UNIVERSITY OF BELGRADE,
FACULTY OF CIVIL ENGINEERING
INSTITUTE FOR HYDRAULIC AND ENVIRONMENTAL ENGINEERING
11000 Belgrade, Bul. kralja Aleksandra 73
Phone:(381 11) 337-0206,
Fax:(381 11) 337-0206

Report on application of sectorization method for creation of District Meter Areas (DMAs) within the water distribution network of the city of Amsterdam

Horizon 2020 Marie Skłodowska-Curie Research and Innovation Staff Exchange project WAT-QUAL

Belgrade, September 2020.

Prepared by:

Željko Vasilić PhD, University of Belgrade – Faculty of Civil Engineering, Serbia

Damjan Ivetić PhD, University of Belgrade – Faculty of Civil Engineering, Serbia

The report has been prepared in collaboration with:

Ralf de Groot, Waternet, Netherlands

Arne Bosch, Waternet, Netherlands

This report has been produced in the framework of the Horizon 2020 Marie Skłodowska-Curie Research and Innovation Staff Exchange project "WAT-QUAL“.

DISCLAIMER

List of Figures

List of Tables

Table of Contents

1 INTRODUCTION

The subject of this report is presentation of abilities of sectorization algorithm to identify and create District Meter Areas (DMAs) within the water distribution network of the city of Amsterdam.

District Meter Area (DMA) 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 (Burrows et al., 2000). Typical DMA design is shown in Figure 1. Sectorization of Water Distribution Network (WDN) into District Meter Areas (DMAs) is a proven measure for proactive leakage and pressure control. It is considered as the most cost-effective strategy for the control of real water losses. Setting up DMAs can be potentially useful even for water utilities that operate WDNs that do not suffer from high volumes of non-revenue water, such as Amsterdam city WDN operated by Waternet. For example, improved control of the contamination spreading can be considered as an additional significant benefit. Sectorization of WDN must be designed carefully, as required network interventions can endanger network's reliability in terms of water supply and pressure distribution.

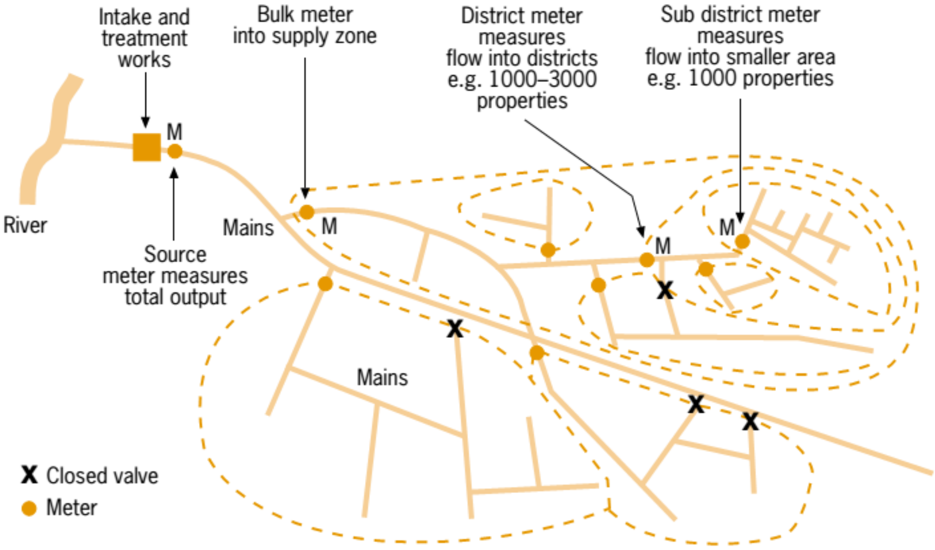

Figure 1. Typical DMA design options (adapted from Farley (2001))

Complexity of the real life WDN results in many different alternatives in which network sectorization can be done. Sectorization of WDN into an optimal system of DMAs is a hard task to achieve, especially for the existing and continuously operating WDNs. Every WDN is unique in its topology, characteristics and key sectorization objectives, so there is no common procedure for sectorization of WDN into DMAs, 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). Sectorization solutions are usually obtained by the "trial and error" technique conducted by a local expert, familiar with all 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 and to allow additional control and isolation of the system in order to improve water security. Beside general criteria given by the aforementioned guidelines (e.g. DMA size, network length), sectorization process should be driven by the case-specific criteria such as required number of feeds, fire flow regulations etc. A more formal approach to sectorization problem, that will enable investigation of alternative sectorization solutions for large WDNs, adopting local design criteria is presented within this research program.

Sectorization method employed in this analysis is named DeNSE (Distribution Network Sectorization). It is developed at the University of Belgrade, Faculty of Civil Engineering (Vasilic, 2018). DeNSE sectorization method is based on newly developed uniformity index metrics (Vasilic et al., 2020) which drives the sectorization process and identifies clusters in the network. Originally, DeNSE method relies on common engineering heuristics for placement of flow meters and valves to create DMAs. Although being able to produce a good set of feasible sectorization solutions, using only engineering heuristics limits the search space of potential solutions. In this research DeNSE methodology has been further improved to include an optimization procedure to ensure finding (sub)optimal sectorization solution within the broader specter of feasible sectorization solutions. Least investment for field implementation and maintaining the same level of WDN's operational efficiency are adopted as sectorization's main design criteria. Additional local design criteria, specific for the Amsterdam water distribution network, have been included also. These upgrades made to the original DeNSE method proved to be significant, which is confirmed by testing it on two real-life case studies which are part of Amsterdam WDN. Reported results prove that developed method can be used as a decision support methodology valuable to practicing engineers commencing implementation of sectorization strategy in WDN.

2 METHODOLOGY

DeNSE method used in this research is based on the Graph Theory for identification of Strong Connected Components (SCCs) and their aggregation into the clusters based on newly presented network uniformity index (U). Sectorization process should start with the definition of key sectorization objectives and design criteria, followed by the identification of indicators that will be used to assess impact of interventions made in the network. 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 adopted here. Such set of design criteria is most appealing to many water utilities, especially in initial stages of DMAs implementation. Number of indicators are adopted to evaluate the effects of the sectorization interventions and the quality of the obtained sectorization solution. Complete list of these indicators will be listed later in the text, but they are related most importantly to the change of network pressure (reflects the network's operational performance) and cost of the solution (reflects the required investment for implementation).

The new method requires hydraulic model of WDN as an input, 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 desired number of feasible sectorization solutions, as shown in Figure 2. 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 network clustering algorithm. WDN decomposition into clusters is done in the second stage, based on the uniformity index (U). Prior to the third stage, and after the network clustering process is finished, user input is required – selection of the preferable clustering solution, which is then subjected to the genetic algorithm (GA) based optimization procedure. GA optimization procedure is employed to determine the (sub)optimal positioning of the valves and flow meters on clusters' boundary links, in order to define DMAs. Extended period hydraulic analysis of the solution and evaluation of the adopted indicators quantifying the quality and "goodness" of the solution, are all part of the third stage. Having in mind that the GA optimization is stochastic procedure in essence, each run of the GA (stage 3) will yield a different positioning of flow meters and valves on clusters' boundary links, thus giving a different definition of DMAs. Therefore, number of runs of the third stage is equal to the desired number of optimized solutions to be compared. Finally, user can rank optimized solutions according to the adopted indicators and select preferable feasible solution.

Each of the three stages will be briefly explained in the following text. For full description of methodology and mathematical background reader is referenced to the relevant publications (Vasilic, 2018; Vasilic et al., 2020).

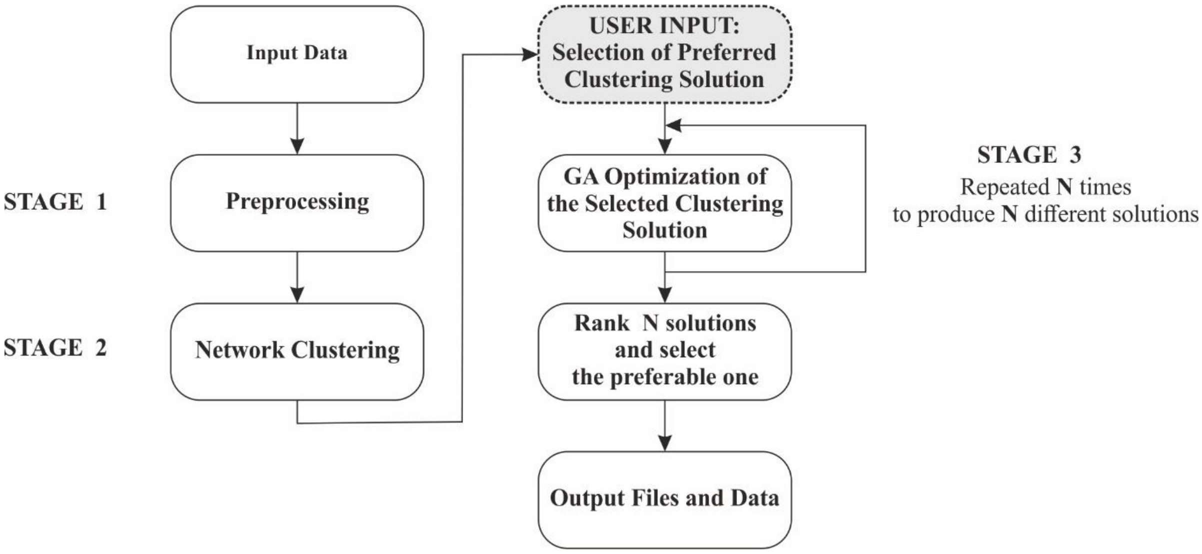

Figure 2. DeNSE sectorization method flow chart

2.1 INPUT DATA

DeNSE 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 ($\mathbf{n_c^{min}}$) and maximum ($\mathbf{n_c^{max}}$) number of property connections per DMA, as well as total number of connections in the network ($\mathbf{n_c}$), if number of connections per node is not available within the network model (which is the case usually). 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 ($\mathbf{D_{MAIN}}$). 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 $\mathbf{D_{MAIN}}$ is network specific, usually being 300-350 mm (Ferrari et al., 2014).

4. Minimum required and maximum allowed pressures in the network, $\mathbf{p_{MIN}}$ and $\mathbf{p_{MAX}}$ and

5. Desired number of alternate solutions for the definition of DMAs (**N**). This number is the number of runs of the optimization algorithm (stage 3). Each run will yield a different number of valves and flow meters, as well as their different positioning. Thus, the cost and other indicators will be different between solutions. It is considered that 5 solutions is minimum number to make a representative multi-criteria ranking, however user can opt for a larger set of solutions to compare.

2.2 PRE-PROCESSING

Pre-processing stage (stage 1) has two phases.

1. In the first phase, transmission mains are defined, based on the $\mathbf{D_{MAIN}}$ value, and excluded from the sectorization process. For this purpose, network is explored using slightly modified Breadth First Search (BFS) algorithm, simultaneously starting from all main source nodes (reservoirs). BFS algorithm is modified to prioritize propagation through the links with diameters equal or greater than $\mathbf{D_{MAIN}}$.

2. 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) is defined. Network links with changing flow directions are identified as non-oriented (or links that can have both flow directions).

Figure 3 shows simple network with 16 nodes, two of which are reservoirs (left). Right side of the figure shows result of the pre-processing done in the first stage. Links connecting reservoirs are identified as transmission mains and are excluded from further analysis. It should be noted that user can add any other link in the network to be excluded from sectorization. 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.

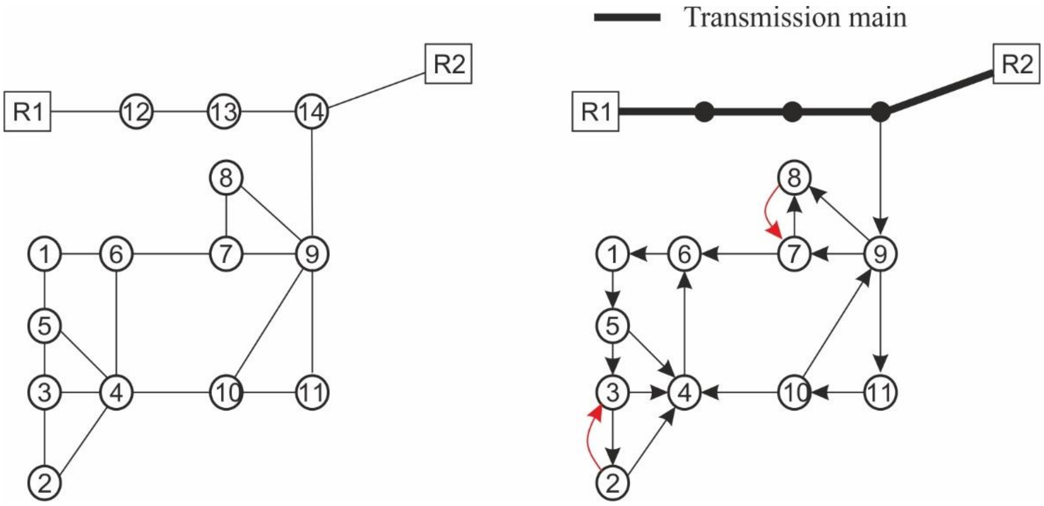

Figure 3. Simple network example (left) and result after preprocessing stage (right)

2.3 NETWORK CLUSTERING

In the second stage of the DeNSE method, partitioning of the WDN into clusters is performed. It is done in three phases which will be briefly listed here, but not described in detail. For full description reader is once again referenced to the relevant literature (Vasilic, 2018; Vasilic et al., 2020). Phases are:

1. First phase is identification of Strongly Connected Components (SCCs) within the DIGRAPH, created previously in the pre-processing stage (stage 1 - Figure 3 right). 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. Newly created graph is not only a directed but also acyclic (i.e. no circular path between the nodes can be made). This graph is called DAG – Directed Acyclic Graph.

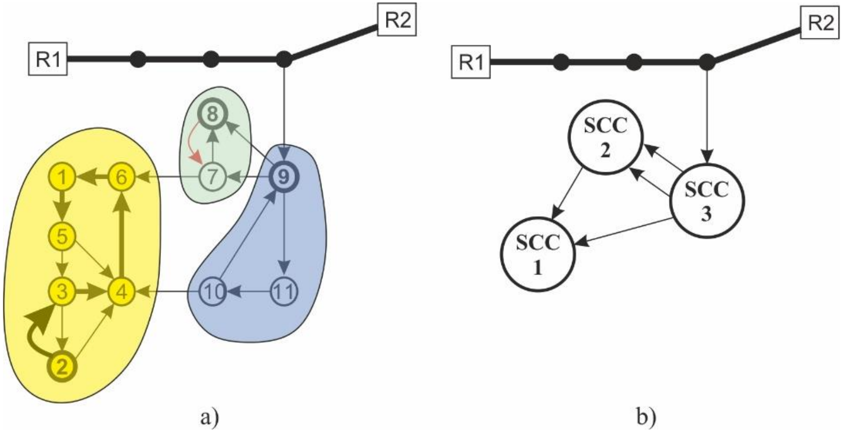

Figure 4. First phase of the network clustering procedure – identified SCCs (a) and DAG representation of the graph (b)

2. Second phase is topological sorting of the DAG identified in the previous phase. 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.

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:

$$\mathbf{U=u_{net}u_{v}w_{agg}} \quad \mathbf{(1)}$$

where $\mathbf{u_{net}}$ is network uniformity in terms of cluster size, $\mathbf{u_v}$ is uniformity of the DMAs size vector and $\mathbf{w_{agg}}$ is relative weight of aggregated links.

Each cluster is characterized with its size (**S_i**), which can be either total demand within the cluster or total number of the connections within the cluster. User decides which characteristic will be used for measuring the size of the cluster. If number of connections for each node in the mathematic model is available, it is probably the way to go, otherwise total demand should be used.

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 $S_{pref} = \left(S_{min} + S_{max}\right)/2$. 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.

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 and aggregating in each step SCCs whose aggregation will contribute the most to the network uniformity (**ΔU_{max}**). All three measures contained in the equation (1), which defines the network uniformity index, take values in the range from 0 to 1, hence uniformity index (**U**) also takes value in the same range. Higher value of network uniformity index indicates that the better uniformity is achieved.

Described aggregation algorithm, which is done in phase 3, is illustrated on a simple example shown in Figure 5. The example is derived from Figure 4-b, 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 (Spref) 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 (**N_{cl}**) corresponding to each aggregation step.

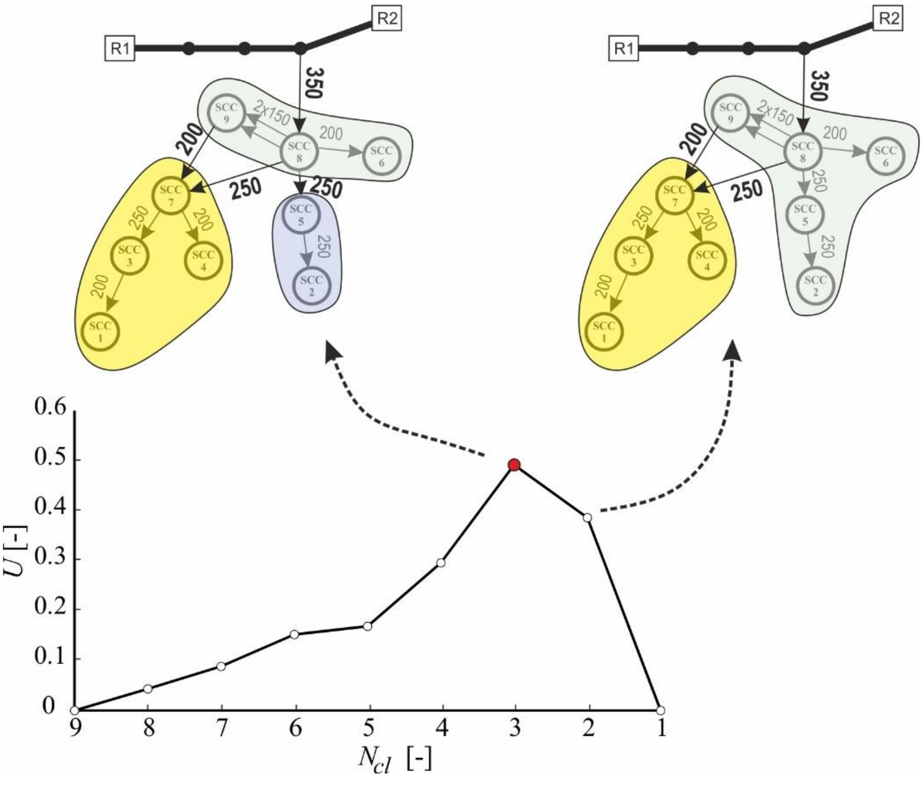

Figure 5. Evolution of network uniformity index (U) during aggregation process

Highest uniformity index value ($\mathbf{U}_{\max}$) 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 $\mathbf{S}_{\max}$. 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.

2.4 DEFINITION OF DMAS USING OPTIMIZATION

After the second stage (i.e. network clustering) user is required to select preferred clustering solution that will be optimized and analyzed further. After the selection of solution has been made stage 3 is evoked. To convert clusters into DMAs (i.e. define DMAs), flow meters and isolation valves have to be positioned on clusters' boundary edges, and this is done by employing optimization procedure based on Genetic Algorithm (GA).

Let's assume that user selected solution with the highest value of uniformity index – solution with 3 clusters and 4 boundary edges between them (see Figure 5). For methodology illustration purposes, another branch of transmission main and additional 4 boundary edges are added to this solution (Figure 6).

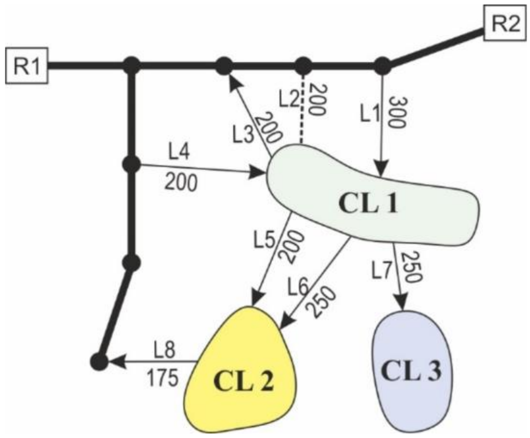

Figure 6. Selected clustering solution to be subjected to the GA optimization

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 (only link L2 in this example). Non-oriented pipes are only those connecting clusters with the transmission main, as identified clusters resulted from the DAG analysis.

We can see from the figure that there are 8 links in total that should be equipped either with valves or with flow meters in order to fully define the DMAs. Prior to GA optimization itself, a two step engineering heuristic procedure is employed in DeNSE method in order to reduce this number of pipes whose status should be determined (open-meter or closed-valve):

1. 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's say that the only one non-oriented link in this example L2 does not satisfy this condition) and

2. Boundary links that always return water from the clusters to the transmission main are closed. These pipes are not on supply paths, and as such can be considered redundant and closed without the effect on system's reliability (in this example on Figure 6 there are 2 such pipes – L3 and L8).

After applying this 2-step heuristic procedure number of pipes whose status should be determined in optimization algorithm is reduced from 8 to 6. This reduction in search space is even more emphasized in real sized network, where there are several tens or hundreds of boundary pipes, and hence can result in significantly lower computational times.

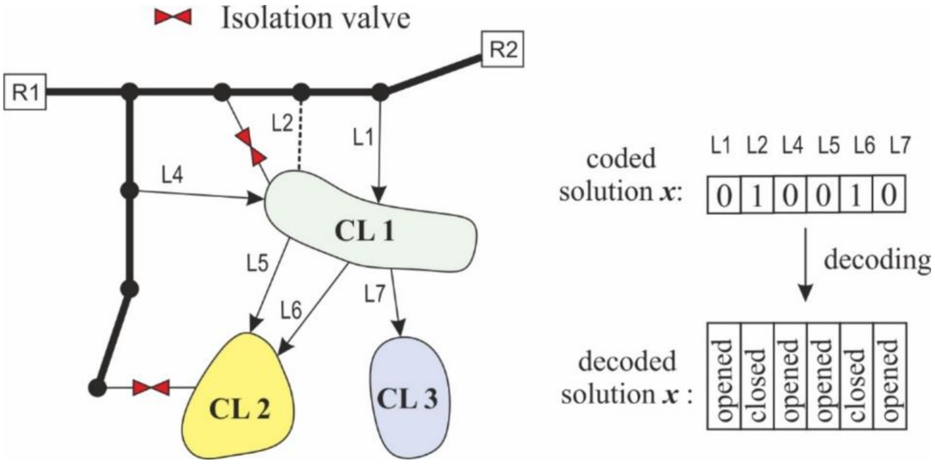

Figure 7. Remaining pipes subjected to the GA optimization

Pipe is either closed by placement of isolation valve, or it remains open and is equipped with flow meter. Pipe status is the only independent variable taking one of two values – opened or closed. In GA, solutions are coded into chromosomes represented with a string of bits (Figure 7). Parts of that string are coded variables (genes). Number of genes equals the number of pipes with unknown statuses. String of 1 bit is sufficient for representation of each gene, as there are only two possibilities for the status of the pipe (e.g. 1 – closed or 0 – opened).

After length of chromosomes is determined, population containing **m** individuals is initialized and its evolution process through generations begins, employing main GA's operations – selection, crossover and mutation. First step is to decode each solution from generation, run hydraulic simulation and evaluate its objective function (OF).

Next step is the selection of best performing individuals – parents to generate offspring. Selection process is responsible for controlled stochastic behavior of GA and it can follow different rules such as roulette-wheel or tournament selection. It means that randomness of the selection process is biased by the fitness of the objective function in the way that better solutions have better chance to be chosen (selected) and take part in the crossover.

Probability of implementing mutation is also the GA parameter and generally should be quite low (0.01) since the aim of GA is to be driven by crossover rather than mutation. By implementing selection, crossover and mutation, the new set of usually better solutions is created (new generation), and the whole process is now repeated (decoding, evaluation, selection, crossover, mutation) until the maximum number of generations is reached.

Finally, summarizing described GA method, parameters that have to be defined for its application are listed:

- Population size (**ps**) – number of individual solutions in population,
- Chromosome length – number of bits coding one solution,
- Crossover probability – usually 0.8 – 0.9,
- Mutation probability – usually < 0.05 and
- Number of generations to evolve (**ng**).

The efficiency of GA will depend on the adopted values for above listed parameters. These values are case specific, since different objective functions will require different set of values for parameters to achieve the same efficiency. Objective function for proposed implementation of the GA considers only economical aspect – solution cost. The informal definition of GA could be that it is optimization method that searches for optimum solution in discrete multidimensional space without constraints. Network sectorization problem is constrained with the request that any implemented interventions do not endanger network's operating reliability, providing feasible sectorization solution.

In DeNSE methodology, feasibility of the solution is imposed through several penalty functions used within the objective function (OF), penalizing each solution with proportional penalty value ($\mathbf{C_iP_i}$) – $\mathbf{C_i}$ being penalty unit cost and $\mathbf{P_i}$ proportional penalty amount:

$$\mathbf{OF = Cost + \sum_{i=1}^{4} C_iP_i} \qquad \mathbf{(2)}$$

Penalty types, their unit cost ($\mathbf{C_i}$) and calculation of proportional penalty amount ($\mathbf{P_i}$) are summarized in the following table:

Table 1. Penalty functions used within the objective function

i	**Type of penalty**	**Description**	**Ci**	**Pi**
1	Unfeasible solution	If there are negative pressures in the network or hydraulic model cannot be solved (does not converge to a solution)	**10^7 €**	$\mathbf{P_i = 1}$
2	Number of feed lines	DMA must have minimum number of feed lines defined according to the number of connections (local regulations)	**5×10^5 €**	$\mathbf{P_i = n_{cl} + \sum_{j=1}^{n_{cl}} (f_j^{req} - f_j^*)}$ n_{cl} – number of clusters that have lower number of feeds than required f_j^{req} – requested number of feeds for j-th cluster f_j^* – achieved number of feeds for j-th cluster
3	Pressure below minimum allowed	If pressure in any node is below minimum allowed (per local regulations), that solution is penalized proportional to the number of such nodes.	**5×10^4 €**	$\mathbf{P_i = n_j}$ n_j – Number of junctions with minimum pressures lower than minimum allowable in the network

4	Lowered pressures in the network	If average pressures in the network are lowered compared to the original network state, that solution is penalized proportional to that lowering of the pressure	**1x10⁴ €/m**	$\mathbf{P_i = p^{orig}_{av\_min} - p^{*}_{av\_min}}$ $p^{orig}_{av\_min}$ - minimum average pressure during 24h in the original network $p^{*}_{av\_min}$ - minimum average pressure during 24h in the network with implemented interventions

Objective function is minimized in the optimization procedure. Solution is feasible if:

a) network pressure is within the predefined range (p_{min}-p_{max}) and

b) number of direct DMAs' feeds is greater than minimum required for the corresponding DMA size.

2.5 INDICATORS FOR RANKING OF SOLUTIONS

After the third stage, solutions have been optimized and user can analyze and investigate all of desired **N** solutions and decide which is preferred one. Decision should be made based on the following indicators for ranking, that measure the effects of interventions made in the network to define the DMAs:

1. Comparison of average network pressure before and after interventions,
2. Relative change of pressure throughout the network caused by the interventions, calculated as:

$$\boldsymbol{\Delta p} = \frac{\sum_{i=1}^{nj}\sum_{t=1}^{24}(p_i^{*t} - p_i^t)}{\sum_{i=1}^{nj}\sum_{t=1}^{24}p_i^t} * 100\ (\%) \tag{3}$$

   where $\mathbf{p_i^t}$ is pressure in i-th node at time t in original network and $\mathbf{p_i^{*t}}$ is pressure in that same node in the same time, but in network with implemented interventions,

3. Solutions' implementation cost calculated per values given in the Table 3 as the installation cost of new valves and flowmeters,
4. Number of flow meters and new valves to be installed (already existing valves are used if possible),
5. Graphical representation of connectivity between the clusters.

First 4 indicators are produced as a 4-in-1 graph, so that the user can more easily compare the solutions (see Figure 8).

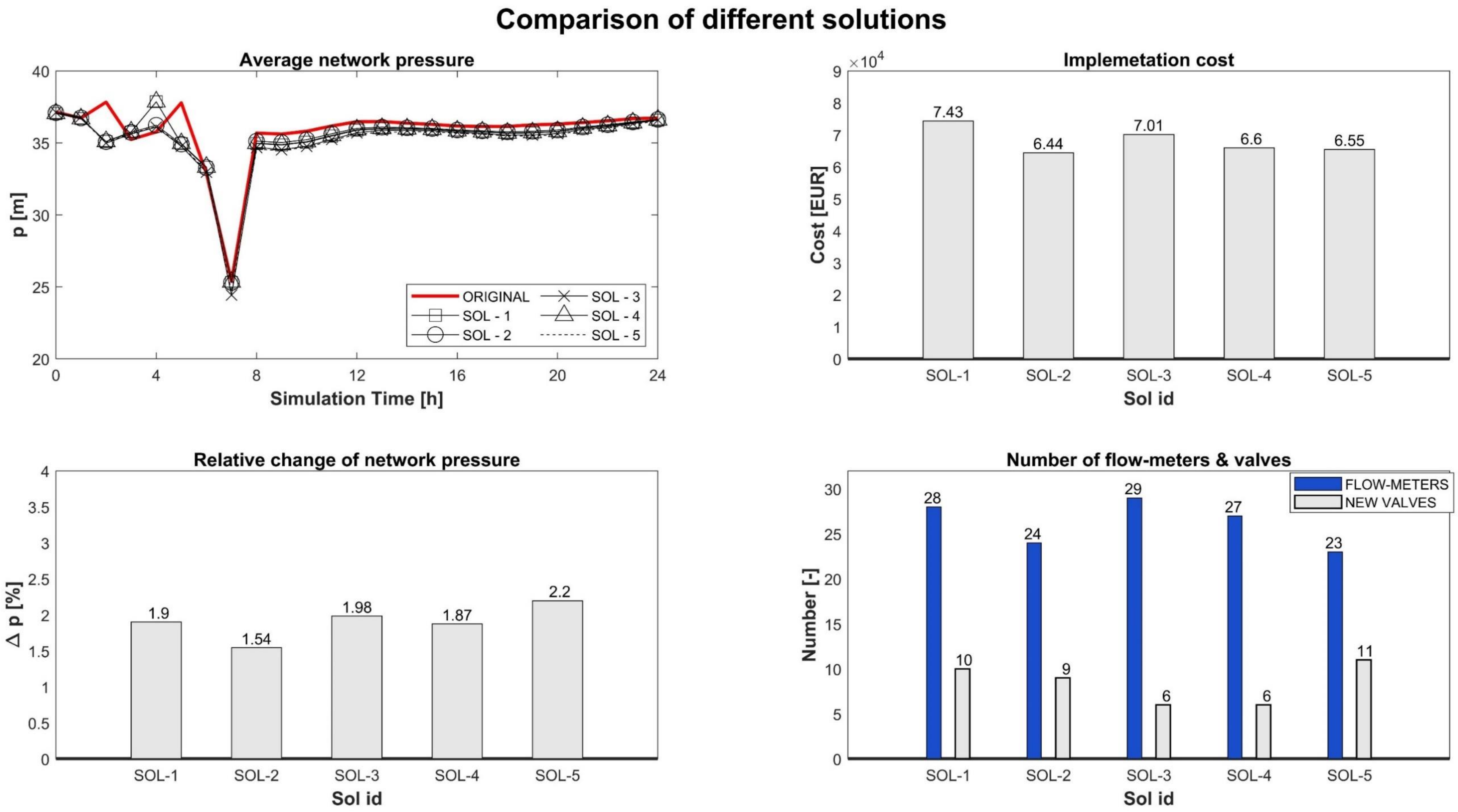

Figure 8. 4-in-1 graph showing 4 comparison indicators for ranking of different solutions

Graphical representation of connectivity is plotted for every solution (Figure 9). Figure shows example plot for one solution in which DMAs are represented as the points of different color connected to the main pipeline and between each other. Blue links with arrows indicate feed lines, while red links are links in which water flows in both direction during the simulation. Thickness of the lines corresponds to their diameter, while diameters are also written on the middle of each link.

Number of required and achieved number of feed lines for each DMA is also plotted on the graph, together with the notation "OK" or "NOT OK", indicating satisfaction of security of water supply design criteria.

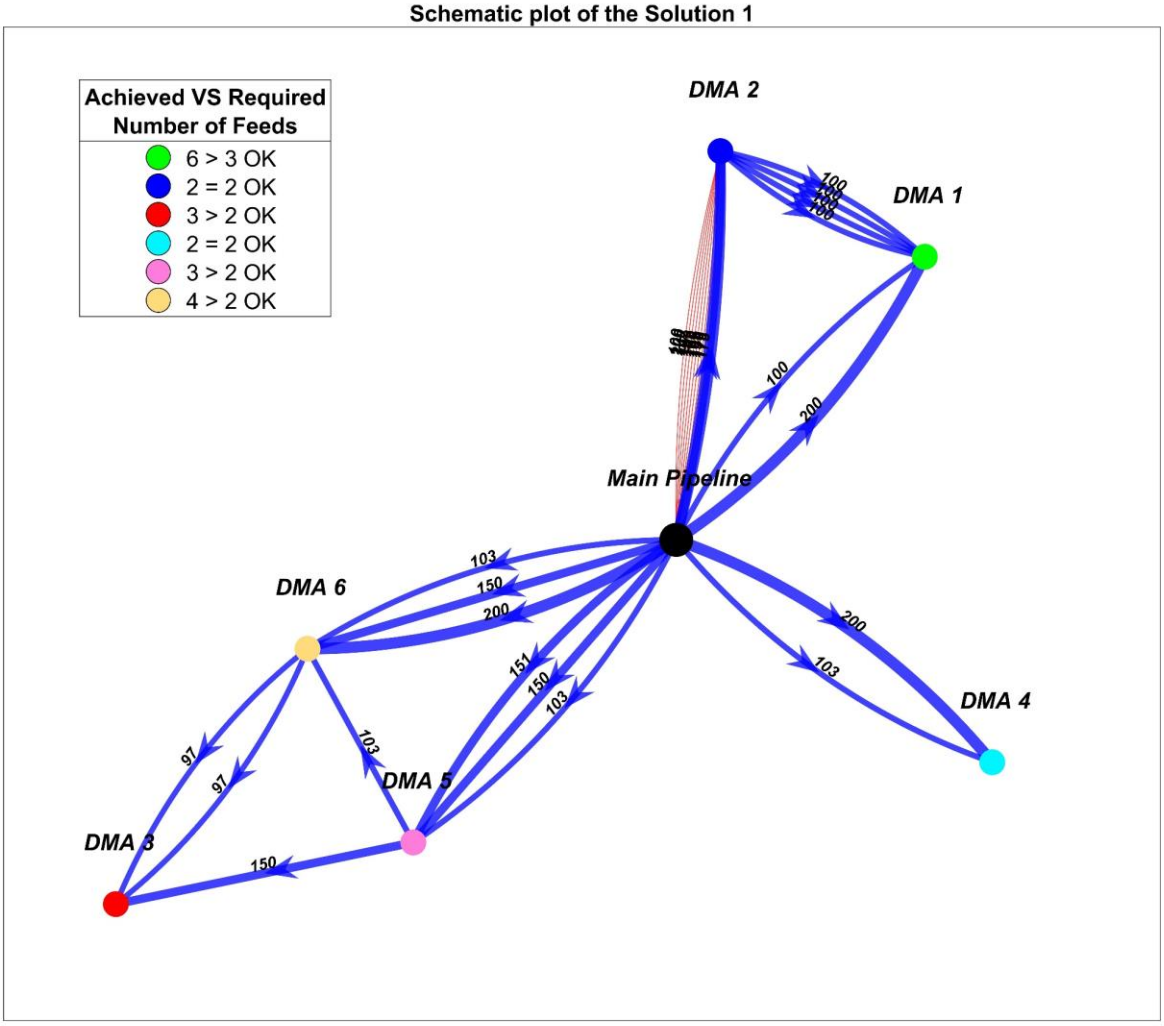

Figure 9. Graphical representation of connectivity between DMAs – plotted for every solution

This graphical representation gives insight to the DMAs' connectivity and directions of water flow. It should help user in selection process between different solutions, instead of him relaying only on the cost of the solution as selection criteria.

2.6 IMPLEMENTATION OF ALGORITHM

Presented methodology is implemented as per Figure 10. Implementation is carried out in Matlab programming environment as central processor with some parts of the methodology developed in different programing environments for computational efficiency. The 2nd Stage of the algorithm

(Network clustering algorithm) is written in C++ programing language to ensure high computational efficiency. It is compiled as a dynamic link library (DLL) that can be used externally to perform clustering. For hydraulic simulations (in stages 1 and 3) EPANET DLL is used, as it is a reference hydraulic solver and can be easily called from Matlab.

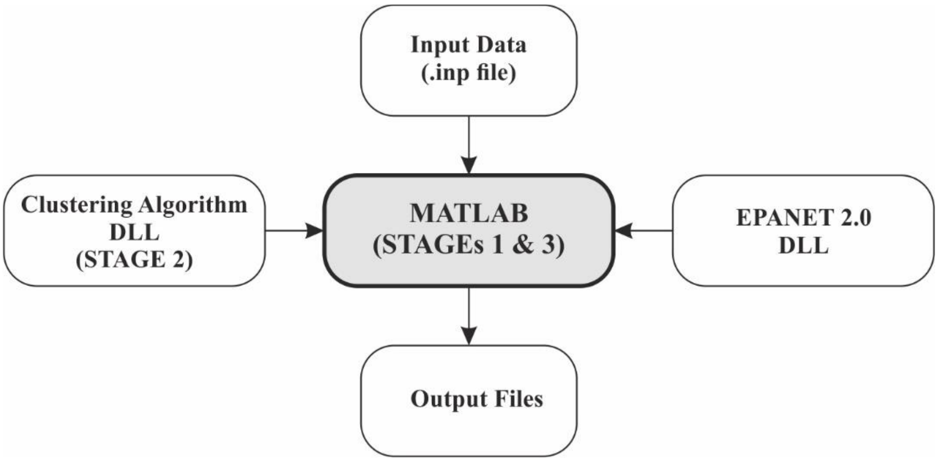

Figure 10. DeNSE method implementation chart

2.7 OUTPUT FILES AND VISUALIZATION OF RESULTS FOR THE SELECTED SOLUTION

After the user selects preferable solution, number of graphical results (graphs) are produced showing characteristics of adopted solution in detail.

Following characteristics of each DMA are shown in these graphs:

1. Average consumption per DMA,
2. Number of connections per DMA,
3. Comparison of requested and achieved number of direct feeds to DMA,
4. Total length of pipeline per DMA,
5. Comparison of average pressures in DMAs, before and after sectorization,
6. Number of flow meters required to define each DMA,
7. Implementation cost of setting up each DMA (individually).

First four characteristics listed above are plotted in 4-in1 graph shown in Figure 11 and the last two in the Figure 12. Additionally, user can plot the selected solution in order to visually inspect positions of flow meters and valves.

It should be noted that the whole procedure is computationally very efficient, as it will be shown later in the results section, so the user can easily investigate thoroughly more than one solution.

Finally, when preferable solution is definitely selected following output files can be produced:

1. **Several Google Earth files (.kml files)** showing the whole network and position of flow meters and valves, which is convenient as it is shown on satellite background,

2. **EPANET file (.inp file)** in which all required interventions are implemented (i.e. all pipes that should be equipped with valves are closed in this model). This means that user can produce other results on his own and further investigate hydraulic model of the solution.

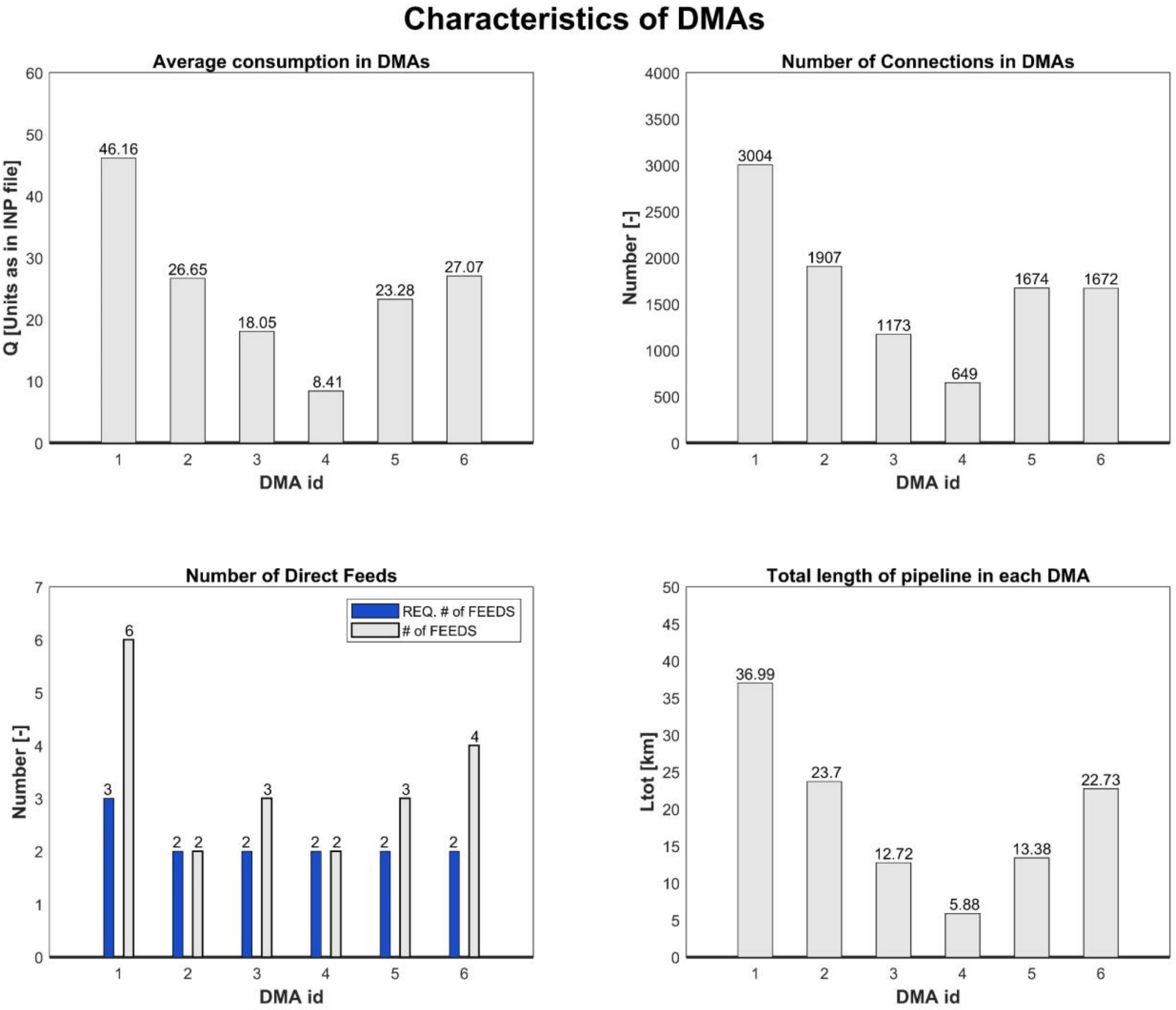

Figure 11. Exemplary plot of characteristics of each DMA in the selected preferable sectorization solution

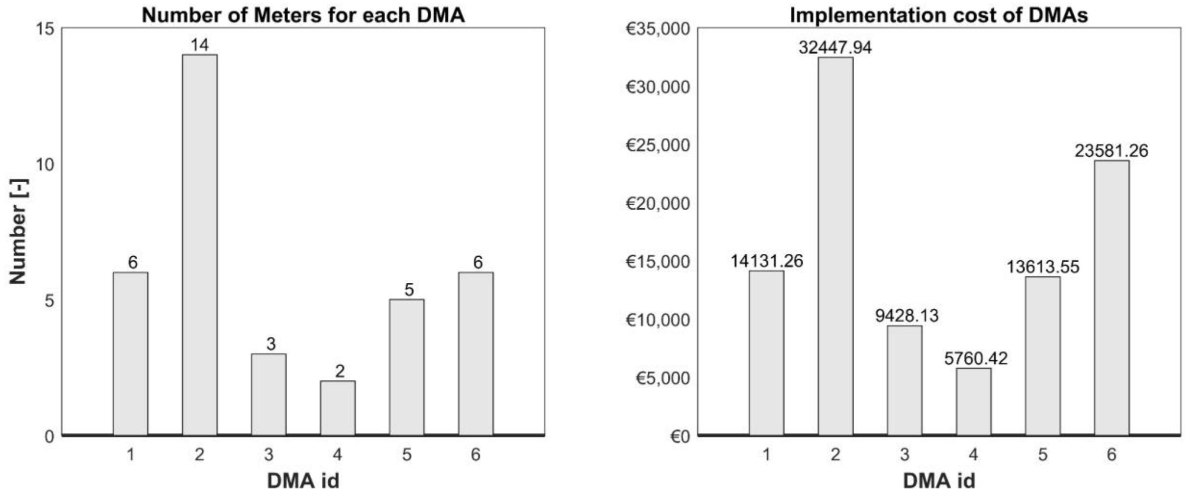

Figure 12. Exemplary plot of flow meters number and implementation cost for each DMA in the selected preferable sectorization solution

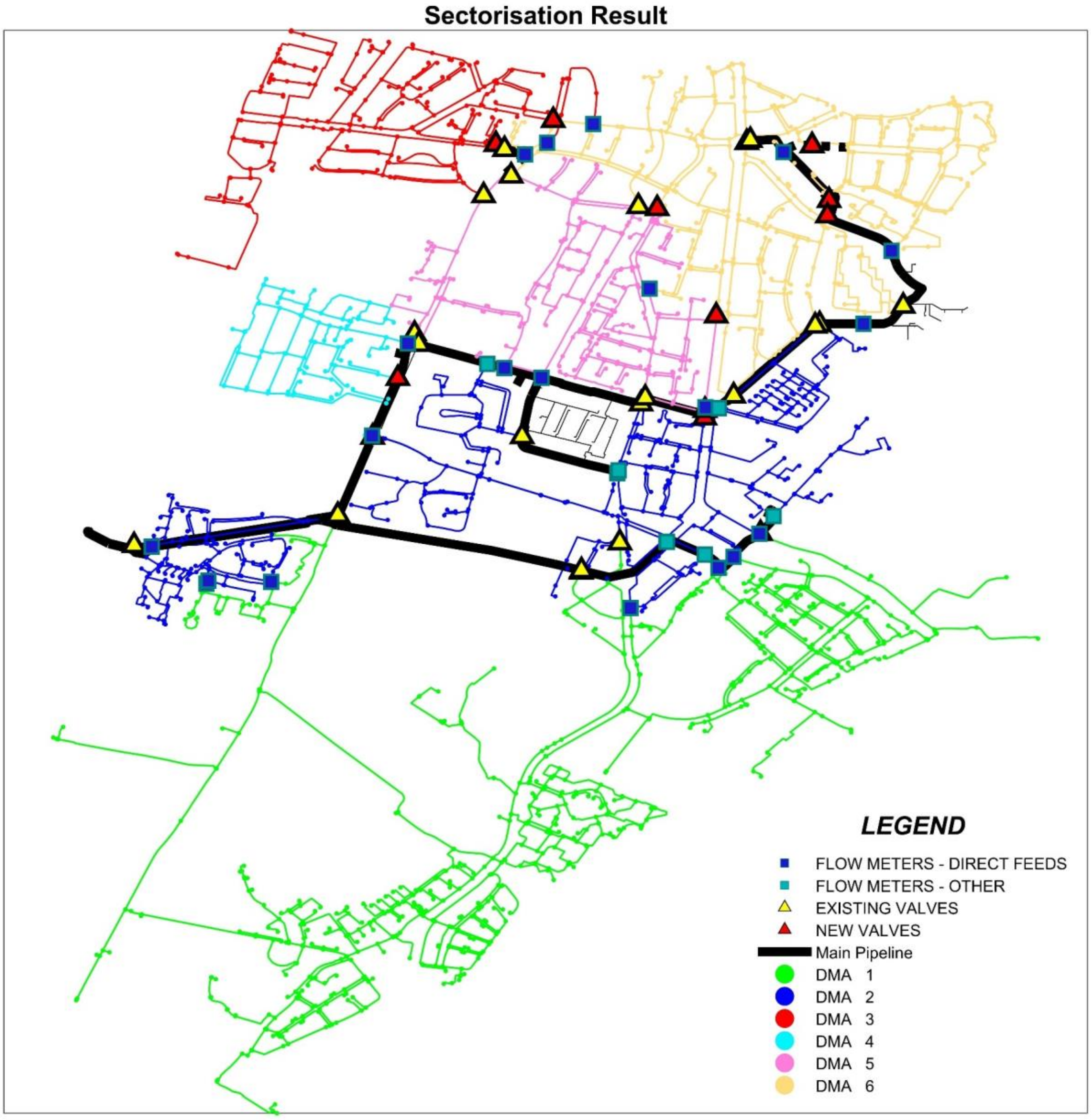

Figure 13. Exemplary plot of the network with the positions of flow meters and valves for the selected preferable sectorization solution

Google Earth (GE) files are completely compatible with the previous plot which is done without background (see Figure 14).

Figure 15 shows comparison of ordinary and Google Earth (GE) plot of one smaller area, and at each pin location on GE plot unique identification of the pipe/valve (ID) is shown. Objects are marked as **EPAValveID** or **EPAPipeID**. This ID corresponds to the name od the pipe in EPANET file, so that the user can easily find these objects in the EPANET file (inp) file, which is also created for the selected solution.

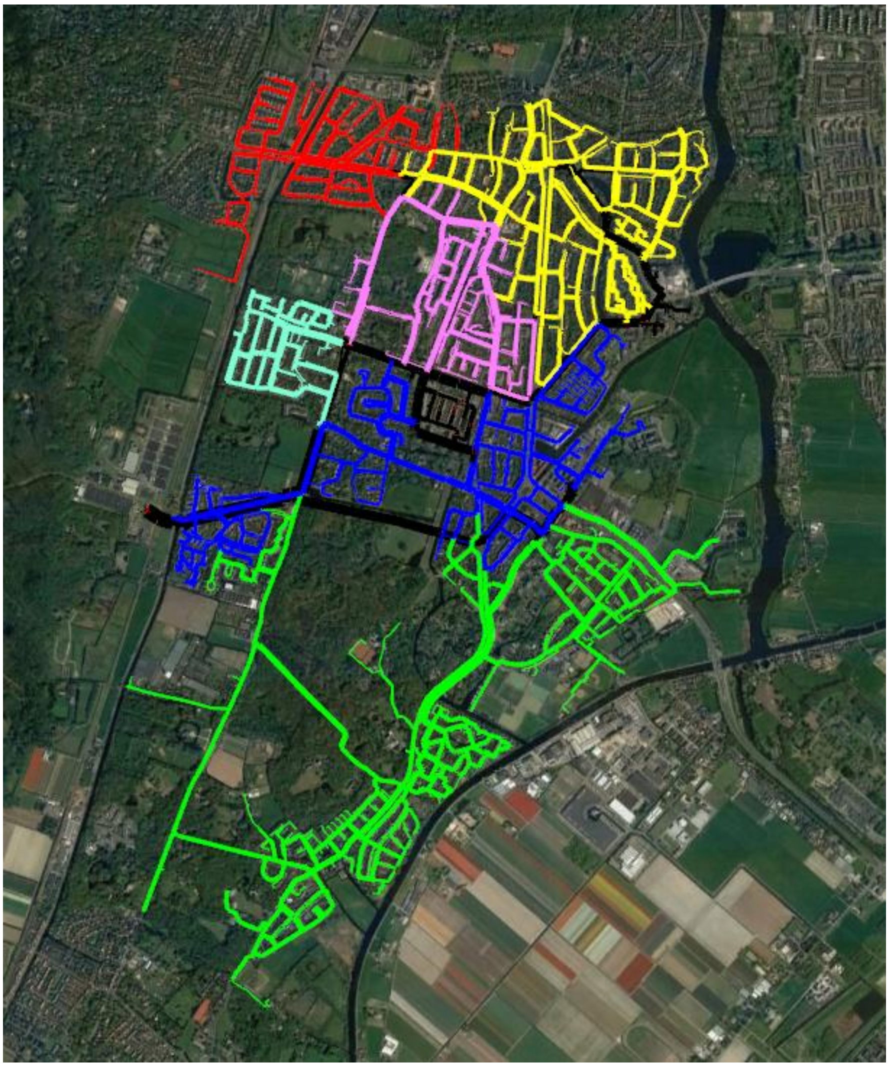

Figure 14. Exemplary Google Earth plot of the network with the positions of flow meters and valves for the selected preferable sectorization solution

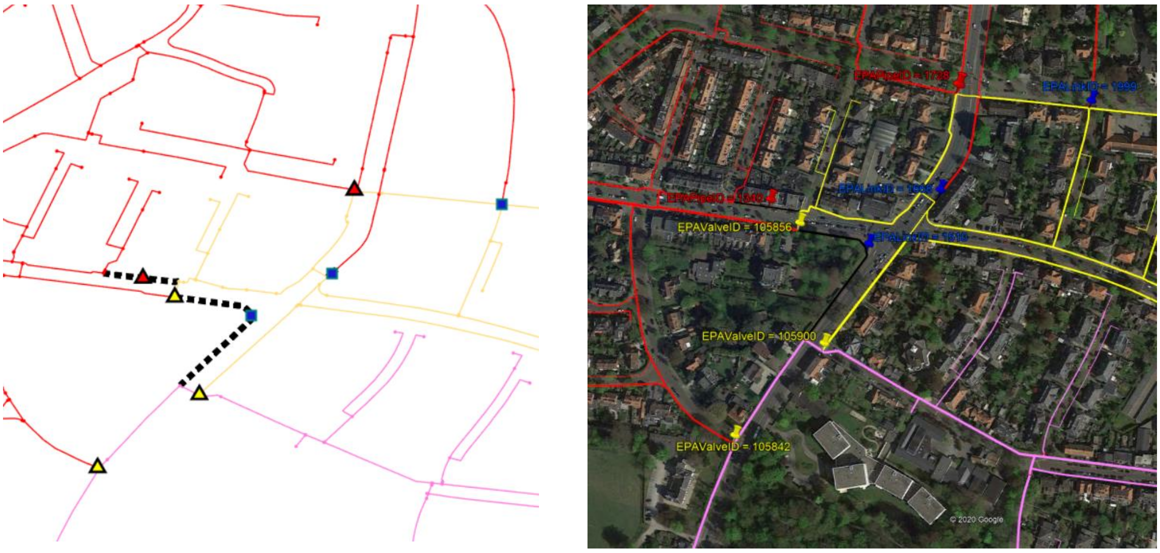

Figure 15. Comparison of ordinary and Google Earth plot of the same area

2.8 ADOPTED LOCAL DESIGN CRITERIA

Main design criteria used in this research are mainly adopted from the document **"Design Criteria Drinking Water Distribution Networks"** supplied by the Waternet personnel. The document serves as a guideline for the design of water distribution networks that are operated by the Waternet company. This applies to the construction of new network as well as replacement and rehabilitation works on the existing ones. Also, some data listed below are not contained in the document but were supplied directly to researchers (e.g. cost of flow meters and valves).

Following design criteria, relevant for the previously described methodology, are adopted and used in this research:

1. **Number of connections** – For the calculation of number of connections per node, based on the total nodal demand, following values are used:
    a. Average lot occupation – **2,1 p/conn** (person per connection),
    b. Use per person – **134 L/p/day** (liters per person daily).
2. **Security of water supply** – Expressed through the number of feeds to an area (DMA) based on the number of connections in that area. In normal operations an area must be fed in the following manner:

Table 2. Required number of feeds for an area to secure water supply

Number of delivery points	Minimum number of feeds per area
Up to 200 connections	1 feed
200 to 2000 connections	2 feeds
More than 2000 connections	3 feeds

3. **Minimum network pressure** – Minimum pressure for the network is **p_{min} = 2.0 bar**, which is 0.5 bar more than the requirement by the law.
4. **Maximum network pressure** – Maximum network pressure is **p_{max} = 6.0 bar**, which corresponds to the design requirement for the pipe material (1.5 x max operating pressure).
5. **Velocities** – Minimum and maximum velocities are defined for tertiary network which has distribution function (diameters are usually less than 125 mm). Daily minimum velocity for self-cleaning is **0.4 m/s** and daily maximum velocity is **1.0 m/s**.
6. **Price of the equipment** – local prices of the valves and flow meters were supplied directly by the Waternet asset management department. Prices include all-in installation (i.e. includes purchase, transport and labor costs). Prices are given in the following table:

Table 3. Installation prices for valves and flow meters

Diameter (mm)	Valve placement cost (all-in)	Flow meter placement costs (all-in)
75	€ 1,575	€ 2,093
90	€ 1,575	€ 2,421
100	€ 2,260	€ 2,690
110	€ 1,785	€ 3,412
150	€ 2,850	€ 3,587
160	€ 2,124	€ 3,635
200	€ 3,119	€ 4,200
315	€ 3,975	€ 6,899
400	€ 4,335	€ 8,761

3 CASE STUDY NETWORK

Water Distribution Network (WDN) operated by Waternet company currently has 6 (six) areas that can be considered DMAs, as they can be hydraulically isolated from the rest of the network. Five (5) of those areas are part of the Amsterdam's WDN and within the city of Amsterdam region, with only one (1) being separate WDN supplying Heemstede municipality located west from the city of Amsterdam. These six (6) areas are the obvious candidates for DMAs as they can be easily spotted as the distinct areas just by looking at the topological map of the Amsterdam WDN. They also significantly vary in size in terms of number of connections:

1. Diemen Noord (~2 900 connections),
2. Prinseneiland (~325 connections),
3. Amstelveen (~35 000 connections),
4. Ouderkerk (~3 500 connections),
5. Heemstede (~10 700 connections),
6. Amsterdam Noord (~42 000 connections).

The rest of the Amsterdam WDN has roughly around 500 000 connections. These connections are mostly concentrated in the central city area, and generally the hardest task is to define DMAs within this densely looped central part of the WDN. For investigation in this research case study network of Heemstede is chosen and shown in Figure 16.

Figure 16. Heemstede network used as a case study in this research

DeNSE methodology, presented in this research, will be used to define a more refined definition of DMAs within this WDN.

Heemstede network is chosen because it is a separate WDN, hydraulically independent with no need for special definition of boundary conditions. Also, it is relatively moderately sized, and it can serve as a model for further development of the methodology.

Table 4. Characteristics of Heemstede network

pipes	nodes	demand nodes	Reservoirs (sources)	tanks	pumping stations	valves	connections	Total demand
3204	3997	2046	1	1	1	989	10741	161.3 m^3/h

Hydraulic model of Heemstede network was made available in the form of EPANET input file (inp file) by Waternet company personnel (Figure 17). Heemstede network model included exact number of house connections for each node.

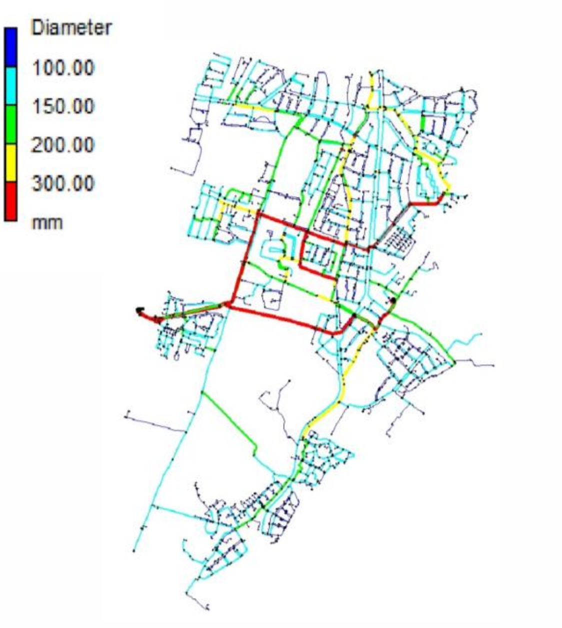

Figure 17. EPANET input file of Heemstede network used in this research

It should be noted that it is not confirmed whether the supplied EPANET model is calibrated or not. Considering that, all results presented in this report should be confirmed and validated once the verified calibrated model is available.

4 RESULTS

This Chapter presents results for the case study network of Heemstede analyzed in this research. Presentation of the results in this Chapter follows the description of the methodology given in the Chapter 2. Input data is recapped first, followed by the analysis of the network clustering algorithm results. Selection of preferable clustering solution is made and elaborated, followed by the ranking of the optimized solutions, finally adopting the preferable solution for DMAs.

4.1 INPUT DATA

Input data used for the sectorization of Heemstede network are as follows:

1. Minimum ($\mathbf{n_c^{min}}$) and maximum ($\mathbf{n_c^{max}}$) number of property connections per DMA are $\mathbf{n_c^{min}}$=500 and $\mathbf{n_c^{max}}$=5000. These are general recommendations about the size of DMAs, given in the relevant guidelines.
2. Total number of connections in the network is $\mathbf{n_c}$**=10741**, taken from the mathematical model.
3. Transmission main threshold diameter used in preprocessing to identify main pipeline is $\mathbf{D_{MAIN}}$ **= 225 mm**, which yielded transmission main shown in Figure 18.
4. Minimum required and maximum allowed pressures in the network are adopted from the local network design criteria (see Section 2.8) - $\mathbf{p_{MIN}}$ **= 2.0 bar** and $\mathbf{p_{MAX}}$ **= 6.0 bar**
5. Desired number of alternate solutions for the definition of DMAs, that will be ranked, is **N = 5**.
6. Parameters used in GA optimization are – number of generations **ng = 35** and population size for each generation **ps = 30**.

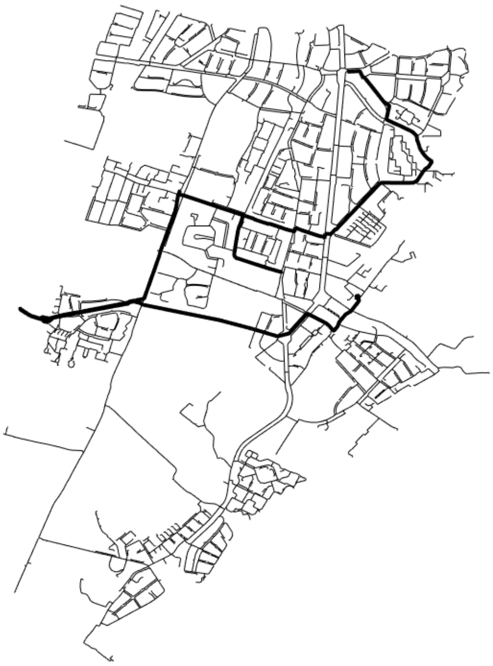

Figure 18. Identified transmission main for Heemstede network

4.2 NETWORK CLUSTERING ALGORITHM RESULTS

Figure 19 shows the evolution of network uniformity index (U) during the aggregation process (network uniformity index is explained in Section 2.3, see equation (1)). Figure shows values of U for the last 12 aggregation steps, as it reaches maximum value and then decreases as the aggregation process terminates.

Maximum value of network uniformity index is U=0.59 and it corresponds to sectorization of the network into six (6) clusters. Obviously, this emerges as preferred clustering solution that will be subjected to the optimization in which position of flow meters and valves will be determined, thus defining DMAs.

However, user can (and should) review plots for several clustering solutions in the vicinity of the one with maximum uniformity value. For example, in this case maximum uniformity corresponds to sectorization into 6 clusters, but the sectorization solutions into 7 (one step before) and 5 (one step after) have similar, slightly lower, values of uniformity index. User should review them as well, as a solution with lower uniformity value can be better than the solution with maximum uniformity value considering local network topology and other constraints and criteria.

Figure 20 shows clustering solutions corresponding to the last 6 steps of aggregation process. Each solution plot in this figure shows clusters marked with different colors with the legend showing number of connections per each cluster. They will be briefly summarized here.

Considering the given input data regarding the minimum and maximum number of connections in the DMA ($\mathbf{n_c^{min}}$=500 and $\mathbf{n_c^{max}}$=5000), preferred number of connections in a DMA is $\mathbf{n_c^{pref}}$ =2750. Achieving this exact number for each cluster in impossible. The whole idea behind the uniformity index is to find the solution in which all clusters are preferably sized in the range $\mathbf{n_c^{min}} - \mathbf{n_c^{max}}$, and as much as possible sized equally and close to $\mathbf{n_c^{pref}}$.

Solution with 8 clusters (Figure 20 a) has 2 extremely small clusters with 194 and 289 connections which are below the $\mathbf{n_c^{min}}$ (cluster 7 and cluster 4), hence the lower value of uniformity index compared to the maximum one. In the next step (solution with 7 clusters - Figure 20 b) one of these small clusters (cluster 7) is aggregated into the larger cluster 1 which improves uniformity of the clusters in size and thus the overall network uniformity index. In the following step the other small cluster (cluster 4) is also aggregated into the cluster 1 (Figure 20 c), further improving uniformity index value which reaches maximum in this step (**U = 0.59**). In this solution all clusters are within the required size range $\mathbf{n_c^{min}} - \mathbf{n_c^{max}}$, with one cluster being slightly above $\mathbf{n_c^{pref}}$ (cluster 1 with 3004 connections).

Two relatively large clusters are merged in the next step (Figure 20 d) – clusters 3 and 6 – which slightly disrupts sizing equality of the clusters and lowers the value of uniformity index. This is because now there are two clusters larger than $\mathbf{n_c^{pref}}$ (cluster 1 with 3004 and cluster 3 with 2845 connections). As the aggregation procedure continues this becomes more emphasized and in the next step (Figure 20 e) clusters 3 and 5 are merged to yield new cluster with 4519 connections (cluster 3). Finally in the last step (Figure 20 f) there are only 3 clusters, with two of them being very close to the $\mathbf{n_c^{max}}$ size and almost double the $\mathbf{n_c^{pref}}$, hence the lowest value of uniformity value. Further aggregation is not possible as these three clusters are connected only to the main pipeline and are not connected between them, thus cannot be aggregated further.

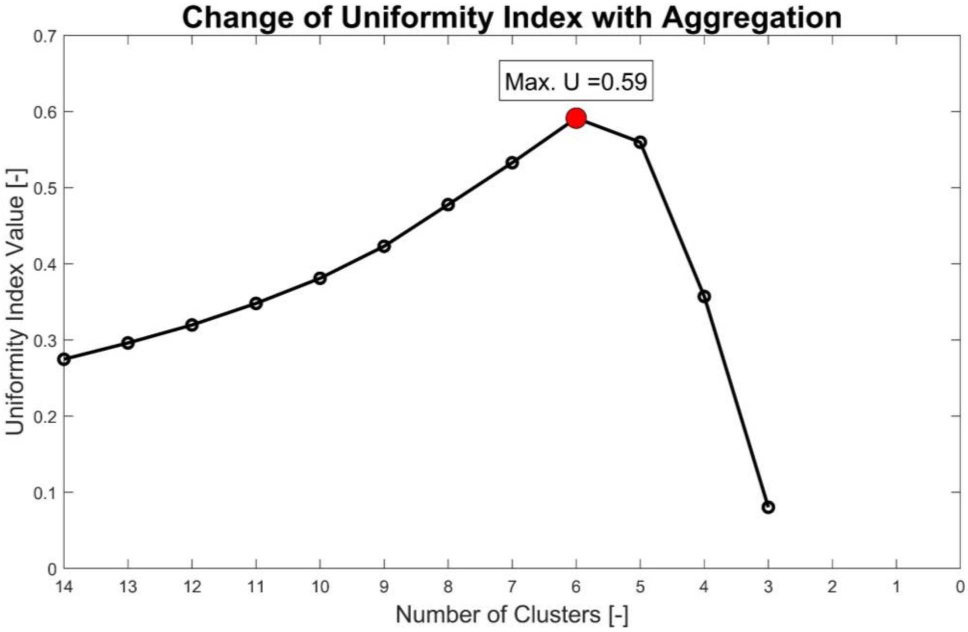

Figure 19. Evolution of network uniformity index (U) during the aggregation process (Heemstede)

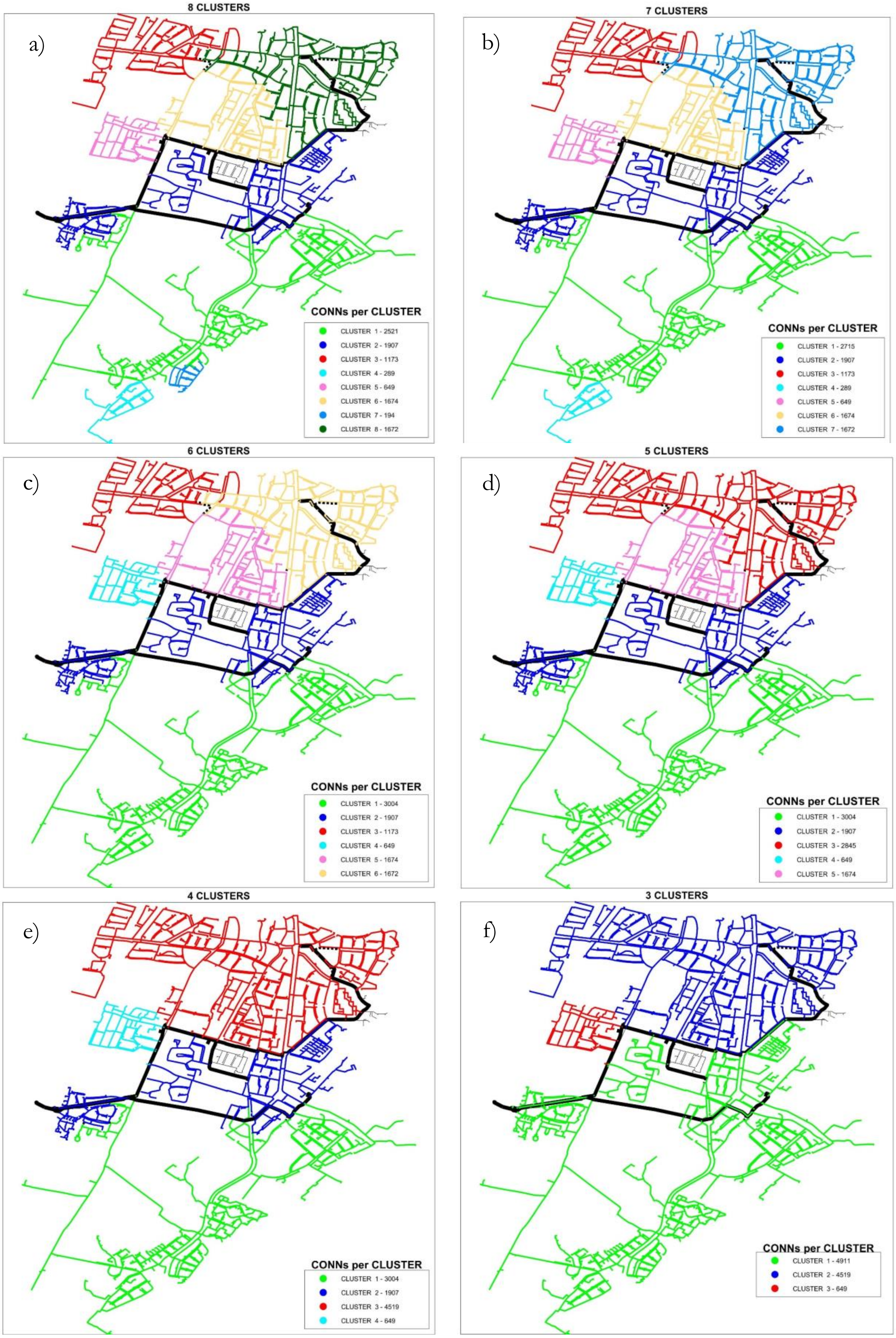

Figure 20. Heemstede clustering solutions corresponding to the last 6 steps of aggregation process

4.3 RANKING OF OPTIMIZED SOLUTIONS

As it was said before, it is up to the user to select preferable clustering solution, which is then subjected to the optimization process. In the previous section number of clustering solutions, were plotted (Figure 20) and described. The solution which has the highest value of uniformity index is selected as preferable in this analysis. That is the solution with 6 clusters (Figure 20c).

Given the desired number of alternate solutions for the definition of DMAs (input data, N=5), optimization algorithm is ran 5 times to yield 5 alternate dispositions of flow meters and valves on clusters boundary links. Ranking of these 5 solutions is done based on the results presented below, and indicators introduced in Section 2.5 ("Indicators for ranking of solutions").

First, let's take a look at the 4 indicators shown in Figure 22: a) average network pressure, b) relative change of network pressure, c) implementation cost and e) number of flow meters and new valves to be installed.

Based on the average network pressure plot, it is safe to say that all solutions do not change much pressure distribution in the network, which was one of the goals (i.e. not to disrupt original network state). In terms of relative change of network pressure solution SOL-2 performs best, followed by the solutions SOL-1 and SOL-4 which perform similarly, while solutions SOL-3 and SOL-5 are lagging behind.

