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
Accepted author manuscript

Published in
Journal of Water Resources Planning and Management

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

Zeljko Vasilic[1], Milos Stanic[2], Zoran Kapelan[3,4], Dusan Prodanovic[5] and Branislav Babic[6]

[1] Assistant Professor, University of Belgrade, Faculty of Civ. Eng., Bul. kr. Aleksandra 73, Belgrade, Serbia, (corresponding author) e-mail: zvasilic@grf.bg.ac.rs

[2] Associate Professor, University of Belgrade, Faculty of Civ. Eng., Bul. kr. Aleksandra 73, Belgrade, Serbia, e-mail: mstanic@grf.bg.ac.rs

[3] Professor, Delft University of Technology, Faculty of Civil Engineering and Geosciences, Building 23, Stevinweg 1, 2628 CN Delft, Netherlands

[4] Professor, University of Exeter, College of Engineering, Mathematics and Physical Sciences, Harrison Building, North Park Road, Exeter EX4 4QF, United Kingdom

[5] Professor, University of Belgrade, Faculty of Civ. Eng., Bul. kr. Aleksandra 73, Belgrade, Serbia

[6] Assistant Professor, University of Belgrade, Faculty of Civ. Eng., Bul. kr. Aleksandra 73, Belgrade, Serbia

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 done using a “trial and error” approach conducted by local experts which often results in arbitrary solutions being identified. A number of methods published recently tried to improve WDN sectorization by automating the process, especially by using optimization. Various sectorization criteria, constraints and limitations are introduced, often neglecting limited funds and shortage of water balance data often encountered in poorly managed WDNs. These methods also suffer from low computational efficiency imposed by optimization methods used. This paper presents a new, Distribution Network SEctorization (DeNSE) method that overcomes these deficiencies. The new method is based on a heuristic procedure where the

WDN sectorization is driven by efficient tracking of water balance and least cost investment for implementation while maintaining the same level of WDN's operational performance. Aforementioned set of criteria is particularly well suited for initial sectorization of poorly managed WDNs, in which great uncertainty in water balance data often leads to poor management decisions. DeNSE method is validated and benchmarked against several literature sectorization methodologies on a real-sized WDN. The results obtained demonstrate the ability of the DeNSE to identify set of good, realistic sectorization solutions that are in some respects better than the corresponding solutions reported in the literature. The new method also enables sectorization to be done in a computationally efficient manner ensuring its applicability to large, real-life sized WDNs.

Key words: Sectorization, DMA, WDN, Uniformity, 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 80's and since it's been implemented in many WDNs worldwide. Sectorization has been done traditionally to address two main objectives: better control of water losses and efficient management of pressures in the network. It is proven that sectorization can be useful for other tasks such as protection against contamination (Chianese et al., 2017; Grayman et al., 2009). Best definition of a DMA, given by Burrows et al. (2000), is that it is 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 WDN into an optimal system of DMAs is a hard task to achieve, especially for the existing and continuously operating WDN. Every WDN is unique in its topology and characteristics and key drivers/objectives so there is no common procedure for performing its sectorization, but rather a series of guidelines provided by the different water and other authorities (Butler, 2000; Farley, 2001;

Morrison et al., 2007; WAA & 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 come up with the solution that will be efficient both in terms of sectorization main objectives and satisfaction of network's hydraulic and other requirements.

Complexity of the real life WDN results in many different alternatives in which network sectorization can be done. Usually, sectorization is governed by the criteria of having zones of "manageable size" in terms of number of consumers, links or network length. It can be also subjected to other criteria (e.g. required number of feeds, fire flow regulations etc.) and limitations. Sectorization solutions are usually obtained by the "trial and error" technique conducted by a local expert, familiar with all of the WDN specifics. Practical application of such approach is illustrated in Grayman et al. (2009) where two large case study networks are redesigned to implement typical DMA design as guidelines provided in Baker (2007) and to allow additional control and isolation of the system in order to improve water security. Need for a more formal approach to sectorization problem, that will enable investigation of alternative sectorization solutions for large WDNs, is recognized early (Tzatchkov et al., 2006).

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 this process (Deuerlein, 2008; Perelman & 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 (1^{st} Step), majority of presented methodologies rely on the Graph Theory algorithms (Alvisi & Franchini, 2014a; Di Nardo et al., 2013; Ferrari et al., 2014; Hajebi et al., 2016; Scarpa et al., 2016), or multi-agent approach and spectral clustering (Di Nardo et al., 2018; Herrera et al., 2010; Herrera et al., 2010), while others are using the modularity index (Ciaponi et al., 2016; Giustolisi & 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 & 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, WDN being one of them, have a property of higher links density within the communities than between them (Fortunato, 2009; Giudicianni et al., 2018). These metrics have been tailored in different ways to be used for the WDN sectorization purpose (Giustolisi & Ridolfi, 2014; Zhang et al., 2017). Although able to determine DMAs, these approaches are sensitive to the selection of links weights (Ciaponi et al., 2016; Diao et al., 2013). So 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 will be also uniform in size as much as possible is aspect addressed in research presented here, hypothesizing that uniformity of DMAs' sizes can be suitable variable to govern the sectorization process.

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

Number and type of PIs, used in 3rd Step to assess the effect of implemented interventions and evaluate the sectorization solution, vary significantly in reported researches. Resilience index, as describe in Todini (2000) is present in almost all researches as a measure of networks post-sectorization reliability.

Water age is usually used to reflect the impact on water quality in network. Some researchers added various other indices to validate feasibility of obtained solutions (e.g. pressure indices are used in Di Nardo et al. (2013) and entropy index is used in Scarpa et al. (2016)).

Some of the drawbacks of available methods for automated sectorization, potentially posing a question of their applicability to real-life WDNs, are associated with: a) comprehensive lists of objectives and constraints used in optimization, b) computational efficiency and c) resolution of sectorization solution.

In the process of developing new methods, various limitations and constraints, important for the proper functioning of the WDN, were 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 were considered (Di Nardo & Di Natale, 2011), with each new method adding new sectorization parameters and network's 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. It may be even commented that these lists have grown too much, exhausting all practical aspects important for normal every day 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 (e.g. algorithms of (Hajebi et al., 2016; 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 are lacking 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 number of DMAs at once, so sectorization process should be planned hierarchically and implemented in phases starting with a few DMAs that can be larger than recommendations given in guidelines. Establishing a few DMAs in WDN should enable tracking of water balance in the network and gathering basic data about system dynamics, without significant effect on network's operational conditions. This could improve operational management of WDN, as management decisions are usually made based on some calculated WDN’s PIs, whose values can be 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 manner economical aspect is addressed as this implies minimization of costs. 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 of sectorization would be closer to engineering perception and more in accordance with the phased creation of DMAs. Either way, hierarchy in sectorization solution should be considered.

From the previous discussion it can be concluded that, despite all recent advancements made, scope exist to further improve existing water network sectorization algorithms. Aspects in which these improvements can be made are: 1) implementation of practical engineering principals, relevant to the WDN, to govern the sectorization process, 2) improving computational efficiency of the algorithm and 3) consider hierarchical sectorization.

In the method presented here, named DeNSE (Distribution Network SEctorization), first aspect 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. Uniformity index also favours sectorization in which cluster’s connecting links are pipes with smaller diameters, indirectly providing economically more favourable solution as installation of valves and flow meters on smaller diameter pipes will be less costly. 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. Network clustering algorithm is evolving in a step by step manner, hence obtained sectorization solution is inherently hierarchically ordered. Furthermore, algorithm presented here does not come up with a single sectorization solution, but with a range of feasible solutions, giving the freedom to the decision makers to select the one best suited for their needs. Algorithm is tested on large real-sized BWSN2 benchmark network (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 method for water Distribution Network SEctorization into DMAs (DeNSE), also able to address hierarchical sectorization. Algorithm is based on the Graph Theory for identification of Strong Connected Components (SCCs) and their aggregation into clusters based on newly presented network uniformity index (*U*). As discussed in Introduction, 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 impact of interventions made in the network. Tracking the water balance in the network is main sectorization objective adopted in DeNSE method. Designing the sectorization solution that requires least investment in the equipment necessary for creation of DMAs (flow meters and isolation valves), while keeping the same level of network's operational efficiency are main design criteria. Such set of design criteria is most appealing to many water utilities, especially in the developing countries, which operate highly inefficient WDNs with significant amount of water and revenue losses. Two PIs are adopted to evaluate the effects of the sectorization on network's operational performance: 1) Resilience Index (*Res*), reflecting post-sectorization reliability of WDN (Todini, 2000) and 2) Water Age (*WA*), surrogate metrics for water quality reflecting water retention rate in the WDN.

The new method requires hydraulic model of WDN as an input, like many other methods relying on it to prove hydraulic feasibility of sectroization solution. The quality of the adopted solution will be better if calibrated hydraulic model is used, and required interventions in the network can be taken with more assurance in preservation of networks hydraulic performance. The method runs through 3 stages to identify the best sectorization solution, as shown in Figure 1. First stage is a pre-processing stage in which all the relevant network data is obtained from the WDN model and prepared for the follow run of the clustering algorithm. WDN decomposition into clusters is done in the second stage, based on the uniformity index. Third stage involves selection of the narrow set of solutions that will be hydraulically analysed. Heuristic, engineering based positioning of the valves and flowmeters on clusters connecting links in order to define DMAs, extended period hydraulic analysis of the solutions and evaluation of adopted PIs, are all part of the third stage. Finally, feasible solutions are ranked and preferable solution is selected. Each of the three stages will be explained in details in the following text.

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 characteristic, 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 mathematical model. Recommendations about these values can be found in number of available guidelines for DMA creation, and usually it is considered that 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 as 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, value of D_{main} is network specific, usually being 300-350 mm (Ferrari et al., 2014).

