To be cited as: Ferrari, G., Savic, D., and Becciu, G. (2014). "Graph-Theoretic Approach and Sound Engineering Principles for Design of District Metered Areas." J. Water Resour. Plann. Manage., 140(12), doi: 10.1061/(ASCE)WR.1943-5452.0000424.

A GRAPH THEORETIC APPROACH AND SOUND ENGINEERING PRINCIPLES FOR DESIGN OF DISTRICT METERED AREAS

Giada Ferrari[1], Dragan Savic[2], Gianfranco Becciu[3]

ABSTRACT

The design of District Metered Areas (*DMA*s) in existing water distribution networks, especially in urban areas, involves a high number of decision variables and the effects of implementing them in districts have to be evaluated, in order not to affect the quality of the service to customers.

A new methodology for designing a given number of districts in looped water distribution networks, is proposed here. It is based on graph theory and takes into account some important *DMA* design criteria: the maximum and minimum size recommended for a district, the connectedness of each district to the water supply source and the absence of links between the districts: therefore it allows the creation of *DMA*s that are independent one from another.

A recursive bisection procedure has been applied to create districts, while an algorithm for graph traversal has been used to verify whether each district can be reached from the water source and connectivity between the nodes.

The successful application of the proposed methodology to a case study has proved its effectiveness for District Metered Areas design in real urban water distribution networks.

---

[1] PhD, Dip. Ingegneria Civile e Ambientale, sez. Scienze e Ingegneria dell'Acqua, Politecnico di Milano, p.za Leonardo da Vinci 32, 20133 Milano, Italy. E-mail: G.Ferrari@exeter.ac.uk

[2] Professor, Centre for Water Systems, College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF, UK. E-mail: D.Savic@exeter.ac.uk

[3] Professor, Dip. Ingegneria Civile e Ambientale, sez. Scienze e Ingegneria dell'Acqua, