Uniformity and Heuristics-Based DeNSE Method for Sectorization of Water Distribution Networks

Vasilic, Željko; Stanic, Miloš; Kapelan, Zoran; Prodanovic, Dušan; Babic, Branislav

DOI
10.1061/(ASCE)WR.1943-5452.0001163

Publication date
2020

Document Version
Final published version

Published in
Journal of Water Resources Planning and Management

Uniformity and Heuristics-Based DeNSE Method for Sectorization of Water Distribution Networks

Željko Vasilic[1]; Miloš Stanic[2]; Zoran Kapelan[3]; Dušan Prodanovic[4]; and Branislav Babic[5]

Abstract: Sectorization of a water distribution network (WDN) into district metered areas (DMAs) is a proven solution for proactive leakage control. Traditionally, WDN sectorization is conducted by local experts using a trial-and-error approach, often resulting in the identification of arbitrary solutions. Some recently published methods try to improve WDN sectorization by automating the process, especially by using optimization. Various sectorization criteria, constraints, and limitations are introduced, which often fail to consider the issues faced by poorly managed WDNs such as limited funds and shortage of water balance data. These methods also have poor computational efficiency imposed by optimization methods used. This paper presents a new distribution network sectorization method (DeNSE), that overcomes these deficiencies. This method is based on a heuristic procedure in which WDN sectorization is driven by efficient tracking of water balance data and determining the lowest cost investment needed to maintain the same level of operational performance. The above-mentioned set of criteria is particularly well suited for initial sectorization of WDNs when major uncertainties in water balance data often lead to poor management decisions. The DeNSE method is validated and benchmarked against other sectorization methodologies in a case study of a large, real-world WDN. The results show that DeNSE can identify sound, realistic sectorization solutions that are in some respects better than corresponding solutions reported in the literature. DeNSE also enables high computational efficiency, ensuring its applicability to real-world WDNs. **DOI: 10.1061/(ASCE)WR.1943-5452.0001163.** *© 2019 American Society of Civil Engineers.*

Author keywords: Sectorization; District metered areas (DMA); Water distribution network (WDN); Uniformity; Distribution network sectorization (DeNSE).

Introduction

Sectorization of a water distribution network (WDN) into zones, sectors, clusters, or district metered areas (DMAs) has become one of the main strategies for efficient management of WDNs. It was introduced in the United Kingdom in the late 1980s and has been implemented in many WDNs worldwide. Sectorization has been done traditionally to meet two main objectives: better control of water losses and efficient management of pressures in the network. Sectorization is proven to be useful for other tasks such as protection against contamination (Chianese et al. 2017; Grayman et al. 2009). Burrows et al. (2000) provide the best definition of a DMA as a distinct hydraulic area of the WDN, separated from the rest of the supply system by isolation valves and one or more metered inlets and outlets.

Sectorization of WDNs into optimal DMA systems is difficult to achieve, especially in an existing and continuously operating WDN. Every WDN is unique in its topology, characteristics, and key drivers/objectives, so there is no standard procedure for sectorization, but rather a series of guidelines provided by the different water and other authorities (Butler 2000; Farley 2001; Morrison et al. 2007; WAA and WRC 1985; UK Water Research Industry 1999). Ideally, planning of DMAs (e.g., their number and size) should be carried out during the new WDN design phase, making it much easier to find the solution that will be efficient both in determining key sectorization objectives and satisfying the network's hydraulic and other requirements.

Complexity of the real-world WDN results in many different alternatives in which network sectorization can be done. Usually, sectorization is governed by the criterion of creating zones of manageable size in terms of number of consumers, links, or network length. Other important criteria (e.g., required number of feeds, fire flow regulations, etc.) and limitations may also apply. Sectorization solutions are usually obtained through trial-and-error methods carried out by local experts who are familiar with all of the WDN specifics. Practical application of such an approach is illustrated in a case study by Grayman et al. (2009), in which two large networks were redesigned to implement typical DMA design and to allow additional control and isolation of the system in order to improve water security. The need for a more formal approach to sectorization problem that will enable investigation of alternative sectorization solutions for large WDNs was recognized early (Tzatchkov et al. 2006).

[1]Assistant Professor, Faculty of Civil Engineering, Univ. of Belgrade, Bul. kr. Aleksandra 73, Belgrade 11000, Serbia (corresponding author). ORCID: https://orcid.org/0000-0002-9574-4509. Email: zvasilic@grf.bg.ac.rs

[2]Associate Professor, Faculty of Civil Engineering, Univ. of Belgrade, Bul. kr. Aleksandra 73, Belgrade 11000, Serbia. Email: mstanic@grf.bg.ac.rs

[3]Professor, Faculty of Civil Engineering and Geosciences, Delft Univ. of Technology, Bldg. 23, Stevinweg 1, Delft 2628 CN, Netherlands; Professor, College of Engineering, Mathematics and Physical Sciences, Univ. of Exeter, Harrison Bldg., North Park Rd., Exeter EX4 4QF, UK.

[4]Professor, Faculty of Civil Engineering, Univ. of Belgrade, Bul. kr. Aleksandra 73, Belgrade 11000, Serbia.

[5]Assistant Professor, Faculty of Civil Engineering, Univ. of Belgrade, Bul. kr. Aleksandra 73, Belgrade 11000, Serbia.

Note. This manuscript was submitted on December 10, 2018; approved on July 23, 2019; published online on December 21, 2019. Discussion period open until May 21, 2020; 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.

Different algorithms for automated sectorization of the WDN into DMAs have been developed and presented in recent years, together with the tools that can be used to support the process (Deuerlein 2008; Perelman and Ostfeld 2012). In general, existing algorithms for automated sectorization have three general steps (Perelman et al. 2015): (1) division of the WDN into clusters, (2) placement of valves and flow meters on cluster's boundary pipes to define the DMAs, and (3) evaluation of solutions based on the previously adopted performance indicators (PIs). For the purpose of initial division of the WDN (Step 1), most presented methodologies rely on the graph theory algorithms (Alvisi and Franchini 2014a; Di Nardo et al. 2013; Ferrari et al. 2014; Hajebi et al. 2016; Scarpa et al. 2016) or multiagent approach and spectral clustering (Di Nardo et al. 2018; Herrera et al. 2010a, 2010b), while others use the modularity index (Ciaponi et al. 2016; Giustolisi and Ridolfi 2014; Laucelli et al. 2016; Campbell et al. 2016) or community structure metrics (Diao et al. 2013; Zhang et al. 2017; Brentan et al. 2017) originally presented by Clauset et al. (2004) and Newman and Girvan (2004). Modularity and community structure metrics are introduced from other fields of research and are based on similarity between clusters based on the weights assigned to the links. Motivation for application of community structure metrics comes from the fact that many complex systems such as WDNs have the property of higher links density within the communities than between them (Fortunato 2010; Giudicianni et al. 2018). These metrics have been tailored in different ways to use for WDN sectorization (Giustolisi and Ridolfi 2014; Zhang et al. 2017). Although these approaches are able to determine DMAs, they are sensitive to the selection of links weights (Ciaponi et al. 2016; Diao et al. 2013). Thus far, presented sectorization methods mainly include cluster (DMA) size range (min−max) and reachability from the transmission main as the sectorization governing variables. Identifying DMAs that are as uniform in size as possible is addressed in research presented here, hypothesizing that uniformity of DMAs' sizes can be a suitable variable to govern the sectorization process.

A large number of possible alternatives exist for positioning the valves and flow meters to define the DMAs (Step 2) in a real-world-sized WDN. Many of these alternatives are not feasible because they do not meet the basic hydraulic requirements for WDN operation. For the purpose of selecting the (near) optimal alternative, a sectorization algorithm is usually coupled with some type of optimization method (Alvisi 2015; Giustolisi and Ridolfi 2014; Hajebi et al. 2016; Laucelli et al. 2016; Zhang et al. 2017) that requires a significant amount of computational resources. Thus far, computational efficiency has been regarded as something of secondary importance, with primary focus on the quality of the obtained solution. Viable alternatives to traditional optimization methods are heuristics-based approaches for positioning of the valves and flow meters (Alvisi and Franchini 2014a; Ciaponi et al. 2016; Diao et al. 2013) or the use of a simplified hydraulic simulator that can quickly find near-optimal solutions (Alvisi and Franchini 2014b).

Number and type of PIs, used in Step 3 to assess the effect of implemented interventions and evaluate the sectorization solution, vary significantly in the research literature. The resilience index, as described in Todini (2000), is present in almost all research as a measure of network postsectorization reliability. Water age is usually used to reflect the impact on water quality in the network. Some researchers added various other indices to validate the feasibility of obtained solutions; for example, pressure indices are used in Di Nardo et al. (2013) and an entropy index is used in Scarpa et al. (2016).

Some of the drawbacks of available methods for automated sectorization that may question their applicability to real-world WDNs are associated with: (1) comprehensive lists of objectives and constraints used in optimization, (2) computational efficiency, and (3) resolution of the sectorization solution.

In the process of developing new methods, various limitations and constraints important for the proper functioning of the WDN have been implemented in optimization procedures (Di Nardo et al. 2017; Gomes et al. 2012; Zhang et al. 2019). Chronologically, only DMA size and network pressure constraints have been considered (Di Nardo and Di Natale 2011), with each new method adding new sectorization parameters and network PIs to their lists of limitations and constraints. Probably the most comprehensive such list is presented in Hajebi et al. (2016), having 13 objectives and 11 constraints. In fact, these lists may have grown too much, exhausting all practical aspects important for normal everyday operation of the WDN. Optimization methods are computationally expensive by their nature, and the addition of new objective functions by each sectorization method only highlights this effect, as in, for example, the algorithms of Hajebi et al. (2016) and Zhang et al. (2017). Solution search space exponentially increases with the complexity of a network, and perhaps this is why recently presented methods employing optimization lack results supporting their application on large-sized networks (Alvisi 2015; Laucelli et al. 2016).

Water utilities operating poorly managed WDNs usually do not have sufficient funds to invest in large numbers of DMAs at once, so the sectorization process should be planned hierarchically and implemented in phases, starting with a few DMAs that can be larger than recommendations given in the guidelines. Establishing a few DMAs in a WDN should enable tracking the water balance in the network and gathering basic data about system dynamics, without significant effect on the network's operational conditions. This could improve operational management of WDNs, as management decisions are usually made based on some calculated WDN's PIs, which can have values significantly influenced by great uncertainty of available water balance data (Babić et al. 2014). With increased resolution of the sectorization, it is usually required that new DMAs keep previously created boundaries of the original DMA layout. In this way, costs are minimized, which addresses the economics of sectorization. Scarpa et al. (2016) considered hierarchical sectorization based on progressive union of initially identified elementary DMAs; this can be viewed as bottom-up approach. A top-down approach to sectorization would be more aligned with an engineering perspective and more in accordance with the phased creation of DMAs. Either way, hierarchy should be considered in sectorization solutions.

From the previous discussion, it can be concluded that, despite all recent advancements made, there are opportunities to further improve existing water network sectorization algorithms. Key areas for improvement are in: (1) implementing practical engineering principles, relevant to the WDN, to govern the sectorization process; (2) improving computational efficiency of the algorithm; and (3) considering hierarchical sectorization.

In the method presented here, distribution network sectorization (DeNSE), the first area is addressed with implementation of a newly presented network uniformity index (Vasilic 2018) that drives WDN decomposition into clusters that are not only within predefined size limits, but also uniform in size as much as possible. The uniformity index also favors sectorization in which the cluster's connecting links are pipes with smaller diameters, indirectly providing an economically more favorable solution because it is less costly to install valves and flow meters on smaller diameter pipes. High computational efficiency is achieved using common-sense engineering heuristics, rather than optimization tools, to position the valves and flow meters on the connecting links and define the DMAs. The network clustering algorithm evolves in a step-by-step manner, and thus the obtained sectorization solution is inherently hierarchically ordered. Furthermore, the algorithm presented here does not determine just one sectorization solution, but a range of feasible solutions, giving decision makers the flexibility to select the one best suited for their needs. The algorithm is tested against a benchmark, a large operating network presented in the Battle of the Water Sensor Networks (BWSN2) (Ostfeld et al. 2008), and results are thoroughly compared with other results previously reported in the literature.

Methodology

This paper presents uniformity and heuristics-based methodology for WDN sectorization into DMAs, called DeNSE, which is also able to address hierarchical sectorization. The algorithm is based on the graph theory for identification of strong connected components (SCCs) and their aggregation into clusters based on a newly presented network uniformity index (U). As discussed in the Introduction, the sectorization process should start with the definition of key sectorization objectives and design criteria, followed by the identification of PIs that will be used to assess the impact of interventions made in the network. Tracking the water balance in the network is the main sectorization objective adopted in the DeNSE method. Designing the sectorization solution that requires the least cost investment in the equipment necessary for creation of DMAs (flow meters and isolation valves), while maintaining the network's existing operational efficiency are the main design criteria. Such design criteria are most appealing to many water utilities, especially in developing countries, which operate highly inefficient WDNs with a significant amount of water and revenue losses. Two PIs are adopted to evaluate the effects of the sectorization on a network's operational performance: (1) resilience index (*Res*), reflecting post-sectorization reliability of the WDN (Todini 2000) and (2) water age (*WA*), surrogate metrics for water quality reflecting water retention rate in the WDN.

The new method requires a hydraulic model of the WDN as an input, like many other methods relying on it to prove hydraulic feasibility of a sectorization solution. The quality of the adopted solution will be better if a calibrated hydraulic model is used, and required interventions in the network can be taken with more assurance in preservation of the network's hydraulic performance. The method runs through three stages to identify the best sectorization solution, as shown in Fig. 1. The first stage is a preprocessing stage in which all the relevant network data are obtained from the WDN model and prepared for the following run of the clustering algorithm. WDN decomposition into clusters is done in the second stage, based on the uniformity index. The third stage involves narrowing choices to the few feasible solutions that will be hydraulically analysed. The third stage includes heuristic, engineering-based positioning of the valves and flow meters on clusters connecting links to define DMAs and extended-period hydraulic analysis of the solutions and evaluation of adopted PIs. Finally, feasible solutions are ranked and the preferred solution is

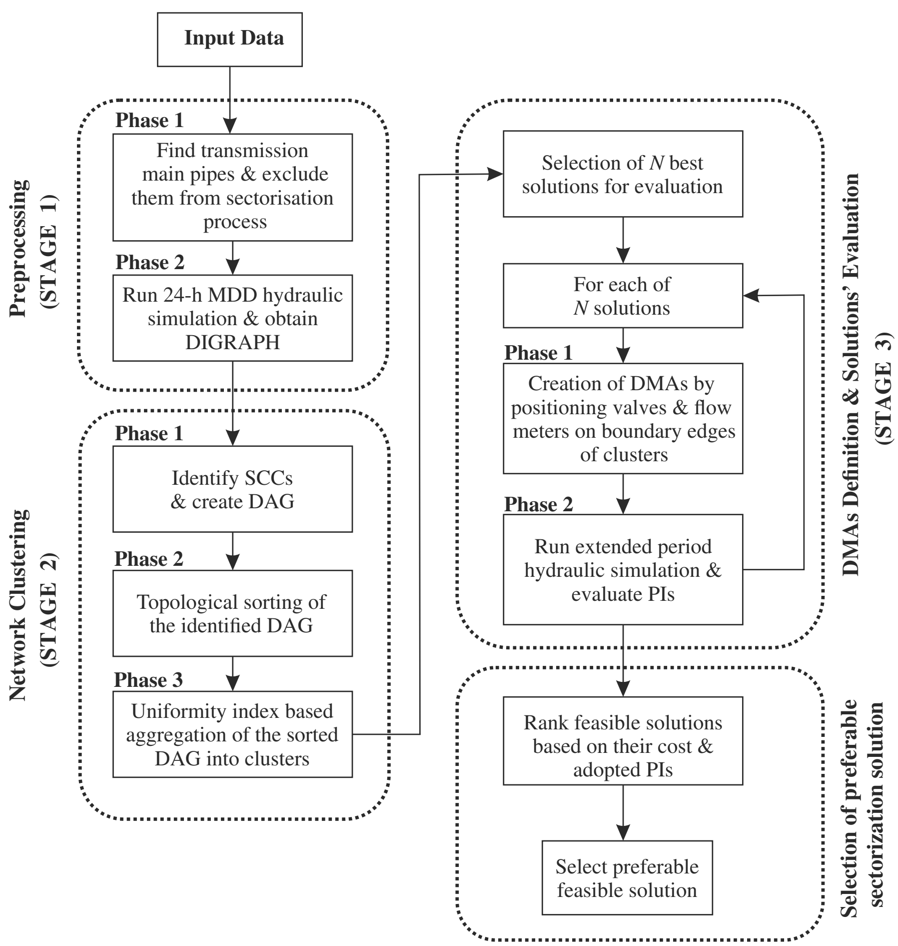

Fig. 1. Flowchart of the DeNSE sectorization method.

selected. Each of the three stages will be explained in detail in the sections below.

Input Data

The new sectorization method requires the following input data:

1. Calibrated WDN network model in the form of EPANET input file, which contains all relevant data (topology, hydraulic characteristics, demand data, etc.).
2. Minimum (n_c^{min}) and maximum (n_c^{max}) number of property connections per DMA, as well as total number of connections in the network (n_c), since number of connections per node is usually not available with the mathematical model. Recommendations about these values can be found in a number of available guidelines for DMA creation; usually, it is considered that the number of connections should be in the range of 500–5,000 (Farley 2001; Morrison et al. 2007). It is considered that having DMAs larger than 5,000 connections is not practical because it becomes difficult to distinguish leakages from the night flow data, while taking more time to allocate them. It should be noted that the preferable DMA size is network specific, influenced by many factors, and should be determined based on a thorough analysis of the specific data relevant to the network in consideration.
3. Transmission main threshold diameter (D_{main}). Large diameter pipes connected in series, running from the network's main source(s), are considered a transmission main. These are the pipes that convey water between the reservoirs and tanks and serve as main supply paths in the network. In this methodology, they are excluded from any interventions. As with the DMA size, the value of D_{main} is network specific, usually 300–350 mm (Ferrari et al. 2014).
4. Pipe closure threshold diameter (D_{tr}). Pipes with diameters equal to or larger than this diameter will not be considered for possible closure for positioning the valves and flow meters (part of Stage 3). By default, the algorithm uses the first class of diameter lower than the D_{tr} (e.g., if D_{main} is 350 mm, D_{tr} will be 300 mm), but the user can specify a different value. However, this will affect the number of isolation valves and flow meters required to create the DMAs and, consequently, the solution cost.
5. Minimum required and maximum allowed pressures in the network, p_{min} and p_{max}, as well as the maximum water age (WA_{max}) allowed in the network as a water quality indicator.
6. Desired number of sectorization solutions (N_{sol}). It is considered that 10–15 solutions is large enough to make representative multicriteria ranking; however, the user can opt for a larger set of solutions to compare.

Preprocessing (Stage 1)

In the first stage, there are two phases (Fig. 1). In the first phase, transmission mains are defined, based on the D_{main} value, and excluded from the sectorization process. For this purpose, the network is explored using a slightly modified breadth first search (BFS) algorithm (Jungnickel 2005), simultaneously starting from all main source nodes (reservoirs). The BFS algorithm is modified to prioritize propagation through the links with diameters equal to or greater than D_{main}. In the second phase, 24-h maximum day demand (MDD) hydraulic simulation of the analyzed WDN is performed to determine the orientation of pipes (based on water flow directions obtained in the simulation). As a result, directional graph (digraph) G is defined with two sets $G = \langle N, C \rangle$, set of network nodes N and set of network links C, where each link is presented

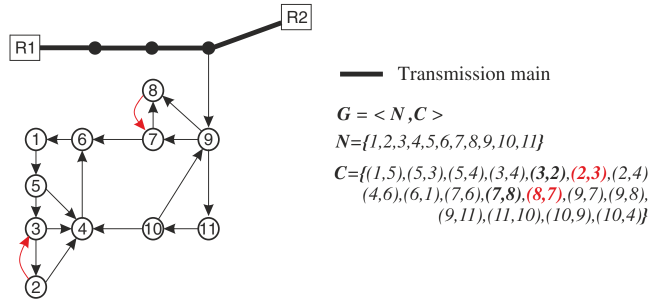

Fig. 2. Digraph presentation of a simple network with two sources and two undirected links.

with ordered pair of nodes. Network links with changing flow directions are identified as nonoriented (or links that can have both flow directions), and are represented with the addition of a fictitious link in the opposite direction. This network representation is used only for identification of SCCs in Stage 2, and original network topology is used for hydraulic simulations. Both phases are illustrated in a simple example network shown in Fig. 2.

The example network consists of 16 nodes, 2 of which are reservoirs, and 21 links. Links connecting reservoirs are identified as transmission mains and are excluded from further analysis. The remaining part of the network, connected to the transmission main with one link in node 9, should be partitioned into DMAs. Illustrated orientations of the remaining links are determined based on the results of the hydraulic analysis. Two of these links are identified as not oriented, and putting that in the context of water networks, these are usually pipes (links) that connect tanks with the rest of the network. So, in an example network, two fictitious links are added (2–3 and 8–7) and nodes 8 and 2 could be tanks.

Network Clustering (Stage 2)

In the second stage of the DeNSE method, partitioning of the WDN into clusters is performed. It is done in three phases (Fig. 1).

Phase 1

The first task is to identify the strongly connected components within the previously created digraph. The SCC, a term from graph theory, is defined as a subgraph in which each node can be reached from any other node within that subgraph (Gabow 2000). Essentially, an SCC is a directed cyclic component in which flow direction within that component can reverse (Perelman and Ostfeld 2012). Therefore, SCCs are parts of the network where water is circulating during the simulation (Vasilić et al. 2016), and thus, control of the water balance and/or water pressure regulation in SCC parts of the network could be difficult to achieve, so SCCs should be detected and treated as aggregated nodes in further network analysis and clustering. Algorithms for the extraction of SCCs from digraphs are well known in graph theory. The Gabow algorithm (Gabow 2000) is used in the methodology shown here. It is chosen due to its linear computational time, which makes it more efficient than others. This is significant because the algorithm must be able to deal with large networks efficiently. Gabow's algorithm requires only one pass through the network (digraph) with a recursive call of the depth first search (DFS) algorithm (Tarjan 1972) with arbitrary selection of the starting node.

For illustration purposes, a simple digraph, shown in Fig. 2, is used. Starting the DFS search from node 2, nodes 3, 4, 6, 1, and 5 are visited [Fig. 3(a)]. During the DFS search, a check is made to determine whether selection of the next node forms a cyclic path.

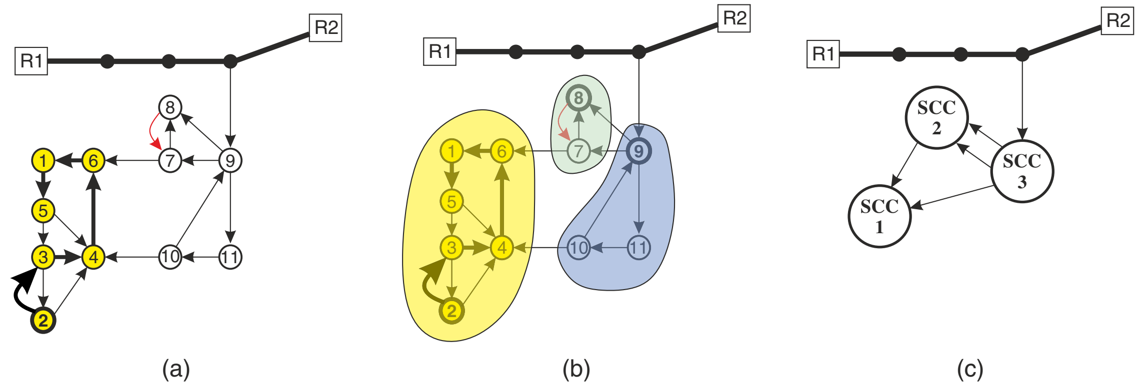

Fig. 3. Digraph transformation to DAG: (a) start the DFS; (b) detected SCCs; and (c) newly formed DAG.

If yes, nodes forming the cyclic path are identified as an SCC. The algorithm continues until no further propagation is possible. In the example shown in Fig. 3, the first SCC component identified is composed of nodes 2, 3, 4, 6, 5, and 1. No further propagation is possible, so the DFS starts again from a randomly selected node, chosen from the set of nodes that were not visited during the first search. Assuming that the randomly selected node is node 9, and after nodes 11 and 10 are visited, the second SCC composed of these three nodes is identified. A DFS search is repeated again starting from node 8, and the third SCC composed of nodes 8 and 7 is detected [Fig. 3(b)]. At the end, the aggregated digraph is composed of three identified SCCs. The digraph can also be viewed as a set of aggregated nodes connected to the transmission main. The most important property of the new aggregated digraph is its acyclicity, indicating it is a digraph without cycles. Such a graph is referred to as a directed acyclic graph (DAG) and, in a water network, it is important because it clearly separates source from the demand nodes and, hence, makes the sectorization of the network easier.

Phase 2

In the second phase, topological sorting of the identified DAG is conducted. DAG nodes, represented with SCCs, are sorted from the downstream end, and this order will be used to drive aggregation of the DAG from the most peripheral SCCs. Again, simple implementation of the recursive DFS algorithm, as explained in Sedgewick and Wayne (2011), is used for this purpose. In an example shown in Fig. 3(c), topological sorting yields the following list of SCCs: SCC1, SCC2, and SCC3.

Phase 3

In this phase, aggregation of the sorted DAG—composed of the SCCs connected between each other and connected to the transmission main—is conducted based on the newly presented network uniformity index (U). The network uniformity index (Vasilic 2018) is defined as follows:

$$U = u_{net} u_v w_{agg} \tag{1}$$

where u_{net} = network uniformity in terms of cluster size; u_v = uniformity of the DMA's size vector; and w_{agg} = relative weight of aggregated links. Each of these variables is explained below, followed by the explanation of the aggregation algorithm itself.

Each cluster is characterized with its size (S_i), calculated as the sum of all nodal demands within that cluster, $S_i = \sum_{j=1}^{N_n^i} q_j$, with N_n^i being number of nodes in ith cluster. Network uniformity (u_{net}) measures average deviation of cluster size from the preferred DMA size (S_{pref}). Ideally, all clusters should have size equal to S_{pref} but, obviously, this is not possible in real networks. Preferred DMA size is calculated based on minimum and maximum DMA size, $S_{\min}$ and $S_{\max}$, as $S_{pref} = \frac{S_{\min}+S_{\max}}{2}$. Minimum and maximum DMA sizes are calculated based on daily average total demand in the WDN (Q_{tot}, available from the WDN hydraulic model), number of minimum and maximum connections in the DMA ($n_c^{\min}$ and $n_c^{\max}$), and total number of connections in the WDN (n_c), given as input data, as follows:

$$S_{\min} = \frac{Q_{tot}}{n_c} n_c^{\min}$$
$$S_{\max} = \frac{Q_{tot}}{n_c} n_c^{\max} \tag{2}$$

Network uniformity is calculated based on the triangular function f that quantifies "quality" of cluster size in the range [0,1] (Fig. 4). If a cluster i has a size $S_i = S_{pref}$, its value of f will be the best, that is, $f_i = 1$. If a cluster has a different size (larger or smaller than S_{pref}), it will have the value of $f_i < 1$. Since the function f is equilateral, both larger and smaller clusters are equally penalized. Extremely large clusters (larger than S_{pref}), are scored with the lowest value of $f_i = 0$. Potentially, other types of function f that will penalize small and large clusters in different rates could be used, but the triangular function currently implemented provided the most consistent results. Finally, network uniformity is calculated as

$$u_{net} = \frac{\sum_{i=1}^{N_{cl}} f_i}{N_{cl}} \tag{3}$$

where N_{cl} = number of clusters for a given sectorization. Note that maximum value of u_{net} is 1 if all clusters are equal to S_{pref}, and minimum value is zero.

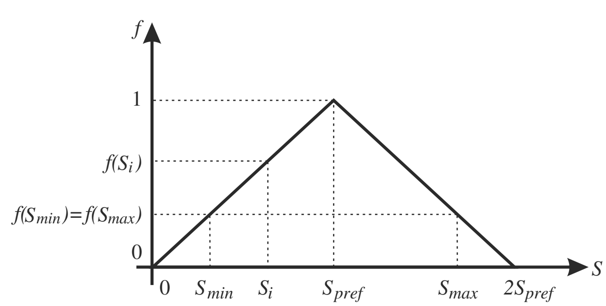

Fig. 4. Triangular function f quantifying cluster size "quality."

Sizing clusters in the range $S_{\min}$–$S_{\max}$, and as close to S_{pref} as possible, is one sectorization objective. Sizing them equally is the other objective. Sizes of all clusters form the normalized size vector of specific sectorization into N_{cl} clusters, $\boldsymbol{S^n} = \{S_1^n, S_2^n, S_3^n \cdots, S_{N_{cl}}^n\}$, where $S_i^n = \frac{S_i}{\sum_{i=1}^{N_{cl}} S_i}$. Uniformity of this vector is calculated as its Euclidean norm (L2 norm) as follows:

$$u_v = \sqrt{\sum_{i=1}^{N_{cl}} (S_i^n)^2} \tag{4}$$

If all clusters are equal in size (e.g., $S_1 = S_2 = S_3 = \ldots = S_{pref}$), which is the most preferable case, uniformity of the size vector is

$$u_v^{best} = \sqrt{\left(\frac{S_1}{N_{cl}S_{pref}}\right)^2 + \left(\frac{S_2}{N_{cl}S_{pref}}\right)^2 + \ldots}$$
$$= \sqrt{\frac{N_{cl}(S_{pref})^2}{N_{cl}^2(S_{pref})^2}} = \sqrt{\frac{1}{N_{cl}}} \tag{5}$$

If all nodes are part of the same cluster—the worst-case scenario in which there is no clustering—uniformity of the size vector is $u_v^{worst} = 1$. To be consistent with the ranging values of network uniformity metrics (u_{net}), where 0 is the minimum value and 1 is maximum, uniformity of the size vector is scaled to the same range to yield the final form of the equation for its calculation, as follows:

$$u_v = \begin{Bmatrix} 1 - \frac{u_v\sqrt{N_{cl}} - 1}{\sqrt{N_{cl}} - 1}; & N_{cl} > 1 \\ 0; & N_{cl} = 1 \end{Bmatrix} \tag{6}$$

Relative weight of aggregated links is calculated as

$$w_{agg} = \frac{\sum_{i=1}^{n_l^{agg}} D_i}{\sum_{i=1}^{n_l} D_i} \tag{7}$$

where n_l = total number of links; n_l^{agg} = number of links within the clusters; and D_i = links diameter. In case of a large number of clusters, there will be more unaggregated connecting links than in the case of a small number of clusters. Hence, the value of w_{agg} will be smaller in the former than in the latter case. The minimum value of w_{agg} is 0 if no aggregation is done, and 1 if all SCCs are aggregated into one cluster.

Aggregation of SCCs into clusters, based on uniformity index metrics described above, is done in a step-by-step manner, propagating upstream through the topologically sorted DAG made of SCCs (obtained in Phase 2) and aggregating in each step SCCs whose aggregation will contribute the most to the network uniformity (Vasilic 2018). Initially, all identified SCCs are considered as individual clusters, meaning that the initial number of clusters corresponds to the number of identified SCCs. Aggregation is iteratively carried out through three steps: (1) identification of candidate SCCs for aggregation, based on topologically sorted DAG; (2) selection and aggregation of the candidate with highest uniformity gain ($\Delta U_{\max}$); and (3) aggregation of remaining downstream SCCs with positive uniformity gain ($\Delta U > 0$). The third step in this iterative aggregation procedure is implemented to avoid the scenario in which small peripheral SCCs remain unaggregated until the late stages of aggregation, which could happen because these SCCs usually have relatively small uniformity gain, and aggregation would continue past them further upstream.

Uniformity index metrics that drive the clustering process is made of three components as given in Eq. (1). Because the aggregation process is driven with the highest uniformity gain ($\Delta U_{\max}$), it is of interest to maximize all three components of the network uniformity index (u_{net}, u_v, and w_{agg}). Maximizing w_{agg} implies that the links with the larger weights (diameters) are aggregated first. In this manner, links with smaller diameters will be left as connecting links between the clusters which, in turn, provides a more economically favorable sectorization solution.

The aggregation algorithm presented here is essentially a greedy optimization method in which aggregation direction is determined based on the highest uniformity index gain ($\Delta U_{\max}$). This is similar to the greedy optimization method, based on highest modularity gain, used to maximize a network's modularity index presented in Clauset et al. (2004). As with all similar types of algorithms, it is not guaranteed that the global optimal solution will be found. However, the benefit is that, generally, a good suboptimal solution can be found with significant computational time savings when compared with other optimization algorithms. The algorithm is deterministic in nature, and it will always provide the same results as long as the same input parameters are given.

Application of the described aggregation algorithm is illustrated in a simple example shown in Fig. 5. The example is derived from Fig. 3(c), adding six more SCCs for illustration purposes. For the sake of simplicity, total demand of 20 L/s is assigned to all nine SCCs. Diameters of the links connecting SCCs are shown in millimeters in Fig. 5. Minimum ($S_{\min}$) and maximum ($S_{\max}$) DMA size are set to 40 and 80 L/s, respectively, which yields the preferred DMA size (S_{pref}) of 60 L/s. Fig. 5 shows evolution of the network uniformity index through the aggregation process of this simple example. Uniformity index (U) is plotted against the number of clusters corresponding to each aggregation step (secondary horizontal axis).

The highest uniformity index value ($U_{\max}$) corresponds to network sectorization into three clusters with total demands of 40, 60, and 80 L/s. Sizes of all three clusters are within predefined DMA size limits (40–80 L/s). Clusters are connected with three links among them. The next aggregation step leads to the solution with two clusters having total demands of 80 and 100 L/s. Obviously, this solution does not meet DMA size constraints, because one cluster is larger than $S_{\max}$. However, there are now two links connecting two clusters, which requires fewer isolation valves and flow meters to isolate them and create DMAs than in the case with three clusters. Fig. 5 also illustrates hierarchical ordering of the sectorization solutions embedded in the clustering algorithm. The solution with three clusters is lower in the hierarchical order, and is easily derived from the solution with two clusters.

Heuristic Device Placement and Evaluation of Solutions (Stage 3)

At the end of Stage 2, the clustering of DAG made out of identified SCCs based on network uniformity index is finished. As described above, clustering is done in a step-by-step manner, preserving data about the clusters' structure at each aggregation step (Fig. 5). Note that number of aggregation steps corresponds to the number of identified clustering solutions. Obviously, not all of the solutions obtained are of interest, only the ones with a high value of network uniformity index.

Prior to execution of Stage 3, selecting the solutions that will be hydraulically analyzed and evaluated for satisfaction of initially adopted PIs is conducted. Number of solutions (N_{sol}) for Stage 3 analysis is specified by the user as an input parameter. Selection of solutions is made based on the network uniformity index values

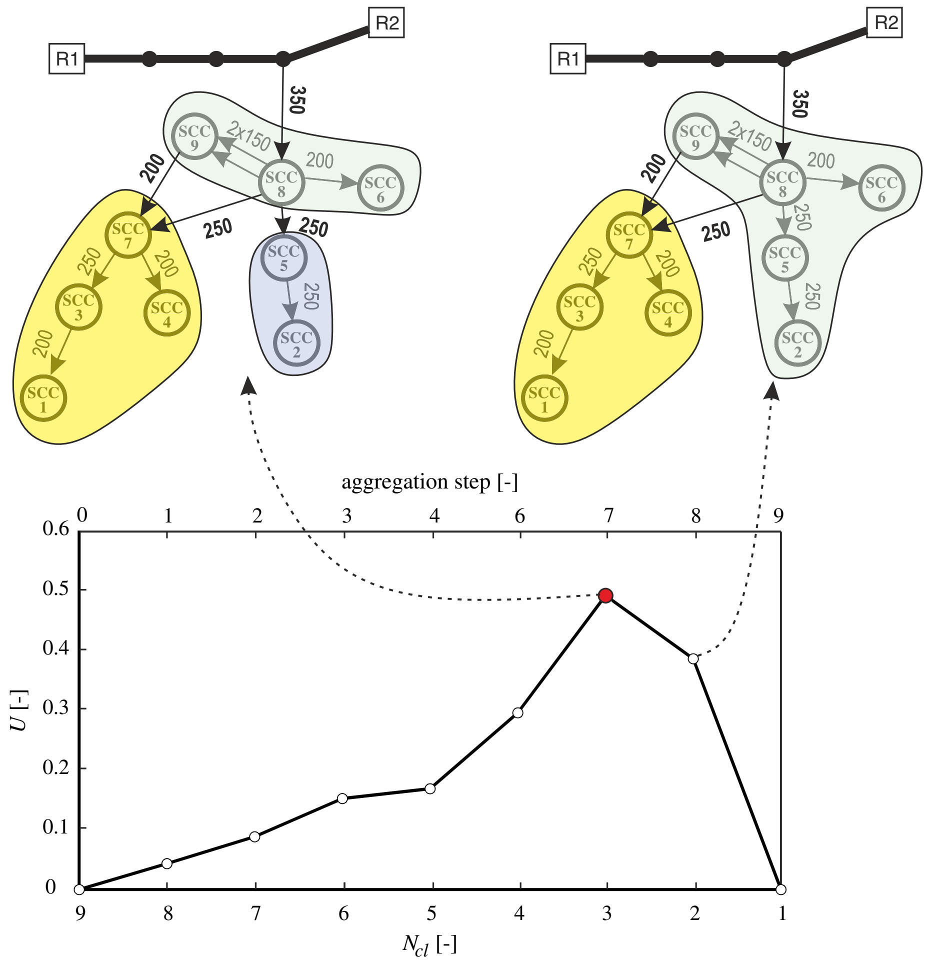

Fig. 5. Evolution of network uniformity index during aggregation process.

obtained at each aggregation step. The solution with the highest uniformity index (best solution) is selected, together with additional $N_{sol} - 1$ solutions from succeeding aggregation steps. Additional solutions are on the descending part of the uniformity index plot (Fig. 5), characterized by the lower value of uniformity index (than the best solution) and by the smaller number of clusters. The described strategy for selection of solutions adopted here is particularly well suited for the application at the initial stages of the DMA design process. For coarser sectorization, solutions can be chosen from the ascending part of the uniformity index plot as well. Clusters connected only to the transmission main and having size smaller than $S_{\min}$ are removed from each solution and excluded from further analysis. Such clusters are below minimum DMA size limit and will not be considered as a DMA.

After the solutions that will be evaluated have been selected, the main part of Stage 3 is conducted. There are two main phases in Stage 3: Phase 1 is conversion of clusters into DMAs and Phase 2 is evaluation of solutions.

Phase 1

To convert clusters into DMAs (i.e., define DMAs), flow meters and isolation valves must be positioned on the clusters' boundary edges. Positioning of the flow meters and valves is done based on engineering heuristics. Continuing from the simple example used to describe the aggregation algorithm (Fig. 5), consider the solution that has the highest network uniformity index value. This solution has three clusters and four boundary edges to be considered for installation of flow meters/valves. For methodology illustration purposes, another branch of the transmission main and four boundary edges are added to this solution [Fig. 6(a)].

Boundary edges are labeled L1 through L8, and numbers show the links' diameters in millimeters. Flow orientations during 24-h MDD hydraulic simulation, obtained in Stage 1-Phase 1, are indicated with arrows. Pipes with a changing direction (nonoriented) are indicated using dashed lines without arrows. Nonoriented pipes are only those that connect clusters with the transmission main, as identified clusters resulted from the DAG analysis (i.e., all other nonoriented pipes are already aggregated with the identification of SCCs in Stage 2-Phase 1). In this case, there is only one such pipe (L2). The heuristic procedure comprises the following three steps:

- Nonoriented pipes are identified, and all pipes in which absolute difference between the maximum and minimum flow rate is less than 0.2 L/s are marked for closure, as this is considered negligible flow rate (hypothetically, let L2 be such a pipe in this example).
- All links connecting clusters with the transmission main, oriented from the clusters to the main, are closed (L3 and L8 in the example shown). These are the pipes always returning the water from the demand nodes into the main, hence it is considered that they are not supply pipes and can be closed without negative effects on system hydraulics.
- Supply pipes of each cluster (oriented toward cluster) are analyzed independently. It is sufficient to analyze only supply pipes

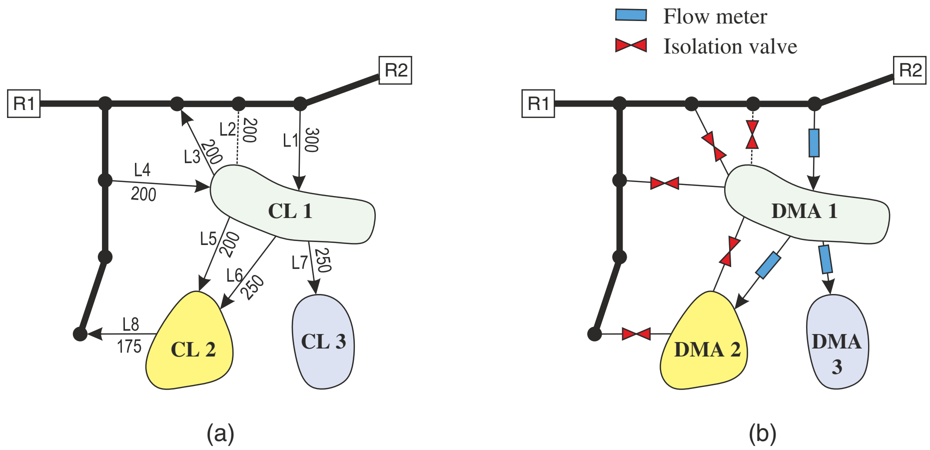

Fig. 6. Heuristics-based placement of flow meters and isolation valves (Stage 3-Step 1).

becasuse the graph in consideration is a DAG and one cluster's output pipes are others' supply pipes. Supply pipes for a cluster are identified; the pipe with the largest maximum inflow to the cluster ($Q_{\max}$) is considered the main supply pipe, and will not be considered for closure. The maximum capacity of this pipe ($C_{\max}$) is calculated based on maximum allowable velocity of 2.0 m/s, and its remaining capacity is $C = C_{\max}–Q_{\max}$. All remaining supply pipes having diameter larger than threshold value, given as an input (D_{tr}), are candidates for closure. Their maximum capacities are calculated in the same manner ($c_{\max}$), and they are analyzed one by one, starting from the link with the lowest maximum flow rate ($q_{\max}$). When a pipe i is considered for closure, resulting residual input capacity (C_{cl}) is calculated subtracting ith pipe capacity as $C_{cl} = C + \sum c_{\max} - c_{\max}(i)$. If reduced capacity is still larger than the maximum flow rate carried by the ith pipe [$C_{cl} \geq q_{\max}(i)$], the pipe is closed by setting its capacity to zero [$c_{\max}(i) = 0$]. Iterating through this procedure, candidate pipes are closed until input capacity is fully exhausted. Hypothetically, applying this to the simple example in Fig. 6 would result in closure of supply pipe L4 for cluster CL1 and pipe L5 for cluster CL2. Cluster CL3 has only one supply link, so it remains open.

Another approach for positioning flow meters and valves is the optimization method, for example, the genetic algorithm (Ivetić et al. 2013), which considers each boundary pipe as closed or open. Because it is not uncommon in real-world WDNs for the number of boundary edges to exceed several tens, the optimization method could be very time consuming, and thus it was not implemented here. At the end of Phase 1, flow meters and isolation valves are positioned on the clusters' boundary edges, converting them into DMAs [Fig. 6(b)].

Phase 2

After definition of its DMA boundaries, each solution is subjected to the extended-period hydraulic simulation to investigate the effects of modifications made to the network. First, feasibility of the solution is considered through evaluation of pressure constraints in each node, as follows:

$$p_{i,t} \geq p_{\min}; \qquad p_{i,t} \leq p_{\max} \tag{8}$$

where $p_{i,t}$ = pressure in ith node in simulation time step t; and $p_{\min}$ and $p_{\max}$ = minimum and maximum allowable pressures in the network. If the solution does not meet pressure constraints, it is considered infeasible and is excluded from further analysis.

For each feasible solution, cost and two adopted PIs are calculated as follows:

1. Cost of the solution is calculated based on the unit cost of devices installed to create the DMAs (flow meters and isolation valves). Unit cost functions are taken from De Paola et al. (2014).
2. Average network resilience index (Todini 2000) is calculated as mean value over the simulation time period (T). The resilience index is represented as the ratio of residual amount of power in the network after satisfaction of nodal demands and maximum amount of power that can be dissipated in the network internally, while satisfying nodal demands and minimal pressure constraints:

$$Res = \underset{T}{mean}\left(\frac{\sum_{i=1}^{n_j} q_i(h_i - h_i^*)}{\sum_{j=1}^{n_r} Q_j H_j + \sum_{k=1}^{n_p} \frac{P_k}{\gamma} - \sum_{i=1}^{n_j} q_i h_i^*}\right) \tag{9}$$

where n_j = number of junctions; n_r = number of reservoirs; n_p = number of pumps; q_i = nodal demand at node i; h_i = nodal head at node i; h_i^* = minimum nodal head at node i; Q_j = discharge from the reservoir j; H_j = head in reservoir j; P_k = amount of power introduced in the network by pump k; and γ = specific weight of the water.

3. Average water age in the network over the last 24 h of extended-period simulation (WA) is calculated as

$$WA = \frac{\sum_{i=1}^{n_j} \sum_{t=T-24}^{T} WA_i^t}{24 n_j} \tag{10}$$

where WA_i^t = water age in junction i at time t. Water age is also often calculated as demand-weighted water age to give more significance to nodes with larger demands. In this research, Eq. (10) is used for WA calculation instead, to be comparable with other methodologies available in the literature.

The above-listed indicators are calculated and used to evaluate solutions based on initially adopted sectorization criteria in this research. However, other PIs can be calculated to address other sets of sectorization criteria (e.g., some type of leakage index).

Selection of Preferred Sectorization Solution

After Stage 3, WDN sectorization is completed, resulting in a set of feasible solutions. This is one of the main advantages of the proposed methodology: It offers an array of alternative DMA designs to the decision maker. One can opt for a solution with a large number of small DMAs or a solution with a small number of large DMAs, or any option in between these two scenarios. This is especially convenient for the analysis of large WDNs without previously established DMAs, where DMA strategic planning should be addressed carefully. It is the role of decision makers to select the optimal sectorization solution for their network based on calculated PIs and other parameters.

Case Study

Description

Methodology presented in this paper has been tested on a large, real-world water distribution network. The analyzed network was originally presented as the second case study network in the Battle of the Water Sensor Networks competition (BWSN2) (Ostfeld et al. 2008). This network has been used as a case study for number of other DMA design algorithms (Diao et al. 2013; Ferrari et al. 2014; Grayman et al. 2009; Hajebi et al. 2016; Zhang et al. 2017). The network consists of 12,523 nodes, 14,822 pipes, 2 reservoirs, 2 tanks, 4 pumps, and 5 valves. Total demand in the network, Q_{tot}, is 1,243 L/s, and total number of connections in the WDN, n_c, is 77,916.

The input data for the DeNSE sectorization method (see "Methodology" section) are carefully set to allow meaningful comparison with previously published methods in the literature where the same network was used. The input data are as follows: (1) network's EPANET input file (Exeter University Centre for Water Systems); (2) minimum number of connections per DMA, $n_c^{\min}$, is 500, and maximum number of connections per DMA, $n_c^{\max}$, is 5,000; (3) transmission main diameter threshold, D_{main}, is 350 mm; (4) pipe closure diameter threshold, D_{tr}, is 300 mm; (5) minimum and maximum operating network pressures are $p_{\min} = 20$ m and $p_{\max} = 75$ m, and maximum allowable water age, $WA_{\max}$, is 48 h; and (6) desired number of sectorization solutions, N_{sol}, is 15.

Based on total demand in the network (Q_{tot}), minimum ($n_c^{\min}$) and maximum ($n_c^{\max}$) number of connections in a DMA, and total number of connections in the network (n_c), minimum and maximum DMA size are calculated using Eq. (2) as $S_{\min} = 8$ L/s and $S_{\max} = 80$ L/s. For hydraulic modeling, 24-h MDD simulation is used, and for water quality modeling (WA calculation), extended-period simulation of 192 h is used.

Network Clustering (Stage 2)

Fig. 7 shows the evolution of the network uniformity index (U) through the network clustering process done in Stage 2, with maximum uniformity index value corresponding to 43 clusters ($U_{\max} = 0.5112$). The minimum number of clusters is 23, in accordance with Ferrari et al. (2014), in which the same transmission main diameter (350 mm) was used and 23 independent districts connected to the main were identified. Fig. 8 shows the evolution of all three components constituted in the network uniformity index $(U) - u_{net}$, u_v, and w_{agg}, in the last 77 aggregation steps (in total, there are 11,708 steps and all three components start from zero). Results illustrate that until maximum uniformity index value is reached, u_{net} is the main parameter driving the clustering process. After that point, large clusters are created, which impacts both u_{net} and u_v, causing them to decrease (seemingly at comparable rates). As the plot suggests, w_{agg} continuously increases as aggregation proceeds, and changes only slightly in the final 77 steps given that most of the links are already aggregated.

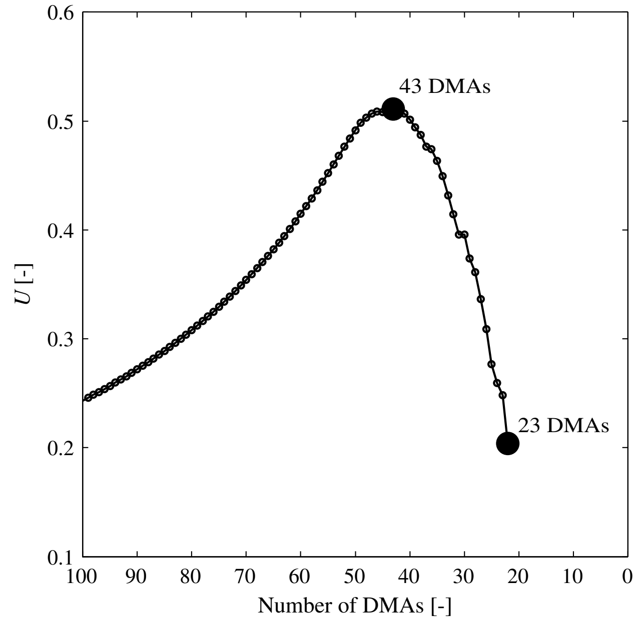

Fig. 7. Evolution of uniformity index during clustering of BWSN2 network.

DMAs Definition and Evaluation (Stage 3)

After Stage 2, 15 solutions are selected for further analysis having between 43 and 29 clusters. In Stage 3, flow meters and isolation valves are positioned to create DMAs and each solution is hydraulically analyzed. The first solution (Sol-1), with 43 DMAs, does not satisfy the pressure constraints and is excluded from further analysis as infeasible.

In addition to adopted PIs used to evaluate the solutions, the following additional indicators are calculated to aid in evaluating solutions using the methods proposed here, and to enable comparison with other methods described in the literature (see "Comparison of Results with Other Methods", below):

- Number of DMAs (N_{DMA}), number of meters (N_M) and number of valves (N_V),
- NL—Number of DMAs larger than maximum DMA size ($S_{\max}$),
- NS—Number of DMAs smaller than minimum DMA size ($S_{\min}$), and
- A_{conn}—Average number of connections per DMA.

Cost, adopted PIs (Res and WA), and above-listed additional indicators for the remaining 14 feasible solutions are shown in Table 1.

As shown in Table 1, all solutions have relatively similar values of two PIs, $WA$and Res. As the number of DMAs in the solution decreases, average the number of connections per DMA increases, meaning that DMAs are larger in size. Consequently, creation of a smaller number of larger DMAs requires fewer flow meters and isolation valves, resulting in lower-cost solutions. The second solution (Sol-2) has one DMA that is smaller than minimum size $S_{\min}$. In solutions Sol-3–Sol-9, all DMAs are within specified $S_{\min}$–$S_{\max}$ range, whereas solutions Sol-10–Sol-15 include one or two DMAs that are larger than $S_{\max}$.

Selection of Preferred Sectorization Solution

The preferred solution is identified by analyzing Sol-3–Sol-9, the solutions that fully satisfy the DMA size constraints. As noted earlier, all feasible solutions have similar impacts on the network's

J. Water Resour. Plann. Manage., 2020, 146(3): 04019079

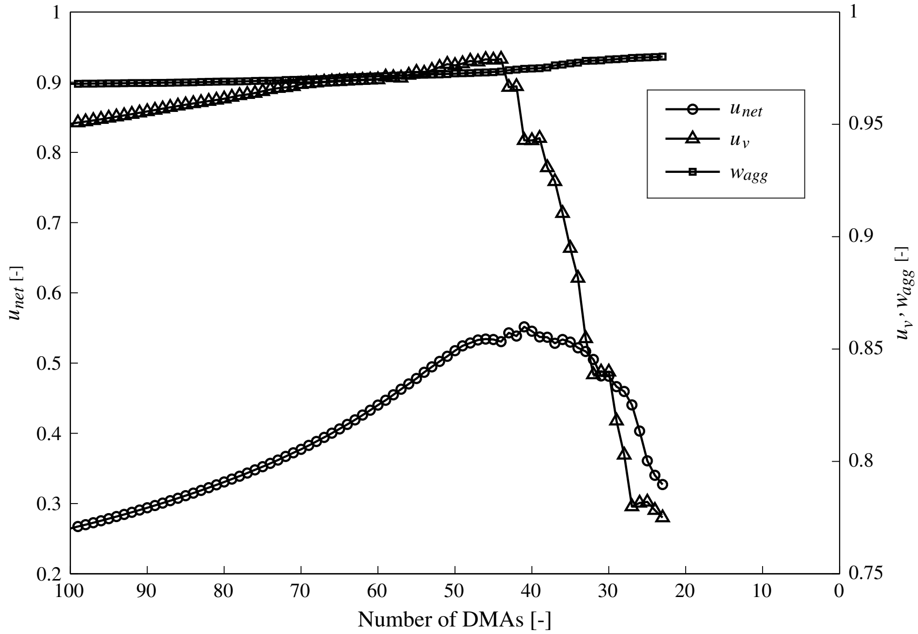

Fig. 8. Evolution of all three components (u_{net}, u_v, and w_{agg}) constituted in the network uniformity index (U) in the last 77 aggregation steps.

Table 1. Evaluation indicators for 14 feasible solutions

Solution No.	N_{DMAs}	NL	NS	A_{conn}	WA (h)	Res	$Cost$ (€)	N_M	N_V	u_{net}	u_v	U
Sol-2	42	0	1	1,655	34.13	0.881	557,405	81	178	0.538	0.967	0.5073
Sol-3	41	0	0	1,696	34.11	0.881	551,215	80	177	0.552	0.943	0.5070
Sol-4	40	0	0	1,738	34.11	0.881	545,870	79	177	0.545	0.943	0.5013
Sol-5	39	0	0	1,783	33.98	0.882	542,210	79	176	0.537	0.944	0.4943
Sol-6	38	0	0	1,830	34.02	0.880	537,920	77	176	0.537	0.931	0.4872
Sol-7	37	0	0	1,879	34.02	0.880	534,500	76	175	0.528	0.925	0.4767
Sol-8	36	0	0	1,931	34.01	0.880	530,995	76	169	0.534	0.910	0.4744
Sol-9	35	0	0	1,987	34.00	0.880	523,685	75	166	0.530	0.895	0.4633
Sol-10	34	1	0	2,045	34.00	0.881	522,565	75	164	0.522	0.882	0.4496
Sol-11	33	1	0	2,107	34.01	0.881	516,375	74	163	0.516	0.855	0.4318
Sol-12	32	2	0	2,173	33.98	0.881	515,815	74	162	0.505	0.839	0.4145
Sol-13	31	2	0	2,243	33.98	0.881	510,470	73	162	0.482	0.840	0.3957
Sol-14	30	2	0	2,318	33.96	0.880	497,205	71	153	0.481	0.840	0.3956
Sol-15	29	2	0	2,398	33.88	0.885	490,470	71	138	0.466	0.818	0.3736

resilience ($Res = 0.880$–0.885) and water age ($WA = 33.88$–34.13 h). Therefore, Sol-9 is preferred over Sol-5 because it is the least costly of the two solutions.

Fig. 9 shows the preferred solution, Sol-9, where the analyzed WDN is sectorized into 35 DMAs, together with the detail of DMA 23 with the positions of valves and flow meters. These positions are identified using the heuristic approach described in Stage 3-Phase 1. Originally, the cluster to which this DMA belongs had six boundary pipes. Three of them were identified as links that always return water to the transmission main (V2, V3, and V4), and as such are marked for closure. The other three boundary pipes are "always input to the zone" pipes, and using the described methodology, pipe V1 ($D = 203.2$ mm) is selected for closure, while the other two pipes with larger diameters ($D = 304.8$ mm) are left open and equipped with flow meters (M1 and M2).

To provide further insight into the selected solution and the effects of network interventions required to create DMAs, in addition to PIs and other indicators characterizing the solution listed above (Table 1), for each DMA in a solution the following PIs are calculated:

1. p_{DMA}^{av} = Mean average pressures over 24-h in a DMA, as a good indicator of network interventions' impacts on pressure distribution, calculated as

$$p_{DMA}^{av} = \frac{\sum_{i=1}^{n_j} \sum_{t=1}^{24} p_i^t}{24\, n_j} \quad \forall\ i \in DMA \tag{11}$$

2. Res_{DMA} = Average resilience index for a DMA, calculated per Eq. (9), but this time accounting for nodes within the considered DMA.
3. WA_{DMA} = Demand-weighted WA for a DMA, averaged over the entire extended period simulation (192 h). Demand weighting is used to account for difference in size between DMAs in terms of demand.

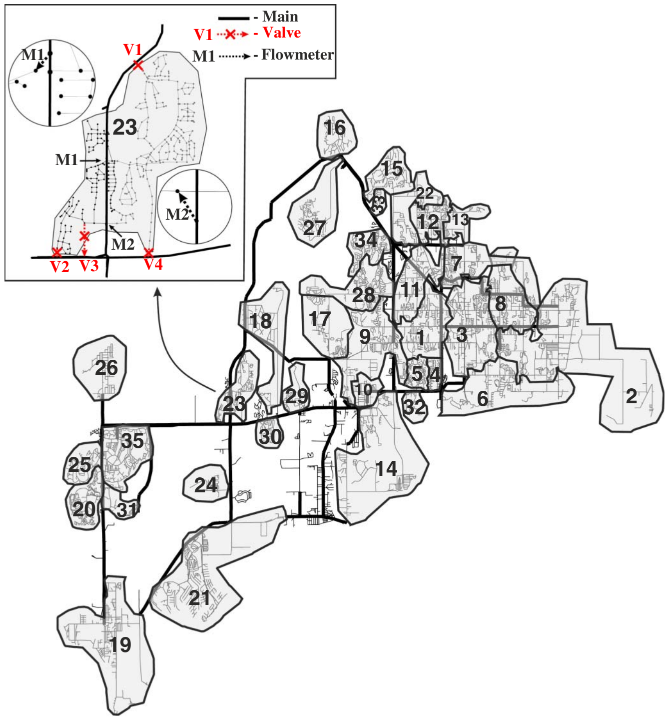

Fig. 9. Preferred sectorization solution, Sol-9, with 35 DMAs and detail of DMA 23.

$$WA_{DMA} = \frac{\sum_{i=1}^{n_j} \sum_{t=1}^{T} WA_i^t q_i^t}{\sum_{i=1}^{n_j} \sum_{t=1}^{T} q_i^t} \quad \forall\ i \in DMA \qquad (12)$$

Figs. 10 and 11 show results for each of 35 created DMAs in selected solution Sol-9. Fig. 10(a) shows average consumption in DMAs, with highlighted minimum and maximum size constraints. As shown in the graph, the 35 DMAs vary in size considerably but always fall within the design limits imposed. Fig. 10(b) shows relative changes in mean average pressure in DMAs compared with mean average pressures in nodes that are part of that DMA in the original nonsectorized network ($\Delta p_{\text{DMA}}^{av}$). For most DMAs, the mean average pressure decreased slightly (up to 4%), whereas a slight increase occurred in six DMAs (up to 1%). Therefore, network sectorization had very limited impact on redistribution of pressure within the WDN. A significant decrease of pressure was observed in DMA 8 (by 13%), but all pressures were still within the required range of $p_{\min}$–$p_{\max}$.

Fig. 11(a) illustrates relative changes in water age in the DMAs, again compared to the original network layout (ΔWA_{DMA}). Maximum decrease of WA is 20%, and increase is almost 30%. Although a decrease in WA is desirable, an increase of 30% may seem a bit high at first. However, plotting absolute values of WA for DMAs in which increase is induced by network interventions (Fig. 12), it is easy to conclude that WA is still well below the set maximum, $WA_{\max}$, of 48 h. Fig. 11(b) shows relative changes in the DMAs' resilience index (ΔRes_{DMA}). Changes in resilience

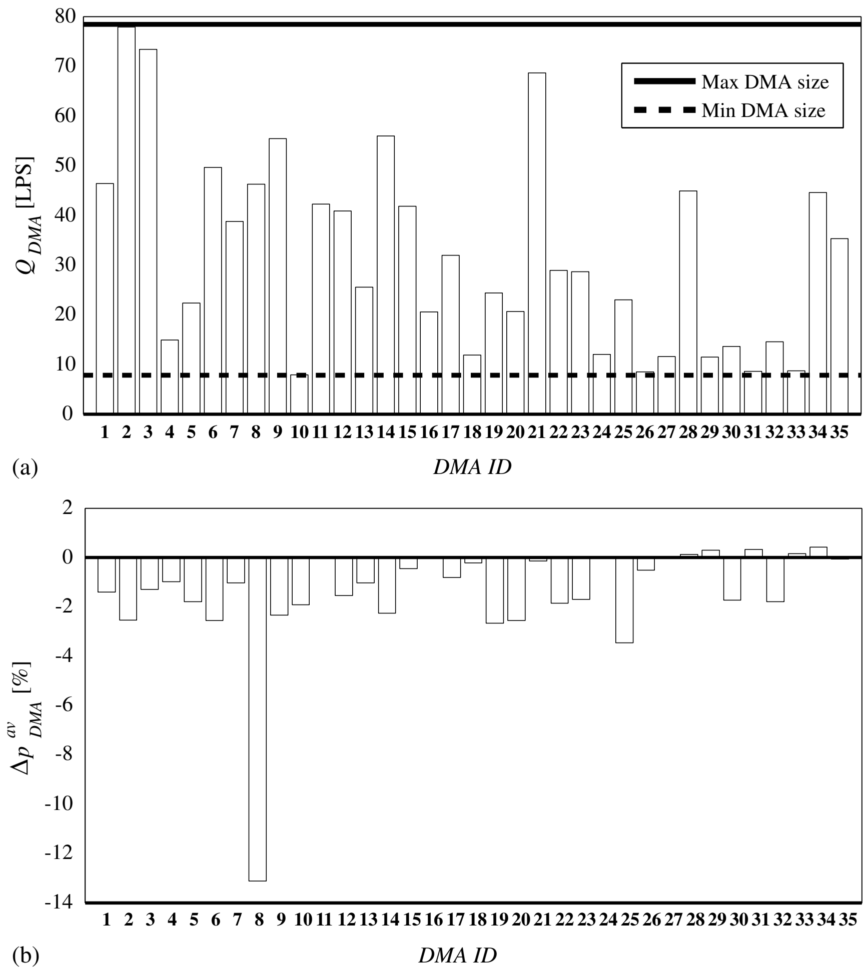

Fig. 10. Results for each DMA in preferred solution, Sol-9: (a) average DMA consumption; and (b) relative change of mean average pressure.

