighted spectral clustering for water distribution network partitioning." *Appl. Network Sci.* 2 (1): 19.

Di Nardo, A., M. Di Natale, C. Giudicianni, G. F. Santonastaso, V. G. Tzatchkov, J. M. R. Varela, and V. H. A. Yamanaka. 2017. "Economic and energy criteria for district meter areas design of water distribution networks." *Water* 9 (463): 1–13.

Diao, K. G., Y. W. Zhou, and W. Rauch. 2013. "Automated creation of district metered areas boundaries in water distribution systems." *J. Water Resour. Plann. Manage.* 139 (2): 184–190. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000247.

Farley, M. 2001. "Leakage monitoring and control." In *Leakage management and control: A best practice training manual*, 58–98. Geneva: World Health Organization.

Ferrari, G., D. A. Savic, and G. Becciu. 2014. "A graph theoretic approach and sound engineering principles for design of district metered areas." *J. Water Resour. Plann. Manage.* 140 (12): 04014036. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000424.

Galdiero, E., F. De Paola, N. Fontana, M. Giugni, and D. Savic. 2016. "Decision support system for the optimal design of district metered areas." *J. Hydroinf.* 18 (1): 49–61. https://doi.org/10.2166/hydro.2015.023.

Giugni, M., N. Fontana, D. Portolano, and D. Romanelli. 2008. "A DMA design for 'Napoli Est' water distribution system." In *Proc., 13th IWRA World Water Congress*. Montpellier, France: International Water Resources Association.

Giustolisi, O., and L. Ridolfi. 2014a. "A new modularity-based approach to segmentation of water distribution network." *J. Hydr. Eng.* 140 (10): 04014049. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000916.

Giustolisi, O., and L. Ridolfi. 2014b. "A novel infrastructure modularity index for the segmentation of water distribution networks." *Water Resour. Res.* 50 (10): 7648–7661. https://doi.org/10.1002/2014WR016067.

Giustolisi, O., and D. Savic. 2010. "Identification of segments and optimal isolation valve system design in water distribution networks." *Urban Water* 7 (1): 1–15. https://doi.org/10.1080/15730620903287530.

Hajebi, S., E. Roshani, and N. Cardozo. 2016. "Water distribution network sectorisation using graph theory and many-objective optimisation." *J. Hydroinf.* 18 (1): 77–95. https://doi.org/10.2166/hydro.2015.144.

Herrera, M., E. Abraham, and I. Stoianov. 2016. "A graph-theoretic framework for assessing the resilience of sectorised water distribution networks." *Water Resour. Manage.* 30 (5): 1685–1699. https://doi.org/10.1007/s11269-016-1245-6.

Jun, H., and G. V. Loganathan. 2007. "Valve-controlled segments in water distribution systems." *J. Water Resour. Plann. Manage.* 133 (2): 145–155. https://doi.org/10.1061/(ASCE)0733-9496(2007)133:2(145).

Kirkpatrick, S., and C. D. Gelatt Jr. 1983. "Optimization by simulated annealing." *Science* 220 (4598): 671–680. https://doi.org/10.1126/science.220.4598.671.

Laucelli, D., B. A. Simone, L. Berardi, and O. Giustolisi. 2017. "Optimal design of district metering areas for the reduction of leakages." *J. Water Resour. Plann. Manage.* 143 (6): 04017017. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000768.

Liu, J., and R. Han. 2018. "Spectral clustering and multicriteria decision for design of district metered areas." *J. Water Resour. Plann. Manage.* 144 (5): 04018013. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000916.

Morrison, J. 2004. "Managing leakage by district metered areas: A practical approach." *Water* 21 (Feb): 45–46.

Newman, M. E. J. 2004a. "Analysis of weighted networks." *Phys. Rev. E* 70 (5): 056131. https://doi.org/10.1103/PhysRevE.70.056131.

Newman, M. E. J. 2004b. "Fast algorithm for detecting community structure in networks." *Phys. Rev. E* 69 (6): 066133. https://doi.org/10.1103/PhysRevE.69.066133.

Perelman, L., and A. Ostfeld. 2011. "Topological clustering for water distribution systems analysis." *Environ. Modell. Software* 26 (7): 969–972. https://doi.org/10.1016/j.envsoft.2011.01.006.