4. Pipe closure threshold diameter (D_{tr}). Pipes having diameter equal or larger than this diameter will not be considered for possible closure for positioning the valves and flowmeters (part of the 3rd stage). By default, algorithm uses first class of diameter lower than the D_{tr} (e.g. if D_{main} is 350 mm, D_{tr} will be 300 mm), but user can specify a different value. However, this will affect the number of isolation valves and flowmeters 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 to 15 solutions is large enough set to make representative multi-criteria ranking, however user can opt for a larger set of solutions to compare.

Pre-processing (STAGE 1)

In the first stage, there are two phases (Figure 1). In the first phase, transmission mains are defined, based on the D_{main} value, and excluded from the sectorization process. For this purpose, network is explored using slightly modified Breadth First Search (BFS) algorithm (Jungnickel, 2005), simultaneously starting from all main source nodes (reservoirs). BFS algorithm is modified to prioritize propagation through the links with diameters equal or greater than D_{main}. In the second phase, 24-hour Maximum Day Demand (MDD) hydraulic simulation of the analysed 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 $G = <N, C>$, set of network nodes N and set of network links C, where each link is presented with ordered pair of nodes. Network links with changing flow directions are identified as non-oriented (or links that can have both flow directions), and are represented with the addition of 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 of these phases are illustrated on a simple example network shown in Figure 2.

The example network consists of 16 nodes, two of which are reservoirs, and 21 links. Links connecting reservoirs are identified as transmission mains and are excluded from further analysis. 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 those links are identified as not oriented, and putting that in the context of water networks, those are usually pipes (links) that are connecting 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 (see Figure 1).

Phase 1. First step is to identify the Strongly connected components within the previously created DIGRAPH. Strongly connected component (SCC) is a term from Graph Theory, and it is defined as a subgraph in which each node can be reached from any other node within that subgraph (Gabow, 2000). Essentially, a SCC is a directed cyclic component in which flow direction within that component can reverse (Perelman & Ostfeld, 2012). Therefore, SCCs are parts of network where water is circulating during the simulation (Vasilić et al., 2016). Due to that fact, control of the water balance and/or water pressure regulation in SCC parts of the network could be difficult to achieve, so the idea is to detect SCCs and treat them as aggregated nodes in further network analysis and clustering. Algorithms for the extraction of SCCs from digraph are well known in the 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 compared to the others. This is significant as algorithm has to be able to deal with large networks efficiently. Gabow's algorithm requires only one pass through the network (DIGRAPH) with recursive call of the Depth First Search (DFS) algorithm (Tarjan, 1971) with arbitrary selection of the starting node.

For illustration purposes, a simple digraph shown in Figure 2 is used. Starting the DFS search from the node 2, nodes 3, 4, 6, 1 and 5 are visited (Figure 3-a). During the DFS search, a check is made weather the selection of the next node forms a cyclic path or not. If yes, nodes forming the cyclic path are identified as a SCC. The algorithm continues until no further propagation is possible. In example shown in Figure 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 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. DFS search is repeated again starting from node 8, and third SCC composed of nodes 8 and 7 is detected (Figure 3-b). At the end, aggregated digraph is composed of three identified SCCs. The digraph can also be viewed as set of aggregated nodes and two remaining connected to transmission main with one link (Figure 3-c). The most important property of new aggregated digraph is its acyclicity, indicating it is a digraph without cycles. Such graph is referred to as Directed Acyclic Graph (DAG), and in terms of water network is very important, because it clearly separates source from the demand nodes and hence, makes the sectorization of 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 recursive DFS algorithm, as explained in Sedgewick & Wayne (2011), is used for this purpose. In an example shown in Figure 3-c, topological sorting yields 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). Network uniformity index (Vasilic, 2018) is defined as follows:

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

where u_{net} is network uniformity in terms of cluster size, u_v is uniformity of the DMAs size vector and w_{agg} is relative weight of aggregated links. Each of these variables are explained in the following paragraphs, followed by the explanation of the aggregation algorithm itself.

Each cluster is characterized with its size (S_i), calculated as sum of all nodal demands within that cluster - $S_i = \sum_{j=1}^{N_n^i} q_j$, N_n^i being number of nodes in i-th cluster. Network uniformity (u_{net}) measures average deviation of clusters 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 size are calculated based on the daily average total demand in the WDN (Q_{tot}, available from the WDN hydraulic model), the number of minimum and maximum connections in the DMA ($n_c^{\min}$ and $n_c^{\max}$) and a total number of connections in the WDN (n_c), given as an input data:

$$\begin{aligned} S_{\min} &= \frac{Q_{tot}}{n_c} n_c^{\min} \\ S_{\max} &= \frac{Q_{tot}}{n_c} n_c^{\max} \end{aligned} \tag{2}$$

Network uniformity is calculated based on the triangular function f that quantifies "quality" of cluster size in the rage [0,1] (Figure 4). If a cluster i has a size $S_i = S_{pref}$, its value of f will be the best, i.e. 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 cluster 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 triangular one 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} is 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.

Sizing clusters in the range S_{min}–S_{max}, and as much as possible close to S_{pref}, is one sectorization objective. Sizing them equally is the other one. Sizes of all clusters form the normalized size vector of specific sectorization into N_{cl} clusters – $\boldsymbol{S^n} = \{S_1^n, S_2^n, S_3^n \ldots, S_{N_{cl}}^n\}$, 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):

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

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

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

If all nodes are part of the same cluster, meaning 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 final form of equation for its calculation:

$$u_v = \begin{cases} 1 - \dfrac{u_v\sqrt{N_{cl}} - 1}{\sqrt{N_{cl}} - 1}; & N_{cl} > 1 \\ 0 \qquad ; & N_{cl} = 1 \end{cases} \tag{6}$$

Relative weight of aggregated links is calculated as:

$$w_{agg} = \frac{\sum_{i=1}^{n_l^{agg}} D_i}{\sum_{i=1}^{n_l} D_i} \quad (7)$$

where n_l is total number of links, n_l^{agg} is number of links within the clusters, and D_i is links diameter. In case of 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. Minimum value of w_{agg} is zero, 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 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 initial number of clusters corresponds to the number of identified SCCs. Aggregation is iteratively carried out through three steps: 1) identification of candidates SCCs for aggregation, based on topologically sorted DAG; 2) selection and aggregation of the candidate with highest uniformity gain ($\Delta U_{\max}$); 3) aggregation of remaining downstream SCCs with positive uniformity gain ($\Delta U>0$). 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. This could happen as such SCCs usually have relatively small uniformity gain and aggregation would continue past them further upstream.

Uniformity index metrics that drives clustering process is made of three components as given with the eq. (1). Since the aggregation process is driven with the highest uniformity gain ($\Delta U_{\max}$), it is of interest to maximize all three components of 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 economically more 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 greedy optimization, based on highest modularity gain, used to maximize network's modularity index presented in Clauset et al.(2004). As with all similar type algorithms, it is not guaranteed that the global optimum solution will be found. However, the benefit is that generally a good sub-optimal solution can be found with significant computational time savings when compared to 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 described aggregation algorithm is illustrated on a simple example shown in Figure 5. The example is derived from Figure 3-c, adding 6 more SCCs for illustration purposes. For the sake of simplicity, total demand of 20 L/s is assigned to all 9 SCCs. Diameters of the links connecting SCCs are shown in Figure 5 in millimetres. Minimum (S_{min}) and maximum (S_{max}) DMA size are set to 40 and 80 L/s respectively, which yields preferred DMA size (S_{pref}) of 60 L/s. Figure 5 shows evolution of network uniformity index through aggregation process of this simple example. Uniformity index (U) is plotted against the number of clusters corresponding to each aggregation step (secondary horizontal axis).

Highest uniformity index value ($U_{\max}$) corresponds to network sectorization into 3 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 between them. Next aggregation step leads to the solution with 2 clusters, having total demands of 80 and 100 L/s. Obviously, this solution does not meet DMA size constraints, as one cluster is larger than S_{max}. However, there are now two links connecting 2 clusters which requires less isolation valves and flow meters to isolate them and create DMAs than in the case with 3 clusters. Figure 5 also illustrates hierarchical ordering of the sectorization solutions, embedded in the clustering algorithm. Solution with 3 clusters is lower in hierarchical order, and is easily derived from the solution with 2 clusters.

Heuristic device placement and solutions' evaluation (STAGE 3)

At the end of the 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 clusters' structure at each aggregation step (Figure 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 high value of network uniformity index are.

Prior to execution of the Stage 3 itself, selection of solutions that will be hydraulically analysed and evaluated for satisfaction of initially adopted PIs is made. Number of solutions (N_{sol}) for the Stage 3 analysis is specified by the user as an input parameter. Selection of solutions is made based on the network uniformity index values obtained at each aggregation step. Solution with the highest uniformity index is selected (best solution), together with additional N_{sol}-1 solutions from succeeding aggregation steps. Additional solutions are on the descending part of uniformity index plot (Figure 5), characterized by the lower value of uniformity index (than the best solution) but also by the smaller number of clusters. Described strategy for selection of solutions is adopted here as it 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 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 selection of solutions for evaluation has been made, main part of the Stage 3 is evoked. There are two main phases in the Stage 3: 1) Conversion of clusters into DMAs (Phase 1) and 2) Evaluation of solutions' PIs (Phase 2).

Phase 1. To convert clusters into DMAs (i.e. define DMAs), flow meters and isolation valves have to be positioned on clusters' boundary edges. Positioning of the flow meters and valves is done based on engineering heuristics. Continuing from the simple example used to describe aggregation algorithm (Figure 5), consider the solution with the highest value network uniformity index. This solution has 3 clusters and 4 boundary edges to be considered for installation of flow meters/valves. For methodology illustration purposes, another branch of transmission main and 4 boundary edges are added to this solution (Figure 6-a).