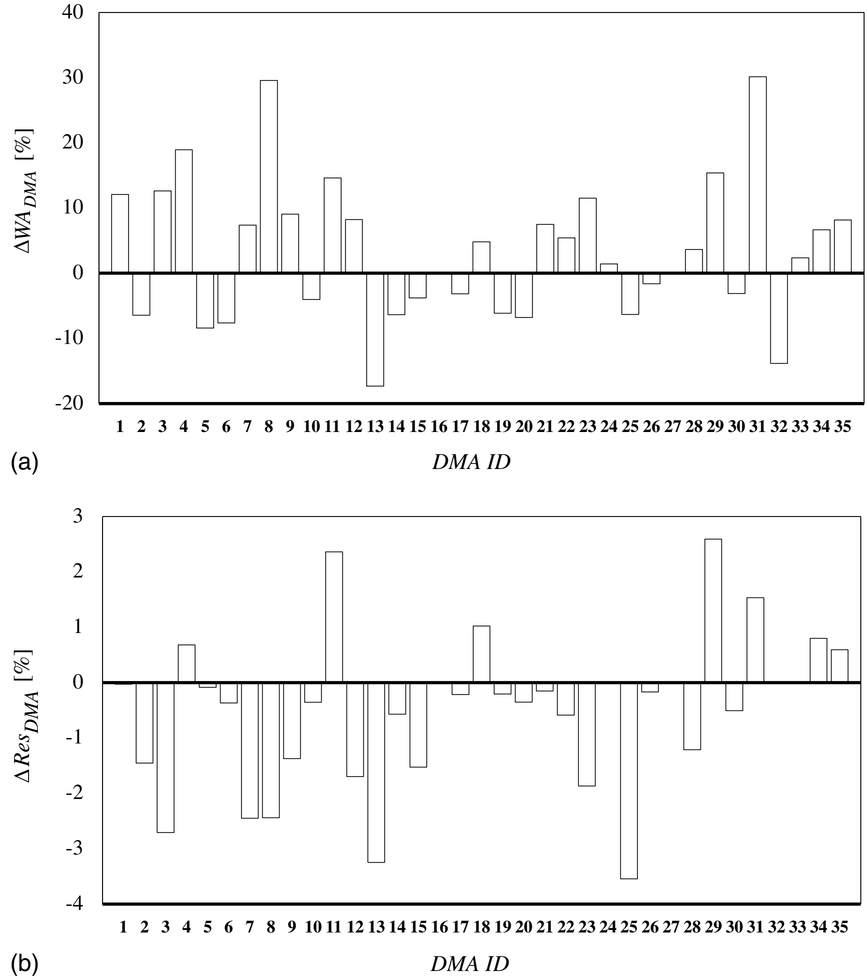

Fig. 11. Results for each DMA in selected preferred solution, Sol-9: (a) relative change of water age; and (b) relative change of resilience index.

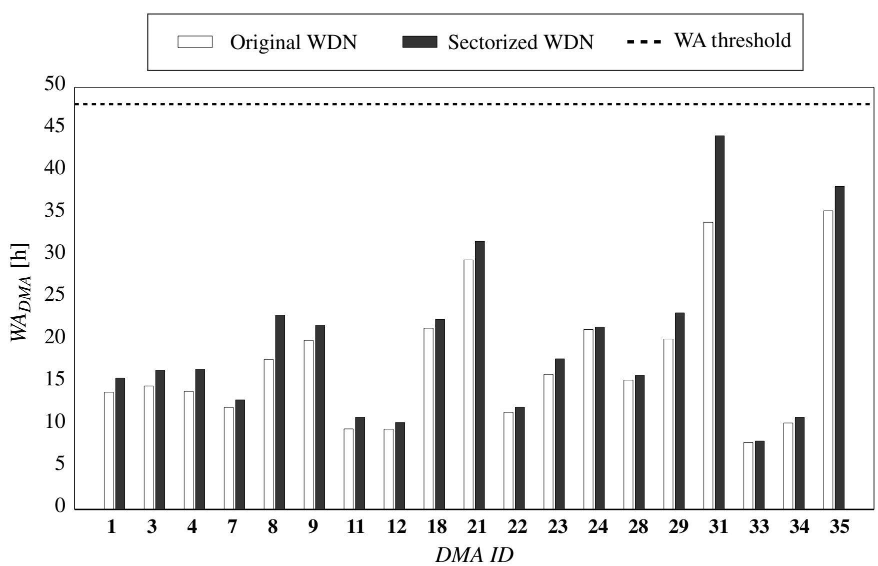

Fig. 12. Values of water age, before and after sectorization, for DMAs with increased water age.

index range from −3.5% to +2.2%, indicating very limited impact of sectorization on the resilience of the WDN.

To summarize, from the results discussed above, it can be concluded that: (1) all DMAs are within required size limits in terms of consumption, (2) the network's hydraulic performance is not endangered given that changes in zone pressures are negligible, (3) water quality requirement, expressed through the *WA*, is satisfied, given that for all DMAs, the *WA* is still below maximum allowed threshold of 48 h, and (4) network reliability is sustained, given that the changes in the resilience index are almost insignificant.

Comparison of Results with Other Methods

Finally, results obtained here are compared against corresponding results obtained from five previously published approaches that addressed the WDN sectorization problem and by using the same case study (Table 2). Comparison is made in terms of number of DMAs (N_{DMAs}), DMAs that are larger (NL) and smaller (NS) than predefined size constraints, number of flow meters (N_M) and isolation valves (N_V), added pipes (P_{add}), average number of connections per DMA (A_{conn}), computational time (*Comp.Time*), and PIs adopted in this research to evaluate the solutions: water age and resilience index (*WA* and *Res*). Direct comparison with other methods in terms of cost could not be made because cost was not explicitly reported in other papers. Reported values of PIs in Table 2 refer to best sectorization solutions reported by each research study. Computational times are given only as a qualitative metric, to illustrate differences in magnitudes between different methods. Table 2 also provides an overview of sectorization methods used in each method for (1) partitioning the WDN and (2) positioning the flow meters and isolation valves.

As shown in Table 2, only the methodologies presented in Hajebi et al. (2016) and in the DeNSE method produce a set of feasible solutions. Hajebi et al. (2016) identified 78 feasible solutions having between 28 and 48 DMAs. Regarding the DMA size constraints, solutions presented by Grayman et al. (2009) and Diao et al. (2013) have DMAs that are both larger and smaller, and in the solution presented by Ferrari et al. (2014), all DMAs fulfil size constraints. In Hajebi et al. (2016), all 78 feasible solutions meet size constraints, while in the methodology presented here, this is case for 7 out of 14 feasible solutions.

Methodologies using multiobjective (MO) optimization to position flow meters and isolation valves (Hajebi et al. 2016; Zhang et al. 2017) require significant amount of computational time (15 and 278 h, respectively). The substantially lower computational time of Hajebi's method compared with the method of Zhang can be attributed to the use of a shorter extended-period simulation time (48 h vs. 192 h) used to calculate WA. The issue of high computational time as a consequence of using MO optimization is addressed in Diao et al. (2013), in which a two-stage heuristic procedure for device placement is applied, resulting in an acceptable running time of approximately 20 min. However, only one solution with 41 DMAs is reported, with three DMAs falling outside of the required size limits. The engineering-based heuristic procedure used in this work takes a similar amount of time (approximately 20 min), but produces a set of feasible solutions compared with the work of Diao et al. (2013). The computational efficiency of the DeNSE method is even more evident when compared with the work of Hajebi et al. (2016). Both methods are able to produce a set of feasible solutions, but DeNSE takes 20 min rather than 15 h, yet it uses a longer extended-period analysis for WA calculation (192 h compared with 48 h).

Methodologies of Ferrari et al. (2014) and Hajebi et al. (2016) ensure connectedness of each DMA to the transmission main (direct access to the water source) and their isolation from other DMAs called isolated DMAs (iDMAs). While methodology presented here does not create iDMAs, the preferred solution presented earlier (Sol-9) fulfils the condition of direct access to the water source. All 35 DMAs are directly connected to the transmission main: 20 DMAs with 1 pipe; 4 with 2; 6 with 3; 4 with 4; and 1 with 6 pipes.

Table 2. Comparison of results with other methods

Reference	Method for											
	WDN partitioning	Device placement	N_{DMAs}	NL	NS	N_M	N_V	P_{add}	A_{conn}	Computational time (min/h)	WA (h)	Res
Grayman et al. (2009)	Manual		43	1	3	53	163	11	1,996	N/A	31.51	0.802
Diao et al. (2013)	Community detection	2-stage heuristic method	41	2	1	N/A	N/A	0	2,044	20 min	32.01	N/A
Ferrari et al. (2014)	Graph-based recursive bisection algorithm		36	0	0	181	152	0	2,317	N/A	N/A	N/A
Hajebi et al. (2016)	Heuristic graph partitioning	MO optimization	28–48	0	0	56–78	66–161	0	1,415–2,423	15 h	31.01	0.830
Zhang et al. (2017)	Community detection	MO optimization	43	N/A	N/A	103	33	0	N/A	278 h	N/A	N/A
DeNSE algorithm	Uniformity-based clustering	Engineering-based heuristic	29–42	0–2	0–1	71–81	138–178	0	1,656–2,398	20 min	34.00	0.880

Note: N/A = not available; WA = water age for best reported solution; and Res = resilience index for best reported solution.

Table 2 also shows the comparison of DeNSE against main PI values for best reported solutions, obtained with different methods: water age (*WA*) and resilience index (*Res*). Presented results show that the DeNSE method achieves a slightly better resilience index value and slightly worse water age value. Reported results are only indicative as different input parameters, affecting values of compared indicators, are used. For water age calculation Grayman et al. (2009), Diao et al. (2013), and the methodology presented here use 192-h extended-period simulation, while Hajebi et al. (2016) uses 48-h simulation. Furthermore, water age value is highly dependent on the adopted time step for water quality simulation and those papers do not supply this information. Grayman et al. (2009) reported increase of 2.61% in *WA* for the DMA system, when compared to the original network (from 30.71 to 31.51 h). In the case of DeNSE method *WA* is increased by 3.31% for the DMA system (from 32.91 to 34 h) which is regarded as insignificant increase and same order of magnitude as achieved in Grayman et al. (2009). Reported resilience indices are influenced by the adopted minimum allowable pressure in the network and time period over which they are averaged. Grayman et al. (2009) adopted minimum pressure of 30 psi (20 m) and 51-h time period. Hajebi et al. (2016) uses 28 m minimum pressure and 48-h time period, while Diao et al. (2013) does not report values of this indicator. Grayman et al. (2009) reported decrease of resilience index of 4.07% for the DMA system, when compared to the original network (from 0.836 to 0.802), while the DeNSE method achieves lower decrease of 2.55% (from 0.903 for the original network to the 0.88 for the DMA system). As noted above, due to the different input parameters, values of PIs presented in Table 2 are not directly comparable, but illustrative and show that in terms of water age and resilience all methods perform similarly.

Conclusions

The new DeNSE sectorization method is introduced in this paper. It was tested and validated on a large, real-world-sized water distribution network BWSN2 (Ostfeld et al. 2008). The results obtained were compared against other sectorization methods that used the same case study network. The following conclusions are drawn:

1. The DeNSE method is able to identify a set of feasible network sectorization solutions for a large water distribution network such as the one used in the case study described herein. DeNSE can find solutions in a computationally efficient manner which, in turn, enables exploring alternative sectorization strategies by changing the method input parameters. High computational efficiency comes mainly from the new heuristic methodology for positioning the flow meters and isolation valves. The advantage of this approach is evident especially when the DeNSE algorithm is compared with other, optimization-based sectorization methods (Hajebi et al. 2016; Zhang et al. 2017).
2. The DeNSE method ensures that sectorization interventions are identified in a way that does not worsen the operational performance of the WDN prior to its sectorization. The method ensures that minimum and maximum network pressures before and after sectorization stay within the same range. The method also ensures that water quality (measured by water age) is not worsened by WDN sectorization.
3. The DeNSE method estimates explicitly the costs involved in WDN sectorization as opposed to other methods where costs are assessed indirectly, for example, by the number of installed new devices or summarized diameters (e.g., Hajebi et al. 2016). Although the proposed method does not make use of optimization, this explicit assessment of costs enables the identification of realistic sectorization solutions that can be compared with the available budgets.
4. The DeNSE method seems particularly well suited for application at the initial stages of the DMA design process and in inefficient WDNs (i.e., WDNs with higher water losses). This is because the method enables: (1) alternative DMA sizes (both small and large) to be considered and analyzed and (2) preservation of network hydraulic performance and reliability which, in turn, enables tracking the network water balance more easily. Other methods seem to focus more on controling the pressures in the network (Zhang et al. 2017).