Implementation cost and number of flow meters/new valves are proportional to some extent, as larger number of devices implies higher price. However, relation is not linear as prices of valves and flow meters differ significantly depending on the diameter (see Table 3). Regarding the installation cost criteria, again solution SOL-2 is preferred and is compared to the solutions SOL-4 and SOL-5, while solutions SOL-3 and SOL-1 are the most expensive. Same order of solutions can be made observing the number of flow meters and new valves plot, as solutions SOL-2, SOL-4 and SOL-4 have the least number of new valves to be installed.

Figure 21 shows absolute minimum pressure observed in the demand nodes of the network for each solution, compared to that same pressure in the original network layout (blue line). All solutions are above minimum required pressure of **p_{MIN} = 2.0 bar,** solutions SOL-2 and SOL-1 being closest to the original minimum pressure.

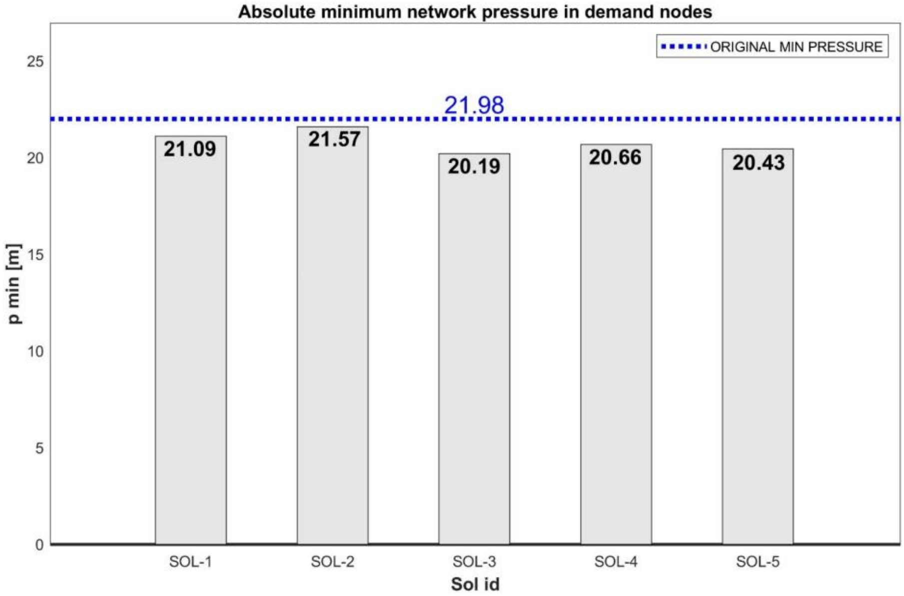

Figure 21. Absolute minimum pressure in demand nodes

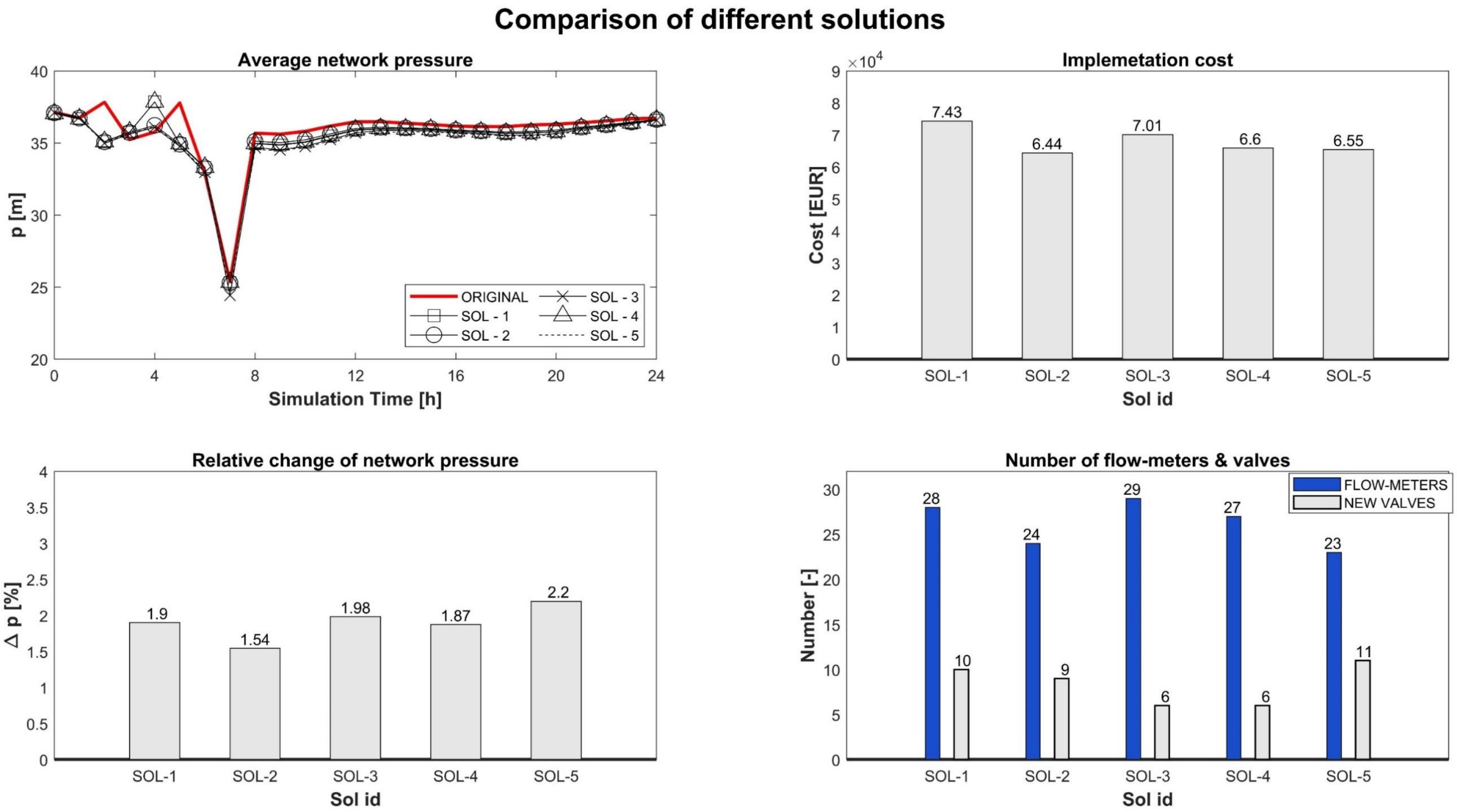

Figure 22. Ranking indicators for Heemstede network

4.4 ADOPTING PREFERABLE CLUSTERING SOLUTION

Results presented in the previous section showed that all solutions are feasible in terms of pressure constraints. Considering discussion made there, solution SOL-2 seem to be the one to select as preferable, as it ranks top in all ranking criteria. However, user should also consider connectivity between DMAs and exact position of flow meters and valves in the network.

Following 5 figures (Figure 23 – Figure 27) show DMAs connectivity for each of the 5 solutions. Legend shows number of achieved feed lines to each DMA, compared to the required number. It is important to notice that in all solutions, achieved number of feed lines is greater than required one (it is marked with OK), meaning that all solutions are feasible from this point of view.

Figures show that there is a difference in diameters of the connecting links and their number as well. For example, in solution SOL-1 DMA 1 is connected to the main pipeline with 2 links (diameters 100 and 200 mm) and in SOL-2 there is only one link with the diameter of 200 mm. Also, in solution SOL-1 there is a connecting link between DMA 5 and DMA 6 and in solution SOL-2 these DMAs are not connected. In SOL-1 feed lines to DMA 6 (4 in total) come from the main pipeline (3) and from the DMA 5 (1). In SOL-2 feed lines to DMA 6 (3 in total) come from the main pipeline. They are also smaller in diameter than in the case of SOL-1, which means that DMA 6 is connected to the main pipeline with different links at different locations.

To gain full insight in DMAs connectivity and layout, both

a) connectivity plots and

b) network plots with exact positions of flow meters and valves for all different solutions,

must be investigated. This will help experts familiar with all network specifics in making the selection of preferable solution and decision-making process easier.

In this report, solution SOL-4 is selected as preferable one after thorough investigation of all available results – ranking indicators, connectivity plots and network layout plots.