Boundary edges are labelled as L1 through L8, and numbers are showing links' diameters in millimetres. Flow orientations during 24-hour MDD hydraulic simulation, obtained in Phase 1 of the Stage 1, are indicated with arrows. Pipes with a changing direction (non-oriented) are indicated using dashed lines without arrows. Non-oriented pipes are only those connecting clusters with the transmission main, as identified clusters resulted from the DAG analysis (i.e. all other non-oriented 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 is comprised of the following three steps:

- Non-oriented pipes are identified, and all such 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 as negligible flow rate (hypothetically, let L2 be such 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's hydraulics.

- Supply pipes of each cluster (oriented towards cluster) are analysed independently. It is sufficient to analyse only supply pipes as graph in consideration is a DAG and one clusters' output pipes are others' supply pipes. Supply pipes for a cluster are identified and pipe with the largest maximum inflow to the cluster (Q_{max}) is considered as main supply pipe, and will not be considered for closure. 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 analysed 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 i-th 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 i-th pipe ($C_{cl} \geq q_{max}(i)$), 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 Figure 6 would result in closure of supply pipe L4 for cluster CL 1 and pipe L5 for cluster CL 2. Cluster CL 3 has only one supply link, so it remains opened.

Another approach for positioning flow meters and valves is the optimization method (e.g. Genetic algorithm - (Ivetić et al., 2013)) which considers each boundary pipe as closed or open. Since it is not uncommon that number of boundary edges exceeds several tens in case of real WDNs, the optimization method could be very time consuming hence it was not implemented here. At the end of the Phase 1, flow meters and isolation valves are positioned on the clusters boundary edges converting them into DMAs (Figure 6-b).

Phase 2. After definition of its' DMAs boundaries, each solution is subjected to the extended period hydraulic simulation to investigate the effects of modifications made to the network. Firstly, feasibility of solution is considered through evaluation of pressure constraints in each node:

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

where $p_{i,t}$ is pressure in i-th node in simulation time step t, and p_{min} and p_{max} are minimum and maximum allowable pressures in network. If solution does not meet pressure constraints it is considered unfeasible and it is excluded from further analysis.

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

1. *Cost* – Cost of the solution 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), calculated as mean value over the simulation time period (T). 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:

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

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

3. Average Water Age in the network over the last 24 hours of extended period simulation (WA):

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

Where WA_i^t is 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, equation (10) is used for WA calculation instead, in order to be comparable with other methodologies available in literature.

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

Selection of preferable sectorization solution

After the Stage 3, WDN sectorization is completed resulting in a set of feasible solutions. This is one of the main advantages of presented methodology, as it gives an array of alternative DMA designs to the decision maker. One can opt for a solution with large number of small DMAs or for a solution with small number of large DMAs, or anything in between. This is especially convenient for the analysis of large WDNs without previously established DMAs, where DMAs strategic planning should be addressed carefully. It is up to a decision maker to select sectorization solution best suitable to his preferences, based on calculated PIs and other parameters.

CASE STUDY

Description

Methodology presented in this paper has been tested on a large water distribution network. The analysed network was originally presented as second case study network in the Battle of the Water Sensor Networks competition (BWSN2 - Ostfeld et al. (2008)). It is a real life WDN slightly modified to preserve its anonymity. 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). Network consists of 12,523 nodes, 14,822 pipes, two reservoirs, two tanks, four pumps and five valves. Total demand in the network is Q_{tot} = 1,243 L/s and total number of connections in the WDN is n_c = 77,916.

The input data for 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 is downloaded from Exeter Centre for Water System (http://emps.exeter.ac.uk/engineering/research/cws/downloads/benchmarks/); 2) minimum number of connections per DMA n_c^{min} = 500, maximum number of connections per DMA n_c^{max} = 5,000; 3) transmission main diameter threshold is D_{main} = 350 mm; 4) pipe closure diameter threshold is D_{tr} = 300 mm; 5) minimum and maximum operating network pressures are set to p_{min} = 20 m and p_{max} = 75 m, maximum allowable water age is WA_{max} = 48 h; desired number of sectorization solutions N_{sol} = 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 equation (2) as S_{min} = 8 L/s and S_{max} = 80 L/s. For hydraulic modelling 24 hours MDD simulation is used, while for water quality modelling (*WA* calculation) extended period simulation of 192 hours is used.

Network clustering (STAGE 2)

Figure 7 shows the evolution of network uniformity index (*U*) through network clustering process done in the Stage 2, with maximum uniformity index value corresponding to 43 clusters (Umax=0.5112). Minimum number of clusters is 23 which is in accordance with research of Ferrari et al. (2014), in which the same transmission main diameter was used (350 mm) and 23 independent districts, connected to the main, were identified. Figure 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 11708 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 rate). As the plot suggests, w_{agg} constantly increase as aggregation proceeds, and changes only slightly in the final 77 steps as most of the links are already aggregated.

DMAs definition and evaluation (STAGE 3)

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

Beside adopted PIs used to evaluate the solutions, the following additional indicators are calculated to aid the evaluation of solutions using the methods proposed here, but also to enable a comparison with other literature methods (see corresponding section below):

1. Number of DMAs (N_{DMA}), number of meters (N_M) and number of valves (N_V),
2. *NL* – Number of DMAs larger than maximum DMA size (S_{max}),
3. *NS* – Number of DMAs smaller than minimum DMA size (S_{min}),
4. 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 it can be seen from Table 1, all solutions have relatively similar values of two PIs, *WA* and *Res*. As the number of DMAs in the solution decreases, average number of connections per DMA increases, meaning that DMAs are larger in size. Consequently, for creation of smaller number of larger DMAs requires less flow meters and isolation valves resulting in lower solutions' cost. Solution Sol-2 has one DMA which is smaller than minimum size S_{min}. In solutions Sol-3 to Sol-9 all DMAs are within specified $S_{min} - S_{max}$ range, while in the solutions Sol-10 to Sol-15 there are one or two DMAs that are larger than S_{max}.

Selection of preferable sectorization solution

The preferable solution is identified by analysing the solutions that fully satisfy the DMA size constraints, i.e. solutions Sol-3 to Sol-9. As noted earlier, all feasible solutions have similar impact on network's resilience (Res = 0.880 – 0.885) and water age (WA = 33.88 – 34.13 h). Therefore, Sol-9 is preferred solution over the Sol-5, as it is the less costly.

Figure 9 shows the preferred solution Sol-9 where the analysed WDN is sectorized into 35 DMAs, together with the detail of DMA #23 with the position of valves and flow meters. These positions are identified using heuristic approach described in Phase 1 of the Stage 3. Originally, the cluster that this DMA belongs too had 6 boundary pipes. Three of them were identified as links that always return water to the transmission main, and as such are marked for closure (V2, V3 and V4). Other three boundary pipes are "always-input to the zone" pipes, and using described methodology pipe V1 (D = 203.2 mm) is selected for closure, while other two pipes with larger diameters (D = 304.8 mm) are left opened 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 solution listed above (see Table 1), for each DMA in a solution following PIs are calculated:

1. p_{DMA}^{av} – mean average pressures over the 24 hours in a DMA, as a good indicator of network interventions' impact on pressure distribution, calculated as:

$$p_{DMA}^{av} = \frac{\sum_{i=1}^{n_j} \sum_{t=1}^{24} p_j^t}{n_j} \qquad \forall n_j \in DMA \qquad (11)$$

2. Res_{DMA} – Average resilience index for a DMA, calculated per equation (9), only this time accounting for nodes within considered DMA and
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 of size between DMAs in terms of demand.

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

Figure 10 and Figure 11 show results for each of 35 created DMAs in selected solution Sol-9. Figure 10-a shows average consumption in DMAs, with highlighted minimum and maximum size constraints. As it can be seen from the graph, identified 35 DMAs vary in size considerably but always within the design limits imposed. Figure 10-b shows relative changes in mean average pressure in DMAs, compared to mean average pressures in nodes that are part of that DMA in the original non-sectorized network (Δp_{DMA}^{av}). For most DMAs the mean average pressure has slightly decreased (up to 4%), whilst slight increase occurs in six DMAs (up to 1%). Therefore, network sectorization had very limited impact on re-distribution of pressure within the WDN. Significant decrease of pressure is observed in DMA #8 (by 13%), but all pressures are still whithin the required range of $p_{min} - p_{max}$.

Figure 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%, while increase is almost 30%. While decrease of *WA* is desirable, 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 (Figure 12) it is easy to conclude that *WA* is still well below set maximum WA_{max} of 48 h. Figure 11-b shows relative changes in DMAs resilience index (ΔRes_{DMA}). Changes in resilience index range from -3.5% to +2.2%, indicating very limited impact of sectorization on the resilience of the WDN.

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

Comparison of results with other methods

Finally, a comparison of results obtained here is made to the corresponding results obtained using 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 the cost was not explicitly reported in other papers. Reported values of PIs in the Table 2 refer to best sectorization solutions reported by each research. 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: a) partitioning the WDN and b) positioning the flow meters and isolation valves.

As it can be seen from Table 2, only the methodology presented in Hajebi et al. (2016) and DeNSE method, presented here, produce a set of feasible solutions. A total of 78 feasible solutions are identified in Hajebi et al. (2016) with solutions having between 28 and 48 DMAs. Regarding the DMA size constraints, solutions presented by Grayman et al. (2009) and Diao et al. (2014) have DMAs that are both larger and smaller, while 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 methodology presented here this is case for 7 out of 14 feasible solutions.

Methodologies using MO optimization to position flow meters and isolation valves (Hajebi et al. 2016, Zhang et al. 2017) require significant amount of computational time (15 hrs and 278 hrs respectively). Substantially lower computational time of Hajebi's method, compared to the method of Zhang, can be attributed to the use of shorter extended period simulation time (48 h compared to 192 h) used to calculate *WA*. Issue of high computational time, as a consequence of using MO optimization, is addressed in Diao et al. (2013) in which two stage heuristic procedure for device placement is applied, resulting in acceptable running time of around 20 min. However only one solution with 41 DMAs is reported with three DMAs out of required size limits. Engineering based heuristic procedure used in methodology presented here takes similar amount of time (about 20 min), but produces a set of feasible solutions compared to the research of Diao et al. (2013). Computational efficiency of DeNSE method is even more emphasized when compared to the method of Hajebi et al (2016). Both methods are able to produce a set of feasible solutions, but DeNSE takes only 20 min (compared to 15 hrs) and yet it uses a longer extended period analysis for WA calculation (192 hrs compared to 48 hrs).

Methodologies of Ferrari et al. (2014) and Hajebi et al. (2016) ensure connectedness of each DMA to the transmission main (direct access to water source) and their isolation from other DMAs (so called isolated DMAs – iDMAs). While methodology presented here does not create iDMAs, preferable solution presented earlier (Sol-9) fulfils condition of direct access to 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 also gives comparison of main performance indicators values for best reported solutions, obtained with different methods – water age (*WA*) and resilience index (*Res*). Presented results show that DeNSE method achieves slightly better value of resilience index and slightly worse value of water age. 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. (2014) and 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 h to 31.51 h). In the case of DeNSE method *WA* is increased by 3.31 % for the DMA system (from 32.91 h 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-life sized water distribution network BWSN2 (Ostfeld et al., 2008). The results obtained were compared to several other literature sectorization methods that used the same case study network. Based on this the following conclusions are drawn:

1. The DeNSE method is able to identify a set of good, feasible network sectorization solutions for a large water distribution network such as the one used in the case study here. The method is able to do this 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 noticeable especially when

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 literature methods where costs are assessed indirectly, e.g. via a number of installed new devices or summarized diameters (e.g. Hajebi et al., 2016) Even though 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 budgets available.
4. The DeNSE method seems particularly well suited for the application at the initial stages of the DMA design process and in low efficient WDNs (i.e. WDNs with higher water losses). This is because the method enables: (a) alternative DMA sizes (both small and large) to be considered and analyzed and (b) preservation of network hydraulic performance and reliability which, in turn, enables tracking the network water balance more easily, as opposed to other literature methods which seem to focus more on controlling 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 etc.

REFERENCES

Alvisi, S. (2015). A New Procedure for Optimal Design of District Metered Areas Based on the Multilevel Balancing and Refinement Algorithm. *Water Resources Management*, *29*(12), 4397–

4409. https://doi.org/10.1007/s11269-015-1066-z

Alvisi, S., & Franchini, M. (2014a). A heuristic procedure for the automatic creation of district metered areas in water distribution systems. *Urban Water Journal*, *11*(2), 137–159. https://doi.org/10.1016/j.proeng.2014.02.006

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

Babić, B., Stanić, M., Prodanović, D., Džodanović, B., & Dukić, A. (2014). Reducing uncertainty of infrastructure leakage index - A case stud. *Procedia Engineering*, *89*, 1577–1584. https://doi.org/10.1016/j.proeng.2014.11.459

Baker. (2007). *Municipal Water Distribution System Secutrity Study: Recomendations for Science and Technology Investments.*

Brentan, B. M., Campbell, E., Meirelles, G. L., Luvizotto, E., & Izquierdo, J. (2017). Social Network Community Detection for DMA Creation: Criteria Analysis through Multilevel Optimization. *Mathematical Problems in Engineering*, 1–12. https://doi.org/10.1155/2017/9053238

Burrows, R., Crowder, G., & Zhang, J. (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 Ld.

Campbell, E., Izquierdo, J., Montalvo, I., & Perez-Garcia, R. (2016). A Novel Water Supply Network Sectorization Methodology Based on a Complete Economic Analysis, Including Uncertainties. *Water*, *8*(5). https://doi.org/10.3390/w8050179

Chianese, S., Nardo, A. Di, Natale, M. Di, Giudicianni, C., Musmarra, D., & Ingegneria, D. (2017). DMA Optimal Layout for Protection of Water Distribution Networks from Malicious Attack. In *CRITIS 2017:International Conference on Critical Information Infrastructures Security* (Vol. 1, pp. 84–96). Springer International Publishing. https://doi.org/10.1007/978-3-319-99843-5

Ciaponi, C., Murari, E., & Todeschini, S. (2016). Modularity-Based Procedure for Partitioning Water

Distribution Systems into Independent Districts. *Water Resources Management*, *30*(6), 2021–2036. https://doi.org/10.1007/s11269-016-1266-1

Clauset, A., Newman, M., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *70*(6 2), 1–6. https://doi.org/10.1103/PhysRevE.70.066111

De Paola, F., Fontana, N., Galdiero, E., Giugni, M., Savic, D., & Sorgenti Degli Uberti, G. (2014). Automatic multi-objective sectorization of a water distribution network. *Procedia Engineering*, *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. *Journal of Hydraulic Engineering*, *134*(6), 822–832. https://doi.org/10.1061/(ASCE)0733-9429(2008)134:6(822)

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

Di Nardo, A., Di Natale, M., Giudicianni, C., Santonastaso, G. F., Tzatchkov, V., & Rodriguez, M. (2017). Economic and Energy Criteria for District Meter Areas Design of Water Distribution Networks, 1–13. https://doi.org/10.3390/w9070463

Di Nardo, A., Di Natale, M., Santonastaso, G. F., & Venticinque, S. (2013). An Automated Tool for Smart Water Network Partitioning. *Water Resources Management*, *27*(13), 4493–4508. https://doi.org/10.1007/s11269-013-0421-1

Di Nardo, A., Giudicianni, C., Greco, R., & Santonastaso, G. F. (2018). Applications of Graph Spectral Techniques to Water Distribution Network Management. *Water*, *10*(1), 1–16. https://doi.org/10.3390/w10010045

Diao, K., Zhou, Y., & Rauch, W. (2013). Automated Creation of District Metered Area Boundaries in Water Distribution Systems. *Journal of Water Resources Planning and Management*, *139*(2), 184–190. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000247

Farley, M. (2001). Leakage Management and Control : A Best Practice Training Manual, 1–169. Retrieved from http://whqlibdoc.who.int/hq/2001/WHO_SDE_WSH_01.1_pp1-98.pdf

Ferrari, G., Savic, D., & Becciu, G. (2014). A Graph Theoretic Approach and Sound Engineering Principles for Design of District Metered Areas. *Journal of Water Resources Planning and Management*, *140*(12), 1–13. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000424

Fortunato, S. (2009). Community detection in graphs. 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. *Information Processing Letters*, *74*, 107–114.

Giudicianni, C., Nardo, A. Di, Natale, M. Di, Greco, R., Santonastaso, G. F., & Id, A. S. (2018). Topological Taxonomy of Water Distribution Networks. *Water*, *10*(4), 1–19. https://doi.org/10.3390/w10040444

Giustolisi, O., & Ridolfi, L. (2014). New Modularity-Based Approach to Segmentation of Water Distribution Networks. *Journal of Hydraulic Engineering*, *140*(10), 04014049. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000916

Gomes, R., Marques, S. S., & Joaquim, S. (2012). Decision support system to divide a large network into suitable District Metered Areas. *Water Science & Technology*, *65*(9), 1667–1676. https://doi.org/10.2166/wst.2012.061

Grayman, W., Murray, R., & Savic, D. (2009). Effects of Redesign of Water Systems for Security and Water Quality Factors. In *World Environmental and Water Resources Congress 2009* (pp. 504–514). https://doi.org/10.1061/41036(342)49

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

Herrera, M., Canu, S., Karatzoglou, A., & Izquierdo, J. (2010). An approach to water supply clusters by semi-supervised learning. In *InternationalEnvironmentalModelling andSoftware Society (iEMSs) 2010International Congress on EnvironmentalModelling andSoftware-Ottawa, Canada*. Retrieved from http://www.iemss.org/iemss2010/index.php?n=Main.Proceedings

Herrera, M., Izquierdo, J., Pérez-garcía, R., & Ayala-cabrera, D. (2010). Water Supply Clusters by Multi-agent Based Approach. In *Water Distribution System Analysis 2010 – WDSA2010, Tucson, AZ, USA, Sept. 12-15, 2010 WATER* (pp. 861–869).

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

Jungnickel, D. (2005). *Graphs, Networks and Algorithms*. (M. Bronstein, A. Cohen, H. Cohen, D. Eisenbud, & B. Sturmfels, Eds.) (Second). Berlin: Springer.

Laucelli, D. B., Simone, A., Berardi, L., & Giustolisi, O. (2016). Optimal Design of District Metering Areas. *Procedia Engineering*, *162*(2014), 403–410. https://doi.org/10.1016/j.proeng.2016.11.081

Morrison, J., Tooms, S., & Rogers, D. (2007). *DMA management guidance notes*. London: IWA.

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *69*(2 2), 1–15. https://doi.org/10.1103/PhysRevE.69.026113

Ostfeld, A., Uber, J. G., Salomons, E., Berry, J. W., Hart, W. E., Phillips, C. a, … Walski, T. (2008). The Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms. *Journal of Water Resources Planning and Management*, *134*(6), 556–568. https://doi.org/10.1061/(ASCE)0733-9496(2008)134:6(556)

Perelman, L., & Ostfeld, A. (2012). *Water-Distribution Systems Simplifications through Clustering*. *Journal of Water Resources Planning and Management* (Vol. 138). https://doi.org/10.1061/(ASCE)WR.1943-5452.0000173

Scarpa, F., Lobba, A., & Becciu, G. (2016). Elementary DMA Design of Looped Water Distribution Networks with Multiple Sources. *Journal of Water Resources Planning and Management*, *142*(6), 04016011. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000639

Sedgewick, R., & Wayne, K. (2011). *Algorithms* (Fourth). Addison-Wesley.

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

Tarjan, R. (1971). Depth-first search and linear graph algorithms. *12th Annual Symposium on Switching and Automata Theory (Swat 1971)*, *1*(2), 146–160. https://doi.org/10.1109/SWAT.1971.10

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., Alcocer-Yamanaka, V., & Bourguett Ortíz, V. (2006). Graph Theory Based Algorithms for Water Distribution Network Sectorization Projects. *Water Distribution Systems Analysis Symposium 2006*, 1–15. https://doi.org/10.1061/40941(247)172

UK Water Research Industry. (1999). *A Manual of DMA Practice*. UK Water Industry Research Limted.

Vasilic, Z. (2018). *Decision Support Algorithms for Sectorization of Water Distribution Networks (PhD Thesis, Belgrade - december 2018)*. University of Belgrade.

Vasilić, Ž., Stanić, M., Prodanović, D., & Kapelan, Z. (2016). Network Sectorisation Through Aggregation of Strong Connected Components. In *18th Conference on Water Distribution System Analysis, WDSA 2016*.

WAA, & WRC. (1985). Water Authorities Association and Water Research Centre: Report 26 Leakage Control Policy & Practice.

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

Zhang, Q., Wu, Z., Zhao, M., & Qi, J. (2017). Automatic Partitioning of Water Distribution Networks Using Multiscale Community Detection and Multiobjective Optimization. *Journal of Water Resources Planning and Management*, *143*(9), 1–14. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000819.

LIST OF TABLES

LIST OF FIGURES

Table 1. Evaluation indicators for 14 feasible solutions

Sol ID [-]	N_{DMAs} [-]	***NL*** [-]	***NS*** [-]	A_{conn} [-]	***WA*** [h]	***Res*** [-]	***Cost*** [€]	N_M [-]	N_V [-]	u_{net} [-]	u_v [-]	***U*** [-]
Sol-2	42	0	1	1655	34.13	0.881	557,405	81	178	0.538	0.967	0.5073
Sol-3	41	0	0	1696	34.11	0.881	551,215	80	177	0.552	0.943	0.5070
Sol-4	40	0	0	1738	34.11	0.881	545,870	79	177	0.545	0.943	0.5013
Sol-5	39	0	0	1783	33.98	0.882	542,210	79	176	0.537	0.944	0.4943
Sol-6	38	0	0	1830	34.02	0.880	537,920	77	176	0.537	0.931	0.4872
Sol-7	37	0	0	1879	34.02	0.880	534,500	76	175	0.528	0.925	0.4767
Sol-8	36	0	0	1931	34.01	0.880	530,995	76	169	0.534	0.910	0.4744
Sol-9	35	0	0	1987	34.00	0.880	523,685	75	166	0.530	0.895	0.4633
Sol-10	34	1	0	2045	34.00	0.881	522,565	75	164	0.522	0.882	0.4496
Sol-11	33	1	0	2107	34.01	0.881	516,375	74	163	0.516	0.855	0.4318
Sol-12	32	2	0	2173	33.98	0.881	515,815	74	162	0.505	0.839	0.4145
Sol-13	31	2	0	2243	33.98	0.881	510,470	73	162	0.482	0.840	0.3957
Sol-14	30	2	0	2318	33.96	0.880	497,205	71	153	0.481	0.840	0.3956
Sol-15	29	2	0	2398	33.88	0.885	490,470	71	138	0.466	0.818	0.3736

Table 2: Comparison of results with other methods

Published in	**Method for**		N_{DMAs}	*NL*	*NS*	N_M	N_V	P_{add}	A_{conn}	***Comp. Time***	***WA***	***Res***
	WDN partitioning	Device placement	[-]	[-]	[-]	[-]	[-]	[-]	[-]	[min/h]	[h]	[-]
Grayman et al. (2009)	Manual		43	1	3	53	163	11	1996	NA	31.51	0.802
Diao et al. (2013)	Comm. detection	2 stage heuristic method	41	2	1	NA	NA	0	2044	20 min	32.01	NA
Ferrari et al. (2014)	Graph based recursive bisection algorithm		36	0	0	181	152	0	2317	NA	NA	NA
Hajebi et al. (2016)	Heuristic graph partitioning	MO optimization	28-48	0	0	56-78	66-161	0	1415-2423	15 h	31.01	0.830
Zhang et al. (2017)	Comm. detection	MO optimization	43	NA	NA	103	33	0	NA	278 h	NA	NA
DeNSE Algorithm	Uniformity based clustering	Engineering based heuristic	29-42	0-2	0-1	71-81	138-178	0	1656-2398	20 min	34.00	0.880

* NA – not available; *WA* – Water Age for best reported solution; *Res* – Resilience Index for best reported solution

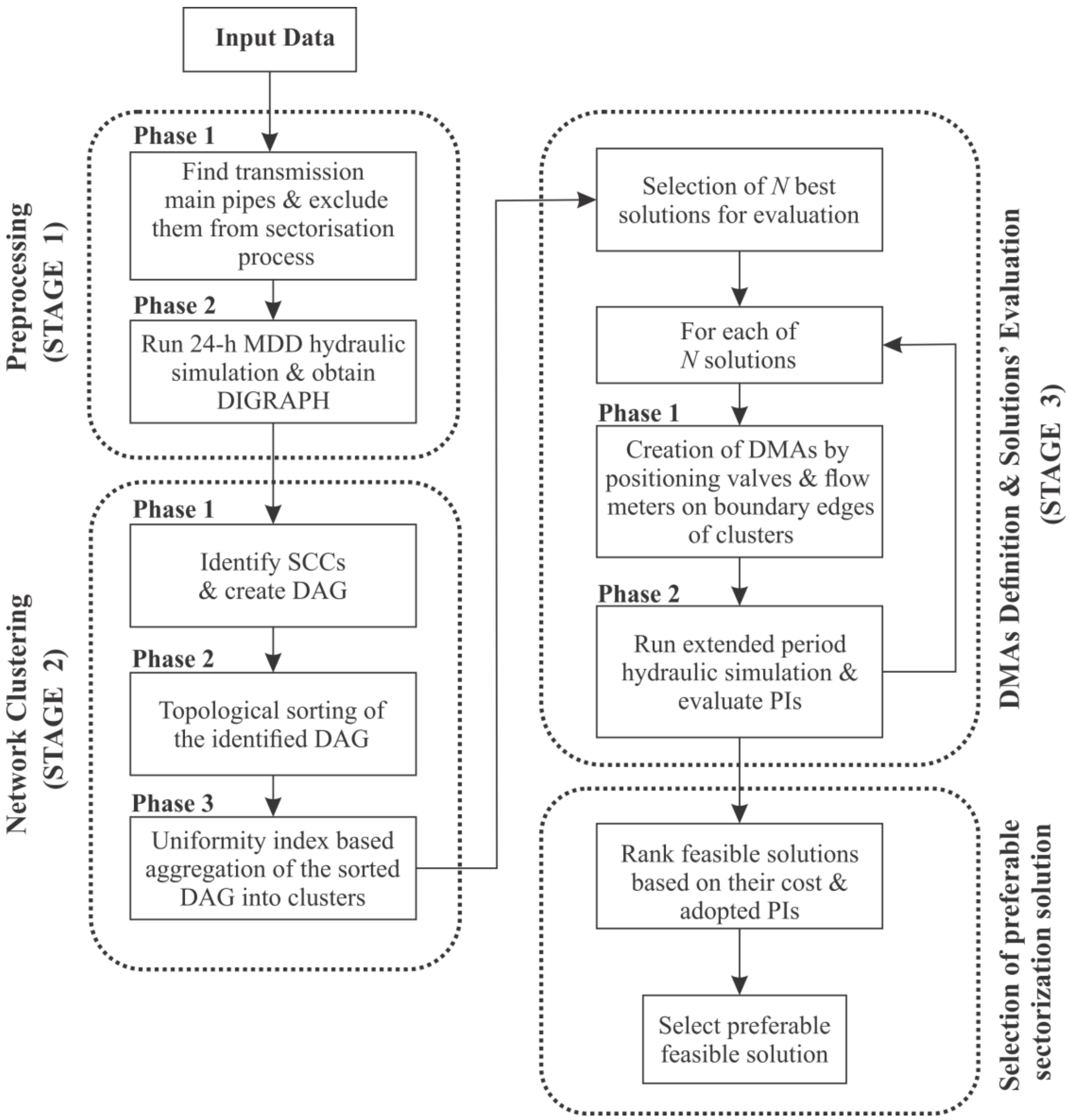


Fig. 1. Flow chart of the DeNSE sectorization method

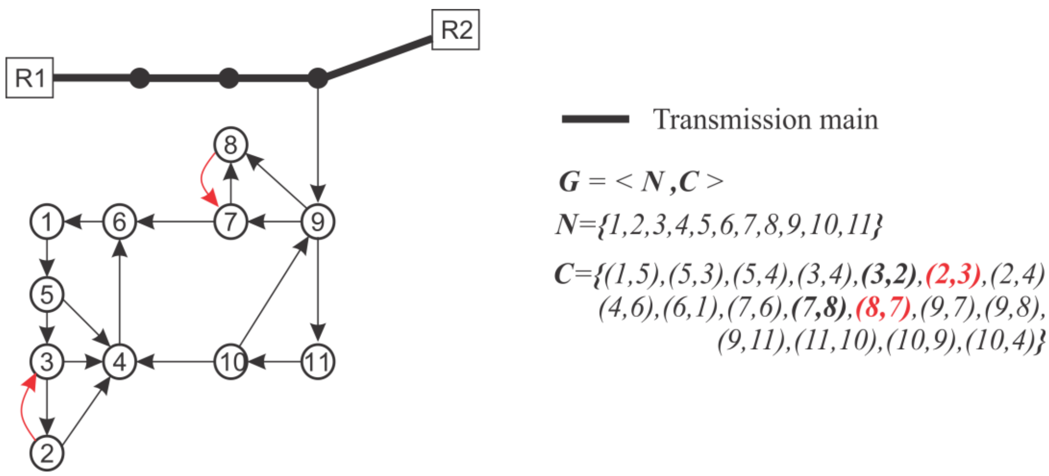


Fig. 2. Digraph presentation of a simple network with 2 sources and 2 undirected links