Future DeNSE development will address adding sectorization criteria such as design for fire flows, specific water quality parameters (e.g., chlorine), design for security, and others.

Data Availability Statement

Some or all data, models, or code generated or used during the study are available in a repository or online in accordance with funder data retention policies (EPANET input file for case study network is available online at http://emps.exeter.ac.uk/engineering/research/cws/downloads/benchmarks/). Some or all data, models, or code generated or used during the study are available from the corresponding author by request (source code developed for implementation of DeNSE method for sectorization).

References

Alvisi, S. 2015. “A new procedure for optimal design of district metered areas based on the multilevel balancing and refinement algorithm.” *Water Resour. Manage.* 29 (12): 4397–4409. https://doi.org/10.1007/s11269-015-1066-z.

Alvisi, S., and M. Franchini. 2014a. “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.

Alvisi, S., and M. Franchini. 2014b. “Water distribution systems: Using linearized hydraulic equations within the framework of ranking-based optimization algorithms to improve their computational efficiency.” *Environ. Modell. Software* 57 (Jul): 33–39. https://doi.org/10.1016/j.envsoft.2014.03.012.

Babić, B., M. Stanić, D. Prodanović, B. Džodanović, and A. Dukić. 2014. “Reducing uncertainty of infrastructure leakage index—A case study.” *Procedia Eng.* 89: 1577–1584. https://doi.org/10.1016/j.proeng.2014.11.459.

Brentan, B. M., E. Campbell, G. L. Meirelles, E. Luvizotto, and J. Izquierdo. 2017. “Social network community detection for DMA creation: Criteria analysis through multilevel optimization.” *Math. Prob. Eng.* 2017: 1–12. https://doi.org/10.1155/2017/9053238.

Burrows, R., G. Crowder, and J. Zhang. 2000. “Utilisation of network modelling in the operational management of water distribution systems.” *Urban Water* 2 (2): 83–95. https://doi.org/10.1016/S1462-0758(00)00046-7.

Butler, D. 2000. *Leakage detection and management*. Cwambran, UK: Palmer Environmental Ltd.

Campbell, E., J. Izquierdo, I. Montalvo, and R. Perez-Garcia. 2016. “A novel water supply network sectorization methodology based on a complete economic analysis, including uncertainties.” *Water* 8 (5): 179. https://doi.org/10.3390/w8050179.

Chianese, S., A. D. Nardo, M. D. Natale, C. Giudicianni, D. Musmarra, and D. Ingegneria. 2017. “DMA optimal layout for protection of water distribution networks from malicious attack.” In Vol. 1 of *Proc., CRITIS 2017: Int. Conf. on Critical Information Infrastructures Security*, 84–96. New York: Springer.

Ciaponi, C., E. Murari, and S. Todeschini. 2016. “Modularity-based procedure for partitioning water distribution systems into independent

J. Water Resour. Plann. Manage., 2020, 146(3): 04019079

districts." *Water Resour. Manage.* 30 (6): 2021–2036. https://doi.org/10.1007/s11269-016-1266-1.

Clauset, A., M. Newman, and C. Moore. 2004. "Finding community structure in very large networks." *Phys. Rev. E: Stat. Nonlinear Soft Matter Phys.* 70 (6): 1–6. https://doi.org/10.1103/PhysRevE.70.066111.

De Paola, F., N. Fontana, E. Galdiero, M. Giugni, D. Savic, and G. Sorgenti Degli Uberti. 2014. "Automatic multi-objective sectorization of a water distribution network." *Procedia Eng.* 89: 1200–1207. https://doi.org/10.1016/j.proeng.2014.11.250.

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

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

Di Nardo, A., and M. Di Natale. 2011. "A heuristic design support methodology based on graph theory for district metering of water supply networks." *Eng. Optim.* 43 (2): 193–211. https://doi.org/10.1080/03052151003789858.

Di Nardo, A., M. Di Natale, C. Giudicianni, G. F. Santonastaso, V. Tzatchkov, and M. Rodriguez. 2017. "Economic and energy criteria for district meter areas design of water distribution networks." *Water* 9 (7): 463. https://doi.org/10.3390/w9070463.

Di Nardo, A., M. Di Natale, G. F. Santonastaso, and S. Venticinque. 2013. "An automated tool for smart water network partitioning." *Water Resour. Manage.* 27 (13): 4493–4508. https://doi.org/10.1007/s11269-013-0421-1.

Di Nardo, A., C. Giudicianni, R. Greco, and G. F. Santonastaso. 2018. "Applications of graph spectral techniques to water distribution network management." *Water* 10 (1): 1–16. https://doi.org/10.3390/w10010045.

Farley, M. 2001. "Leakage management and control: A best practice training manual." Accessed November 21, 2019. http://whqlibdoc.who.int/hq/2001/WHO_SDE_WSH_01.1_pp1-98.pdf.

Ferrari, G., D. 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): 1–13. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000424.

Fortunato, S. 2010. "Community detection in graphs." *Phys. Rep.* 486 (3–5): 75–174. https://doi.org/10.1016/j.physrep.2009.11.002.

Gabow, H. N. 2000. "Path-based depth-rst search for strong and biconnected components 1 introduction 2 strong components." *Inf. Process. Lett.* 74 (3–4): 107–114. https://doi.org/10.1016/S0020-0190(00)00051-X.

Giudicianni, C., A. D. Nardo, M. D. Natale, R. Greco, G. F. Santonastaso, and A. S. Id. 2018. "Topological taxonomy of water distribution networks." *Water* 10 (4): 1–19. https://doi.org/10.3390/w10040444.

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

Gomes, R., S. S. Marques, and S. Joaquim. 2012. "Decision support system to divide a large network into suitable district metered areas." *Water Sci. Technol.* 65 (9): 1667–1675. https://doi.org/10.2166/wst.2012.061.

Grayman, W., R. Murray, and D. Savic. 2009. "Effects of redesign of water systems for security and water quality factors." In *Proc., World Environmental and Water Resources Congress 2009*, 504–514. Reston, VA: ASCE.

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

Herrera, M., S. Canu, A. Karatzoglou, and J. Izquierdo. 2010a. "An approach to water supply clusters by semi-supervised learning." In *Proc., Int. Environmental Modelling and Software Society (iEMSs) 2010 Int. Congress on Environmental Modelling and Software*. Reston, VA: United States Geological Survey.

Herrera, M., J. Izquierdo, R. Pérez-García, and D. Ayala-Cabrera. 2010b. "Water supply clusters by multi-agent based approach." In *Proc., Water Distribution System Analysis 2010–WDSA 2010*, 861–869. Reston, VA: ASCE.

Ivetić, D., Ž. Vasilić, M. Stanić, and D. Prodanović. 2013. "Optimizacija mreža pod pritiskom modeliranih ΔQ metodom." *Vodoprivreda* 264–266 (4–6): 265–274.

Jungnickel, D. 2005. *Graphs, networks and algorithms*. 2nd ed. Edited by M. Bronstein, A. Cohen, H. Cohen, D. Eisenbud, and B. Sturmfels. Berlin: Springer.

Laucelli, D. B., A. Simone, L. Berardi, and O. Giustolisi. 2016. "Optimal design of district metering areas." *Procedia Eng.* 162 (2014): 403–410. https://doi.org/10.1016/j.proeng.2016.11.081.

Morrison, J., S. Tooms, and D. Rogers. 2007. *DMA management guidance notes*. London: International Water Association.

Newman, M. E. J., and M. Girvan. 2004. "Finding and evaluating community structure in networks." *Phys. Rev. E: Stat. Nonlinear Soft Matter Phys.* 69 (2): 1–15. https://doi.org/10.1103/PhysRevE.69.026113.

Ostfeld, A., et al. 2008. "The battle of the water sensor networks (BWSN): A design challenge for engineers and algorithms." *J. Water Resour. Plann. Manage.* 134 (6): 556–568. https://doi.org/10.1061/(ASCE)0733-9496(2008)134:6(556).

Perelman, L., M. Allen, A. Preis, M. Iqbal, and A. J. Whittle. 2015. "Automated sub-zoning of water distribution systems." *Environ. Modell. Software* 65 (Mar): 1–14. https://doi.org/10.1016/j.envsoft.2014.11.025.

Perelman, L., and A. Ostfeld. 2012. "Water-distribution systems simplifications through clustering." *J. Water Resour. Plann. Manage.* 138 (3): 218–229. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000173.

Scarpa, F., A. Lobba, and G. Becciu. 2016. "Elementary DMA design of looped water distribution networks with multiple sources." *J. Water Resour. Plann. Manage.* 142 (6): 04016011. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000639.

Sedgewick, R., and K. Wayne. 2011. *Algorithms*. 4th ed. Boston: Addison-Wesley.

Tarjan, R. 1972. "Depth-first search and linear graph algorithms." *SIAM J. Comput.* 1 (2): 146–160. https://doi.org/10.1137/0201010.

Todini, E. 2000. "Looped water distribution networks design using a resilience index based heuristic approach." *Urban Water* 2 (2): 115–122. https://doi.org/10.1016/S1462-0758(00)00049-2.

Tzatchkov, V., V. Alcocer-Yamanaka, and V. Bourguett Ortíz. 2006. "Graph theory based algorithms for water distribution network sectorization projects." In *Proc., Water Distribution Systems Analysis Symp. 2006*, 1–15. Reston, VA: ASCE.

UK Water Industry Research. 1999. *A manual of DMA practice*. London: UK Water Industry Research Limited.

Vasilic, Z. 2018. "Decision support algorithms for sectorization of water distribution networks." Ph.D. thesis, Faculty of Civil Engineering, Dept. for Hydraulic and Environmental Engineering, Univ. of Belgrade.

Vasilić, Ž., M. Stanić, D. Prodanović, and Z. Kapelan. 2016. "Network sectorisation through aggregation of strong connected components." In *Proc., 18th Conf. on Water Distribution System Analysis, WDSA 2016*. Amsterdam, Netherlands: Elsevier.

WAA and WRC (Water Authorities Association and Water Research Centre). 1985. *Leakage control policy & practice*. Rep. No. 26. London: Water Authorities Association.

Zhang, K., H. Yan, H. Zeng, K. Xin, and T. Tao. 2019. "A practical multi-objective optimization sectorization method for water distribution network." *Sci. Total Environ.* 656 (Mar): 1401–1412. https://doi.org/10.1016/j.scitotenv.2018.11.273.

Zhang, Q., Z. Wu, M. Zhao, and J. Qi. 2017. "Automatic partitioning of water distribution networks using multiscale community detection and multiobjective optimization." *J. Water Resour. Plann. Manage.* 143 (9): 1–14. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000819.