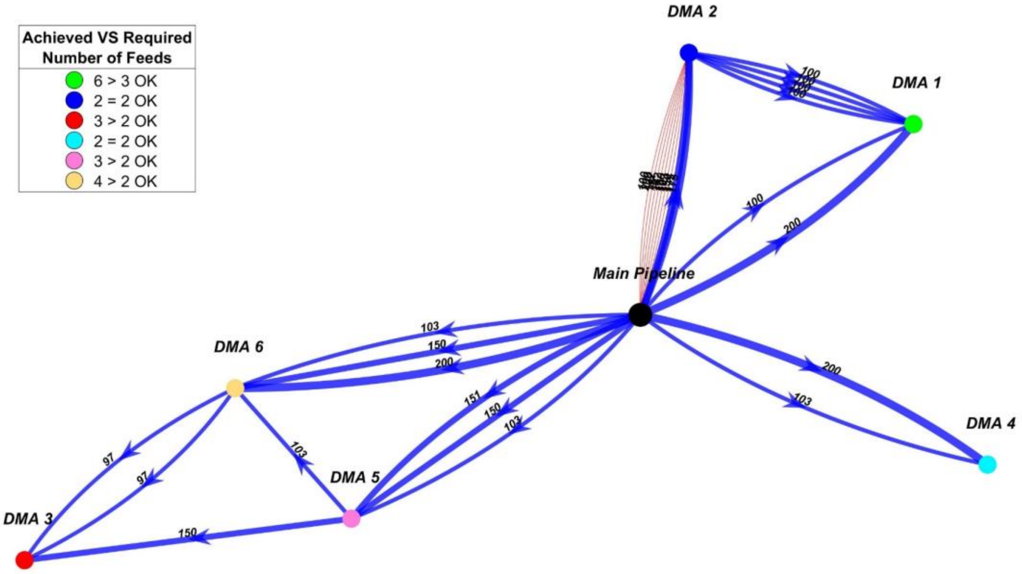

Figure 23. Schematic plot of DMAs connectivity – Heemstede solution SOL-1

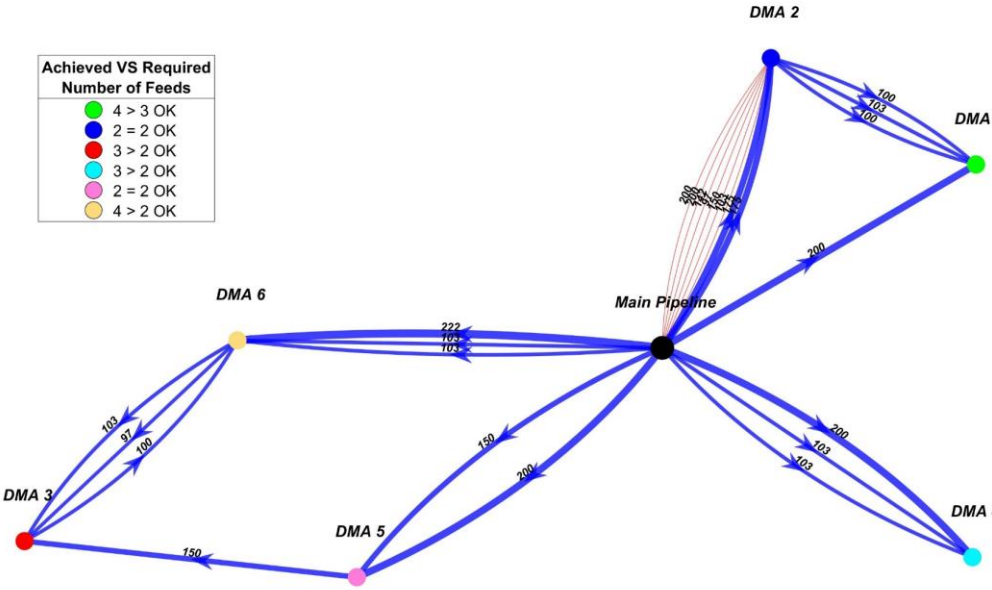

Figure 24. Schematic plot of DMAs connectivity – Heemstede solution SOL-2

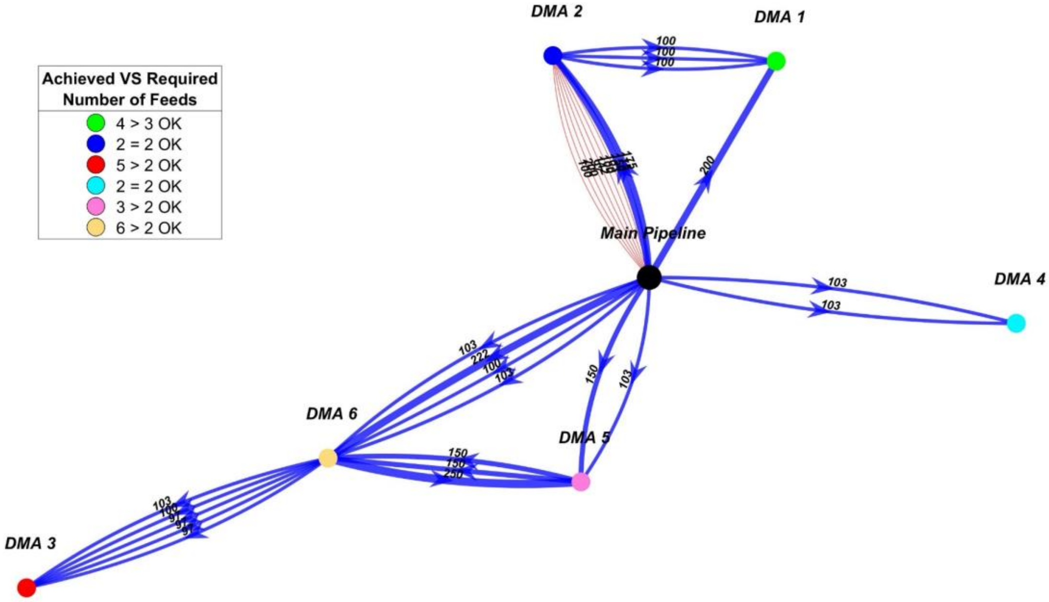

Figure 25. Schematic plot of DMAs connectivity – Heemstede solution SOL-3

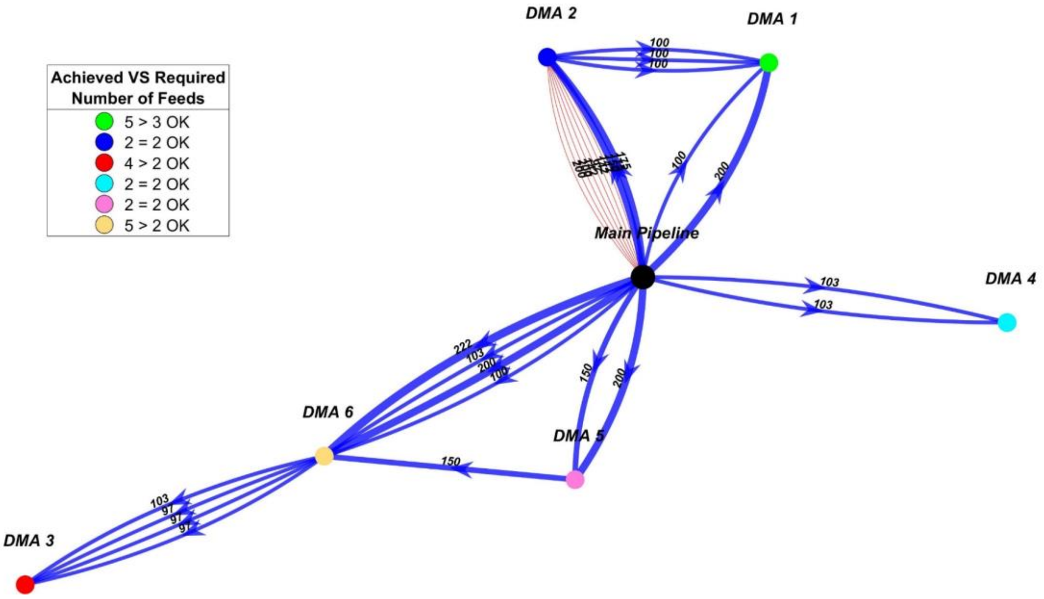

Figure 26. Schematic plot of DMAs connectivity – Heemstede solution SOL-4

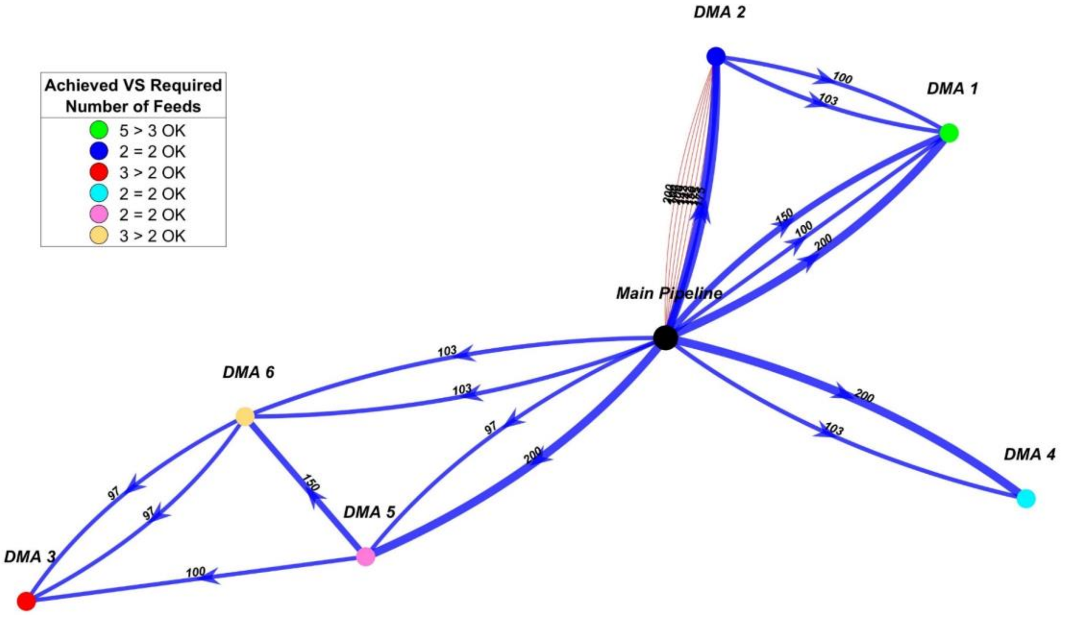

Figure 27. Schematic plot of DMAs connectivity – Heemstede solution SOL-5

4.5 ADOPTED SECTORIZATION SOLUTION

Figure 28 shows network plot for the adopted solution (SOL-4). Figure shows exact positions of flow meters and valves. Flow meters are shown as

a. flow meters on direct feeds (marked with dark blue square markers) - i.e. flow meters on pipes where water does not change its flow direction during simulation) and
b. flow meters on other pipes (marked with light blue square markers) – i.e. flow meters on pipes that do change flow direction.

There are two types of valves:

c. existing valves (marked with yellow triangle markers) – i.e. a valve already exists on pipe marked for closure, so there is no need to install a new one and
d. new valves to be installed (marked with red triangle markers).

Figure 28 should be looked at in conjunction with schematic plot of DMAs connectivity of considered solution (in this case see Figure 26 for solution 4).

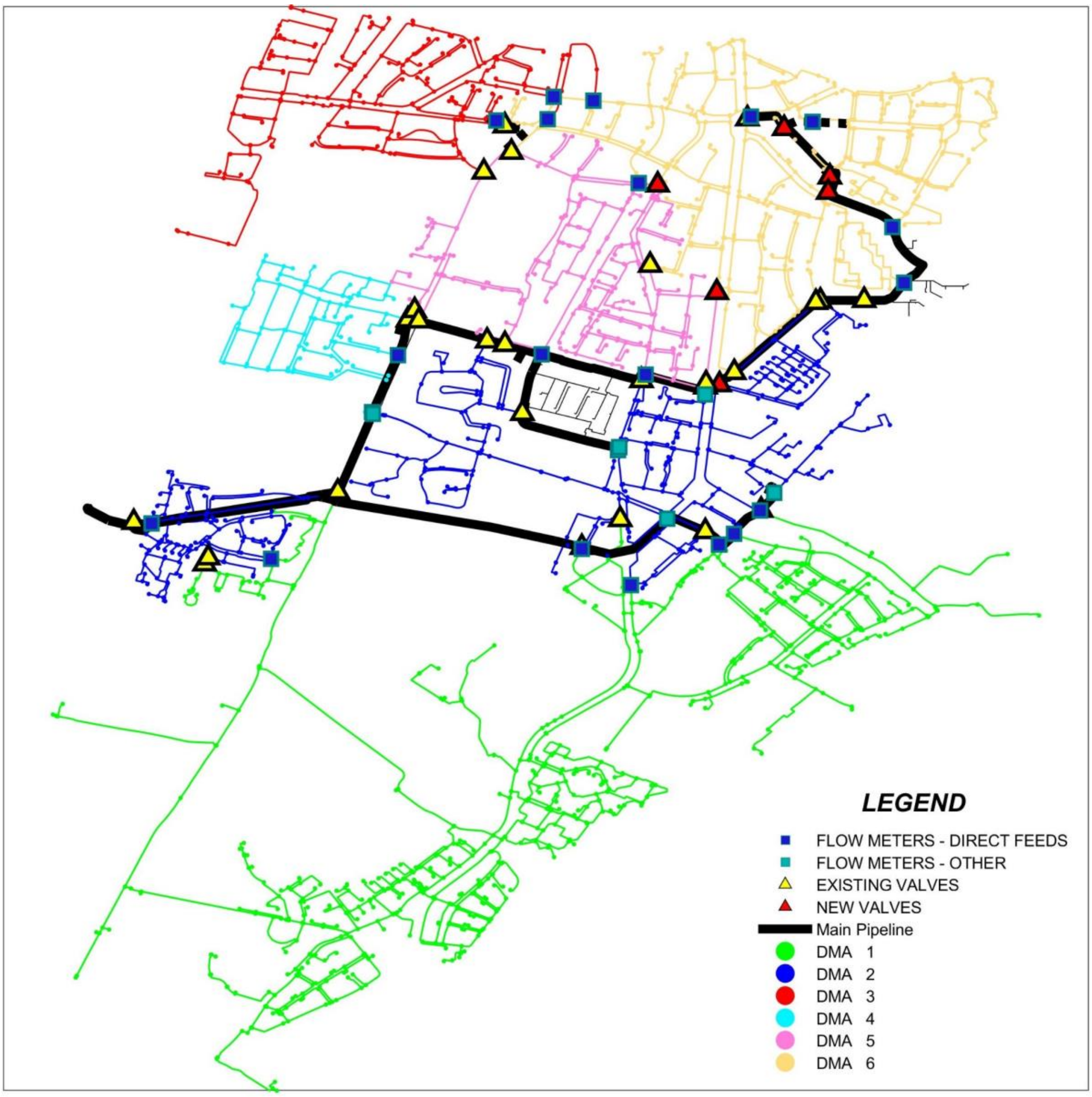

Figure 28. Adopted solution for Heemstee network –plot with positions of flow meters and valves

Figure 29 shows generated Google Earth plot of adopted sectorization solution. One kml file is generated for each type of object of interest – DMAs, position of direct feeds, position of other flow meters, position of existing valves and position of new valves. This allows user to investigate spatial positioning of devices and assess possibility for field implementation.

From the menu on the left window user can select which object layer will be displayed. Object IDs that are displayed in Google Earth files correspond to those in EPANET output file.

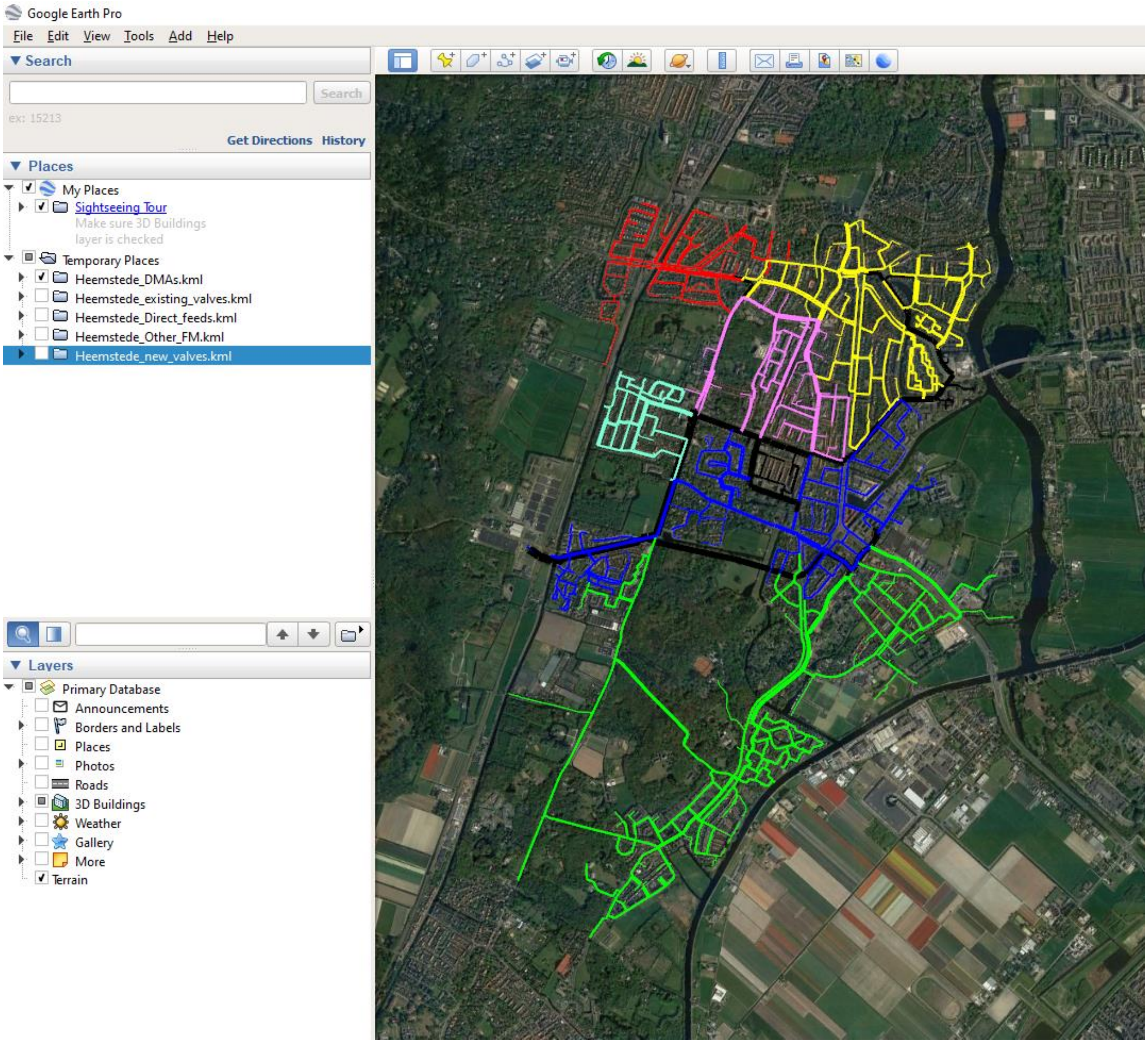

Figure 29. Adopted solution for Heemstede network – generated Google Earth plot (kml file)

Figure 30 shows different characteristics of DMAs in adopted solution. For each DMA figure shows a) average consumption, b) number of connections, c) comparison of achieved and required direct feeds and d) total length of pipeline. DMAs 2, 3, 5 and 6 are fairly equally sized in terms of number of connections (which was one sestorization goal). DMA 4 is significantly smaller, but that is due to the network layout – i.e. this DMA is connected only to the main and it cannot be aggregated with any other part of the network. DMA 1 is slightly larger, but still within the limits of $\mathbf{n_c^{min}} - \mathbf{n_c^{max}}$, as are all the others.

Figure 31 shows comparison of average pressure in DMAs during the simulation, before and after sectorization. Pressure remains mainly unchanged in DMAs 1, 2, 4 and 5. It is slightly lowered in DMA 6 and a bit more in DMA 3. Maximum reduction of pressure in DMA 3 is around 2.5 m, while the pressure is well above 3.0 bar, hence it can be concluded that this will not affect supply security in this area.

Figure 32 shows number of flow meters and new valves that must be installed on boundaries of each DMA to completely define it. It also shows pricing for setting up each DMA individually. It must be said that this diagram shows pricing for initial setting of DMAs. In other words, sum of these pricing is not the total cost for setting up all DMAs. For example, DMA 4 is cheapest, and it can be set up first. It is not connected to any other DMAs (but only to the main) so it does not affect the cost of setting up other DMAs. Next, DMA 3 is second cheapest to set up, but it is connected to DMAs 5 and 6 (i.e. they share some flow meters and valves), so the pricing for these DMAs has to be recalculated (reduced). Total pricing for setting up DMAs in whole network is 66.000,00 EUR as Figure 22 shows (see cost for adopted solution SOL-4).