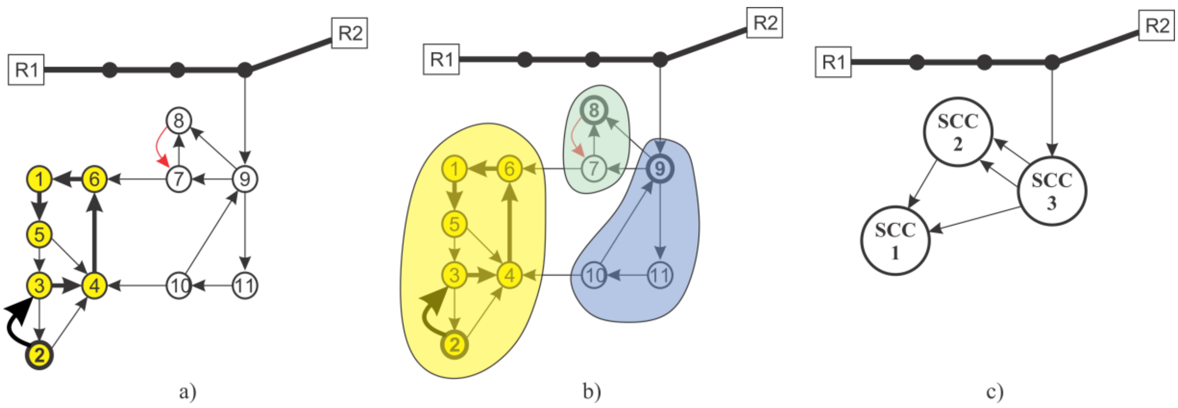


Fig. 3. DIGRAPH transformation to DAG: a) Start the DFS; b) Detected SCCs; c) Newly formed DAG

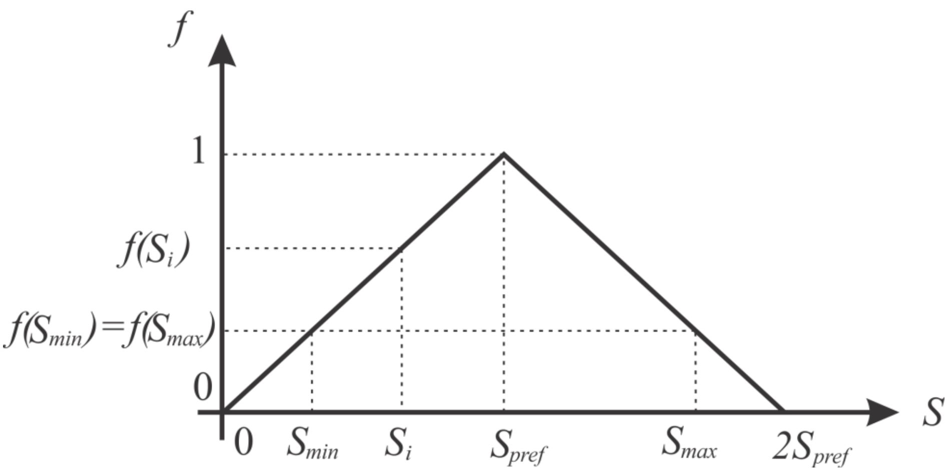


Fig. 4. Triangular function *f* quantifying cluster size "quality"

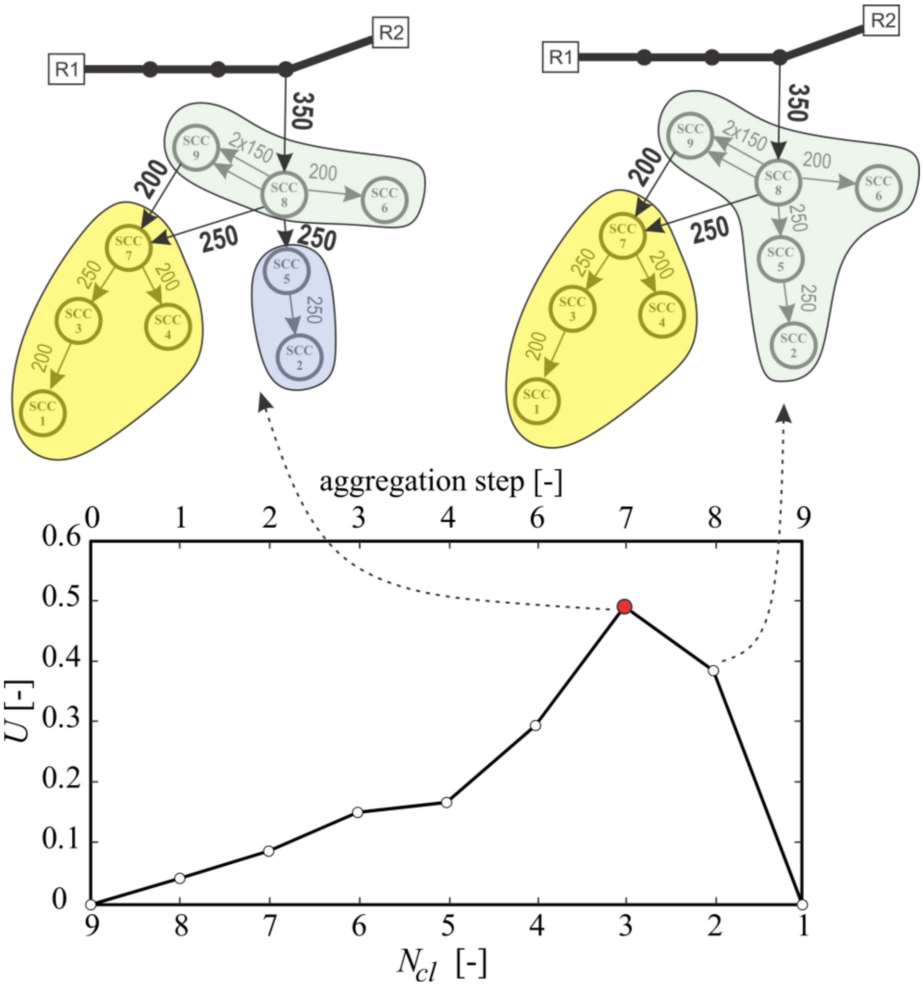
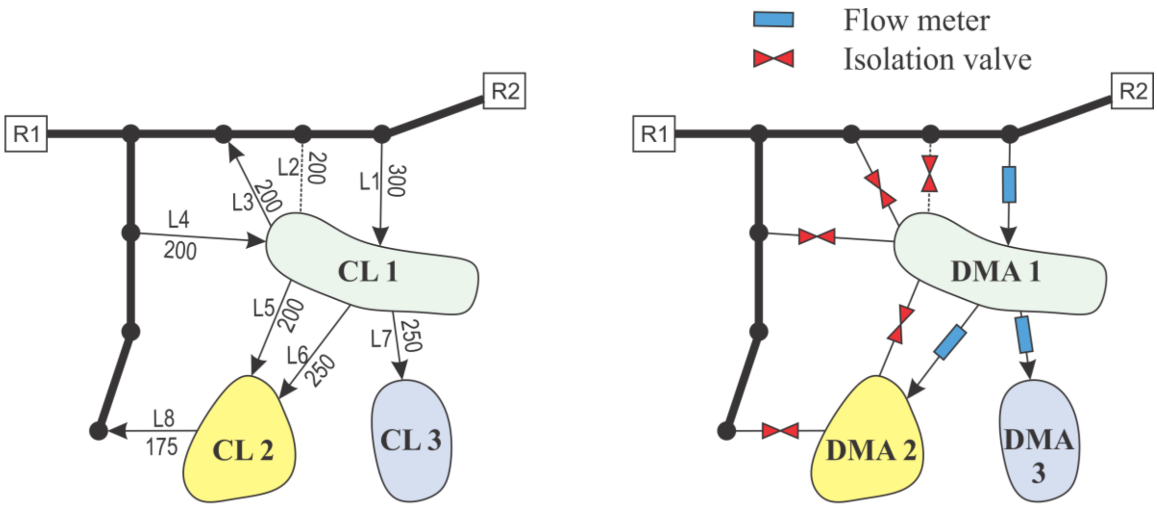


Fig. 5. Evolution of network uniformity index during aggregation process



a) b)

Fig. 6. Heuristic based placement of flow meters and isolation valves (Stage 3 – Phase 1)