Perelman, L. S., M. Allen, A. Preis, M. Iqbal, and A. J. Whittle. 2015. "Flexible reconfiguration of existing urban water infrastructure systems." *Environ. Sci. Technol.* 49 (22): 13378–13384. https://doi.org/10.1021/acs.est.5b03331.

Walski, T. M., D. V. Chase, D. A. Savic, W. Grayman, S. Beckwith, and E. Koelle. 2003. "Introduction to water distribution modeling; using models for water distribution system design." In Vol. 4 of *Advanced water distribution modeling and management*. 1st ed., 333–337. Waterbury, CT: Haestad Press.

Zhang, Q. 2017. "Automatic partitioning of water distribution networks using multiscale community detection and multiobjective optimization." *J. Water Resour. Plann. Manage.* 143 (9): 04017057. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000819.

Zheng, F., A. R. Simpson, A. C. Zecchin, and J. W. Deuerlein. 2013. "A graph decomposition based approach for water distribution network optimization." *Water Resour. Res.* 49 (4): 2093–2109. https://doi.org/10.1002/wrcr.20175.

Queries

1. Please provide the ASCE Membership Grades for the authors who are members.

2. Please provide the postal code for the author "M. Cunha" in affiliations.

3. Please provide expansion for "INESC" for the author "M. Cunha" in affiliations.

4. Please check the hierarchy of section heading levels.

5. [ASCE Open Access: Authors may choose to publish their papers through ASCE Open Access, making the paper freely available to all readers via the ASCE Library website. ASCE Open Access papers will be published under the Creative Commons-Attribution Only (CC-BY) License. The fee for this service is $1750, and must be paid prior to publication. If you indicate Yes, you will receive a follow-up message with payment instructions. If you indicate No, your paper will be published in the typical subscribed-access section of the Journal.]

6. ASCE style for math is to set all mathematical variables in italic font. Please check all math variables throughout the paper, both in equations and throughout the text, to ensure all conform to ASCE style.

7. Please check all figures, figure citations, and figure captions to ensure they match and are in the correct order.

8. ASCE asks that paragraphs contain more than one sentence. As such, the sentence beginning "The logic . . . " was added to the previous paragraph. Please confirm no errors were introduced.

9. ASCE asks that paragraphs contain more than one sentence. As such, the sentence beginning "The sequence . . . " was added to the previous paragraph. Please confirm no errors have been introduced.

10. Please be aware that the print version of this article will appear in black and white. As such all references to color in Fig. 6 have been removed. Please confirm no errors were introduced.

11. By "Weibul frequency" do you mean "Weibull frequency" or something else?

12. The terms "graph a" and "graph b" were changed to "Fig. 8(a)" and "Fig. 8(b)." Please confirm no errors were introduced.

13. Please confirm that Weibul is correct in the caption for Fig. 8 (do you mean Weibull?).

14. ASCE asks that paragraphs contain more than one sentence. As such, the sentence beginning "If the trade-off . . . " was added to the subsequent paragraph. Please confirm no errors were introduced.

15. Please confirm that the term " N_disstart-tuple" is correct, and that no information is missing.

16. Please confirm that " N_bp − C_v,Dd " is correct and that this is an equation with a minus sign, not a range.

17. For Fig. 11, part label "a" was added based on the subcaption within the figure. This subcaption was then removed from the figure. Please check and confirm no errors were introduced.

18. ASCE asks that all abbreviations be defined on first reference. Please provide the expanded form, or definition of "GRF."

19. The print version of this article will print in black and white. As such, the phrase "the light blue and black DMAs" was changed to "the two DMAs." Please confirm this change is acceptable and no errors were introduced.

20. A check of online databases found an error in this reference. Please confirm the year has been changed to'2013'.

21. For reference Creaco et al. (2017), Please provide date of access in the following format: (Mon. DD, YYYY).

22. This query was generated by an automatic reference checking system. This reference could not be located in the databases used by the system. While the reference may be correct, we ask that you check it so we can provide as many links to the referenced articles as possible.

23. This query was generated by an automatic reference checking system. This reference could not be located in the databases used by the system. While the reference may be correct, we ask that you check it so we can provide as many links to the referenced articles as possible.

24. This query was generated by an automatic reference checking system. This reference could not be located in the databases used by the system. While the reference may be correct, we ask that you check it so we can provide as many links to the referenced articles as possible.