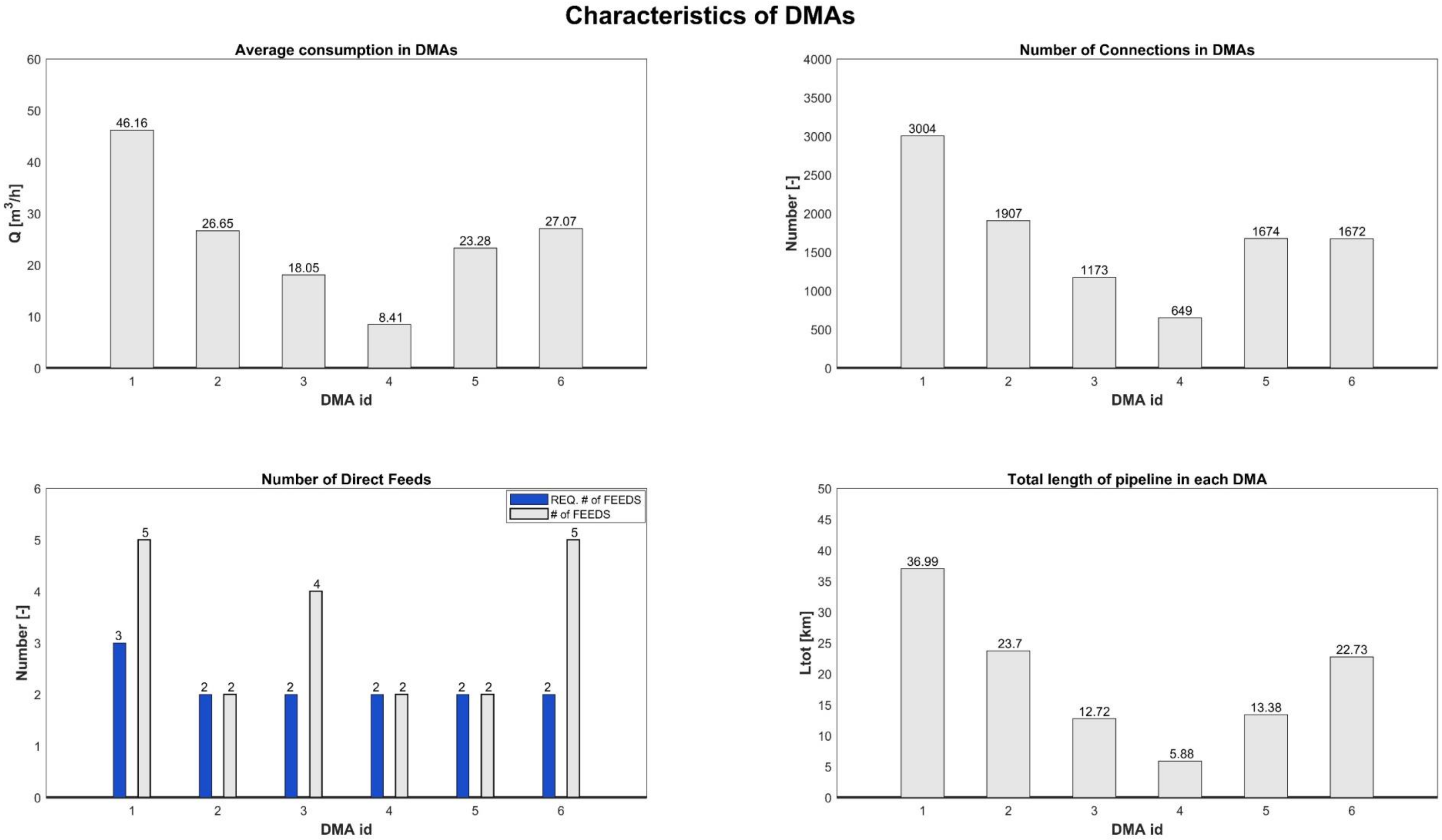

Figure 30. Characteristic of DMAs for the adopted solution (Heemstede)

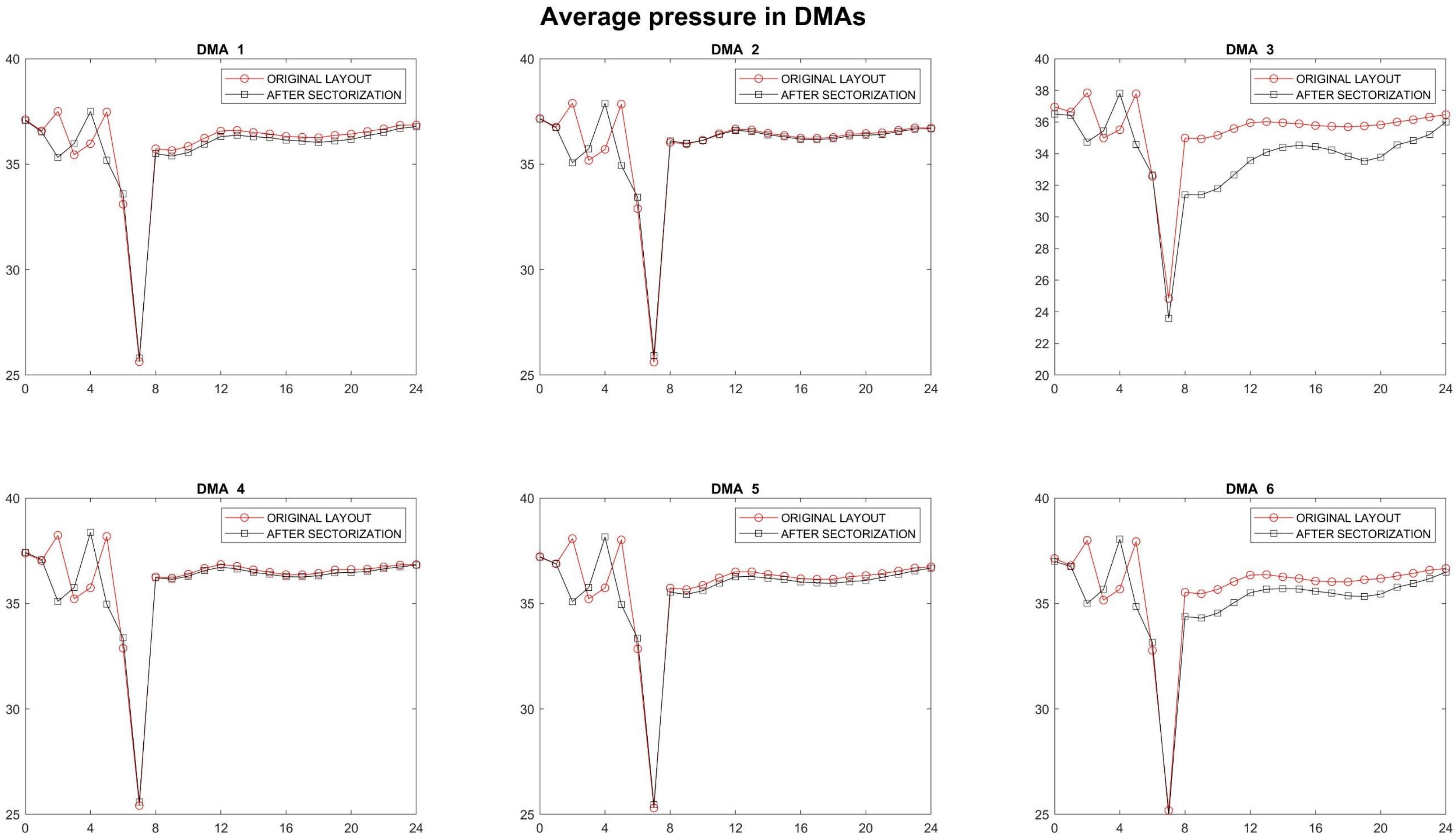

Figure 31. Comparison of average pressures in DMAs, before and after sectorization (Heemstede)

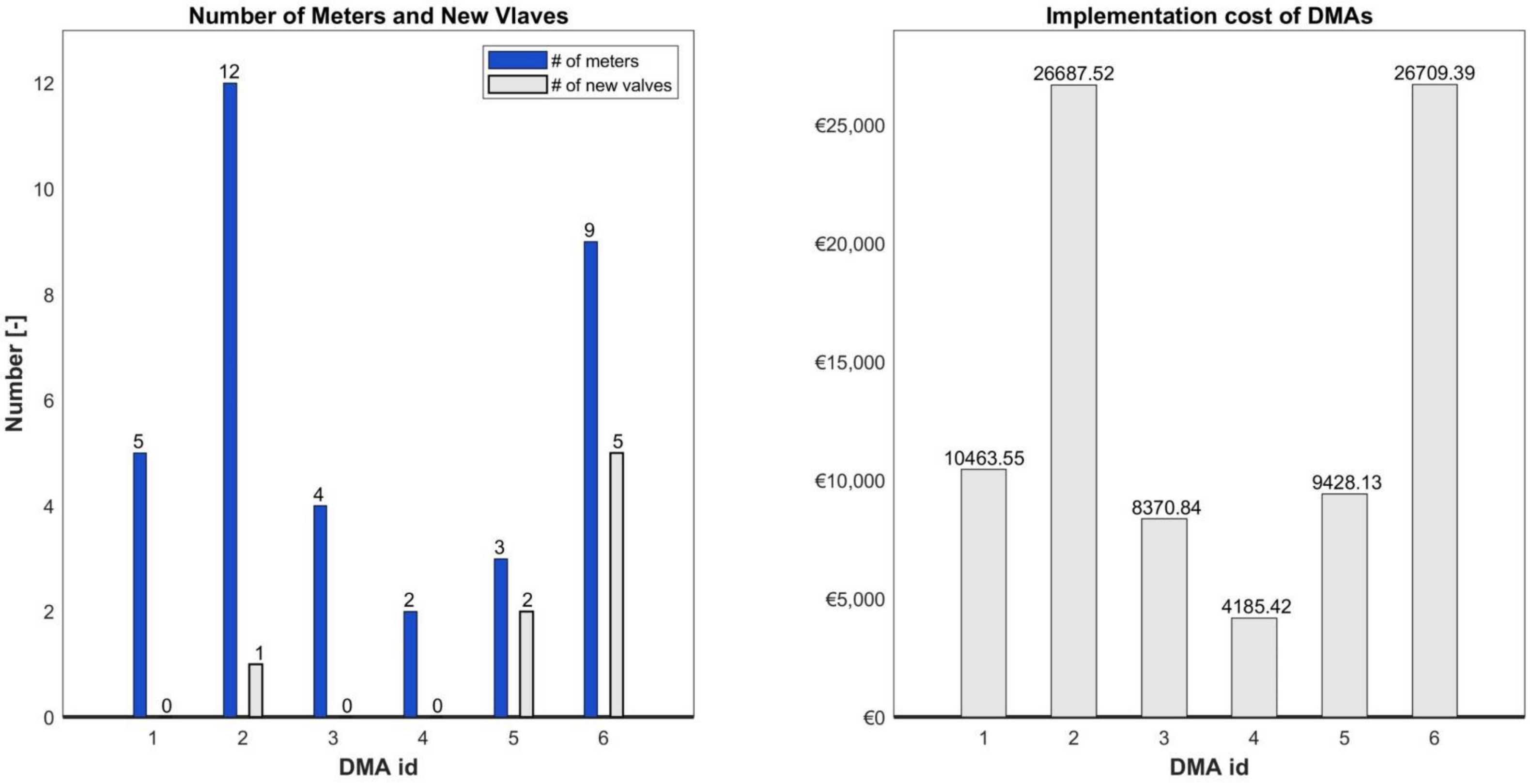

Figure 32. Number of meters for each DMA and its' implementation cost (Heemstede)

5 CONCLUDING REMARKS

This report investigated the problem of water distribution network (WDN) sectorization into the District Meter Areas (DMAs). Main goal of research was to develop sectorization method, test its abilities on several case study networks (i.e. parts of water distribution network of the city of Amsterdam) and assess its possible use in decision making process in the future.

Sectorization method of choice is Distribution Network Sectorization method (DeNSE), originally developed at the University of Belgrade-Faculty of Civil Engineering (Vasilic et al., 2020). 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 implemented in DeNSE methodology. Such set of design criteria is most appealing to many water utilities, especially ones commencing implementation of sectorization strategy in WDN.

DeNSE sectorization method is significantly upgraded during Horizon 2020 research project Wat-Qual as presented in this report. Main improvement made to DeNSE method is implementation of GA optimization procedure, employed to search broader set of potential sectorization solutions. Local design criteria, adopted from the supplied legislative documentation, were implemented as well.

Methodology has been tested on Heemstede network operated by Waternet company. Heemstede served as a pilot network, convenient for the development of the methodology due to its relatively small size. Results presented in this research are very encouraging regarding the potential use of DeNSE method as a decision support tool in the future. Method ensures that set of feasible solutions will be identified, while enabling user to give an important heuristic input based on engineering experience (e.g. which pipes must be excluded from sectorization due to any reason). Provision of set of solutions, high computational efficiency of the method and visualization of results through satellite-based imagery (Google Earth files) enable decision makers to delicately investigate several potential solutions and select the one best suitable to their preferences. Furthermore, provided EPANET files can be used to finely tune the selected solution and investigate different hydraulic scenarios to prove validity of the solution.

It should be noted that all results presented here assume that supplied mathematical model is calibrated. There is a great amount of uncertainty regarding this assumption, as some inconsistencies and doubtful data were found on initial investigation of the mathematical model. It is recommended to thoroughly check mathematical model and whether it is properly calibrated.

DeNSE sectorization method can be used to create hierarchical partitioning of the WDN. Initially defined DMAs can be larger than size recommended by different guidelines. In the following stages, as the knowledge of the system increases and more reliable data is obtained, originally established DMAs can be partitioned further. This increase in sectorization resolution implies more delicate network interventions that could possibly significantly affect water supply, water quality and overall system reliability. DeNSE method could be successfully applied for this task as well but should probably include additional sectorization criteria such as design for fire flows.

6 REFERENCES

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*. Palmer Environmental Ld.

Farley, M. (2001). *Leakage Management and Control: A Best Practice Training Manual*. 1–169. 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

Gabow, H. N. (2000). Path-Based Depth-first Search for Strong and Biconnected Components 1 Introduction 2 Strong Components. *Information Processing Letters*, *74*, 107–114.

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

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

Perelman, L., & Ostfeld, A. (2012). Water-Distribution Systems Simplifications through Clustering. In *Journal of Water Resources Planning and Management* (Vol. 138, Issue 3). https://doi.org/10.1061/(ASCE)WR.1943-5452.0000173

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

Vasilic, Z., Stanic, M., Kapelan, Z., Prodanovic, D., & Babic, B. (2020). Uniformity and Heuristics-Based DeNSE Method for Sectorization of Water Distribution Networks. *Journal of Water Resources Planning and Management*, *146*(3). https://doi.org/10.1061/(ASCE)WR.1943-5452.0001163

Vasilić, Ž., Stanić, M., Prodanović, D., & Kapelan, Z. (2016). Network Sectorisation Through Aggregation of Strong Connected Components. *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*.