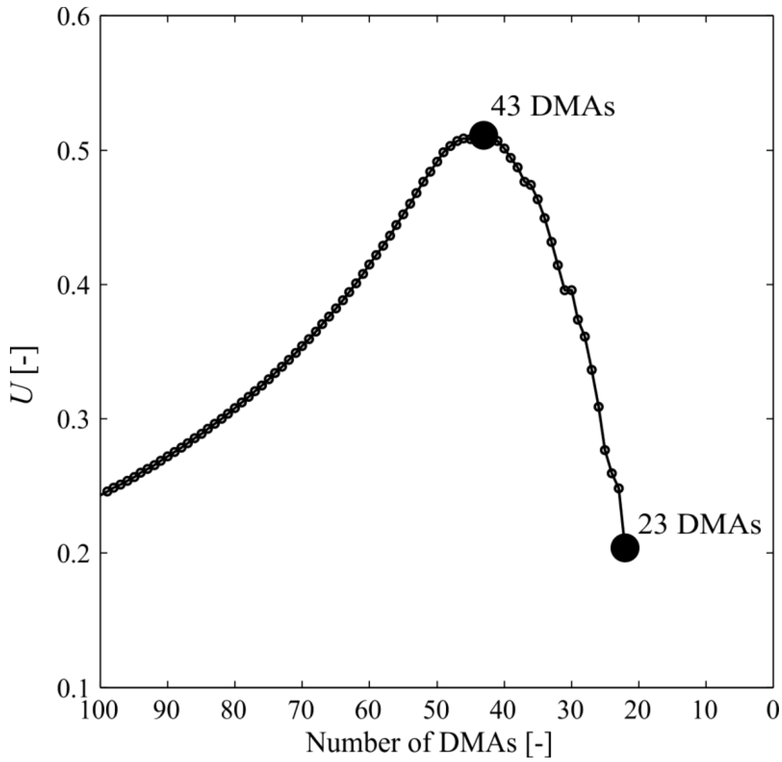


Fig. 7. Evolution of Uniformity Index during clustering of BWSN2 network

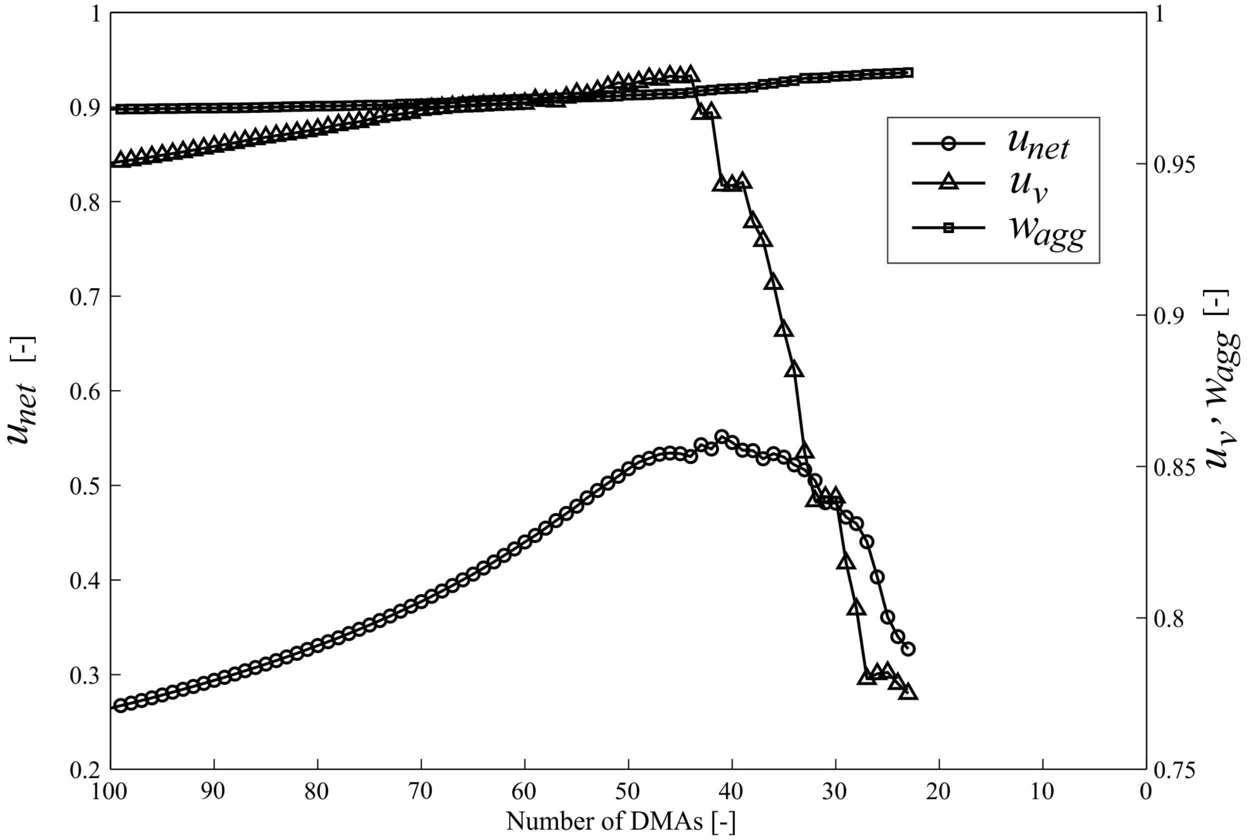


Fig. 8. Evolution of all three components (u_{net}, u_v and w_{agg}) constituted in the network uniformity index (U) in the last 77 aggregation steps

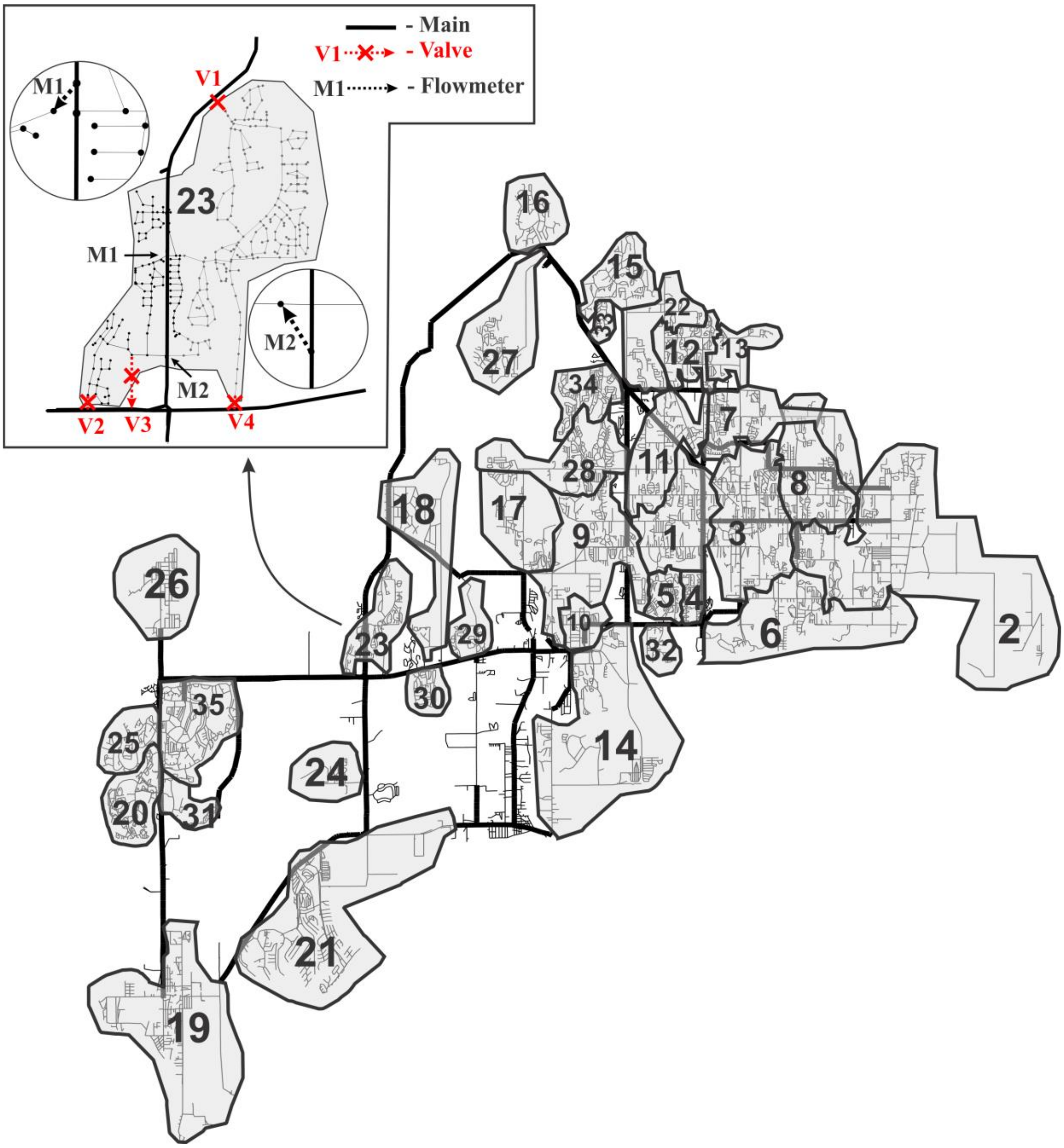


Fig. 9. Preferred sectorization solution (Sol-9) with 35 DMAs and detail of DMA #23

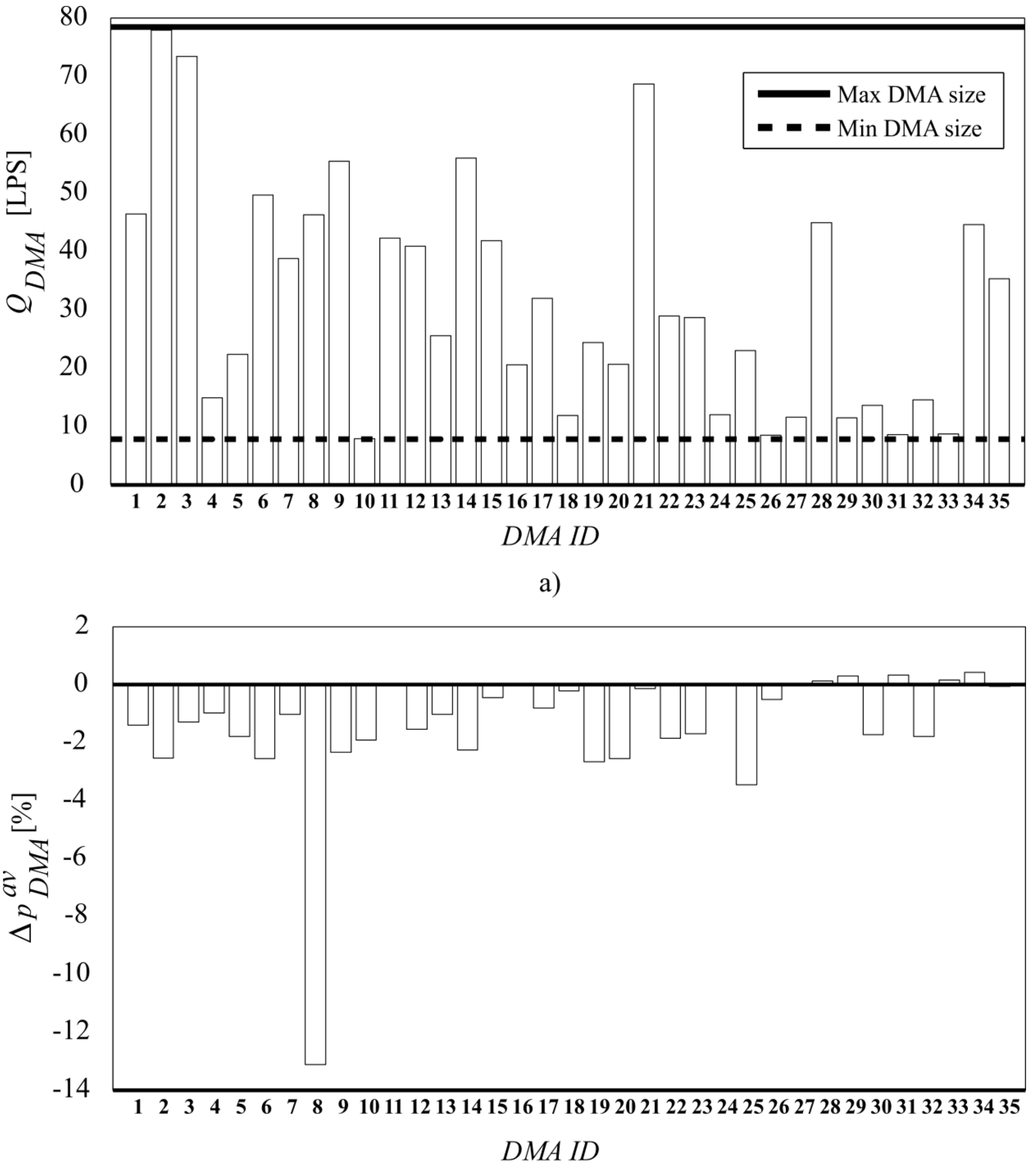
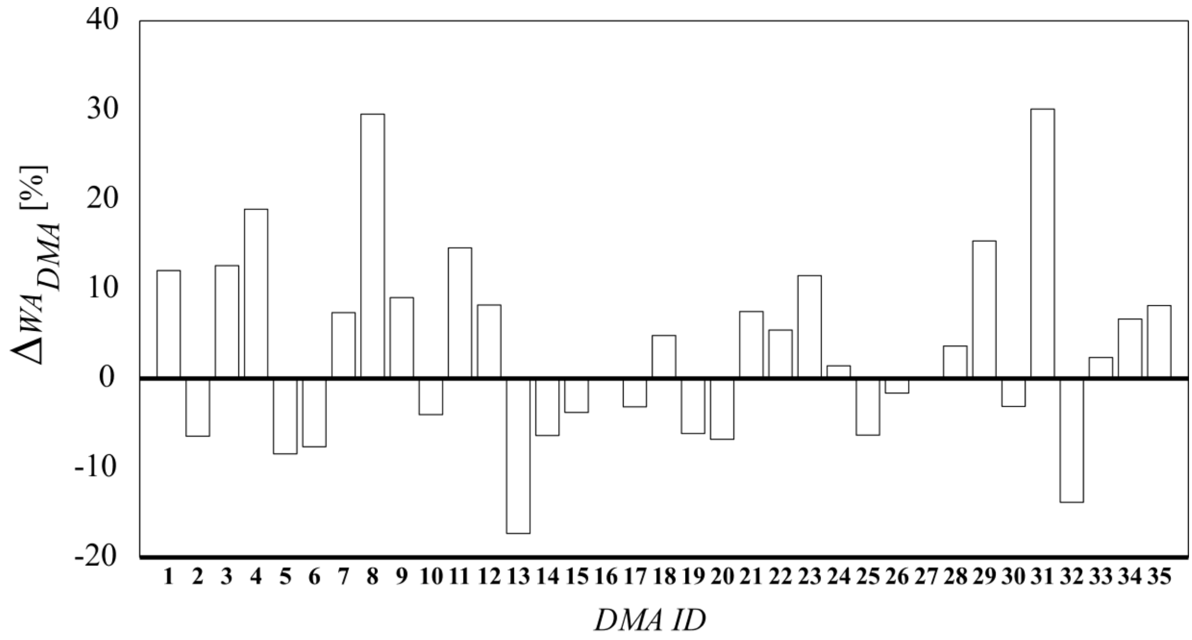
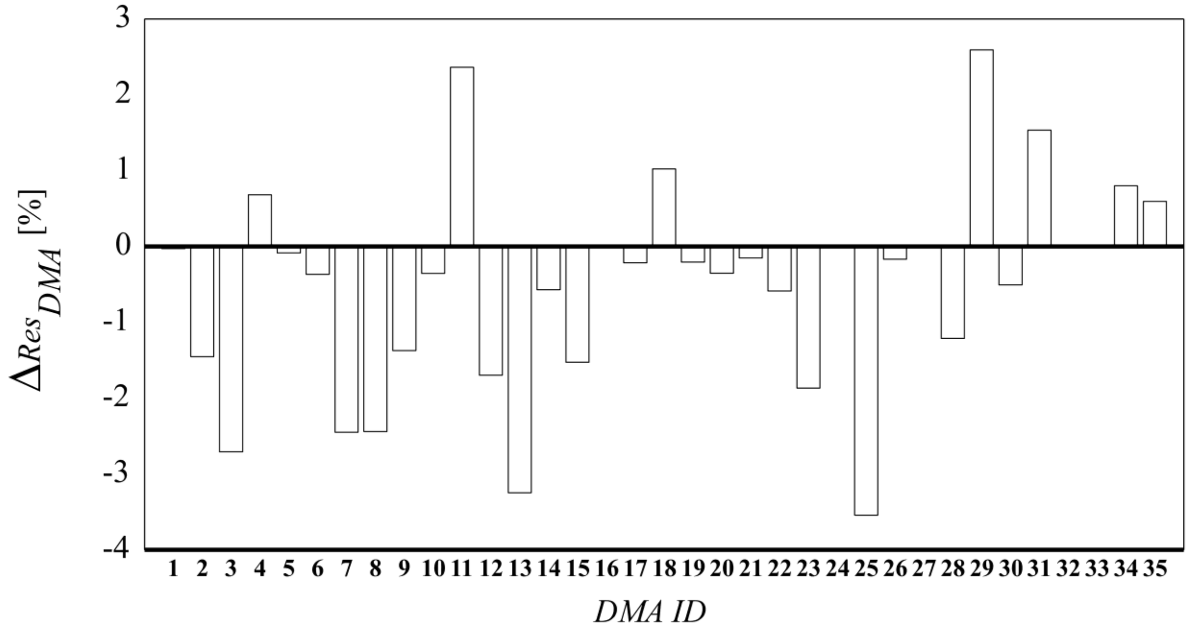


Fig. 10. Results for each DMA in selected preferable solution (Sol-9): a) average DMA consumption; b) relative change of mean average pressure



a)



b)

Fig. 11. Results for each DMA in selected preferable solution (Sol-9): a) relative change of Water Age; b) relative change of Resilience Index

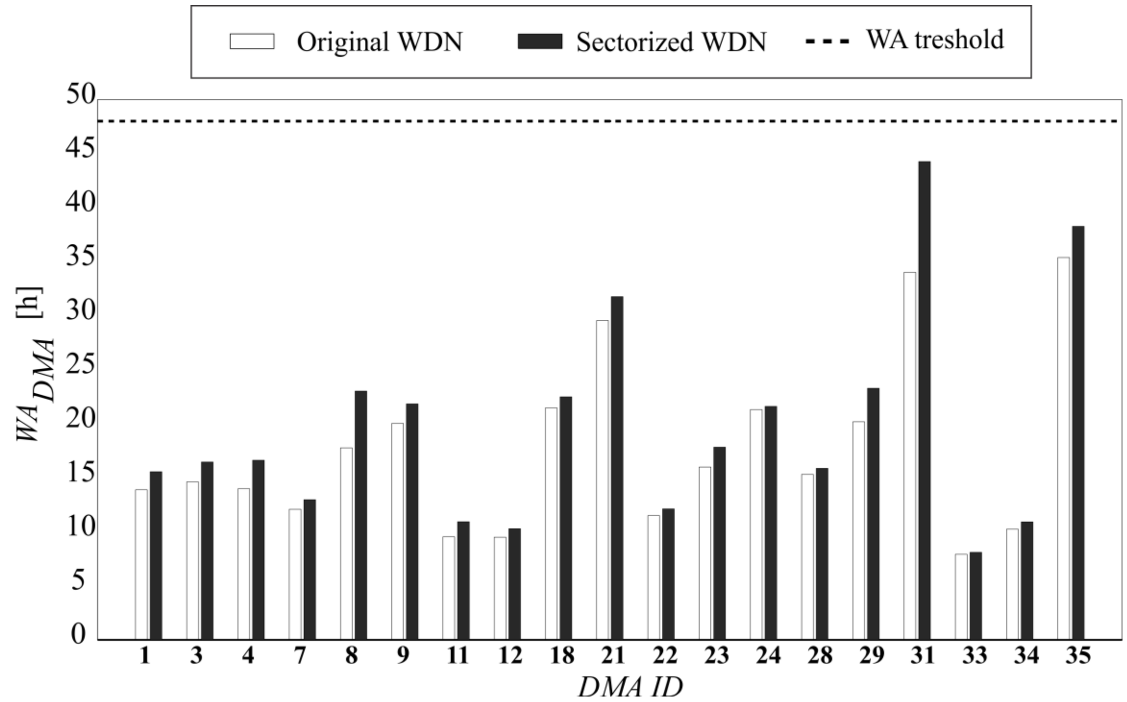


Fig. 12. Values of water age, before and after sectorization, for DMAs with increased water age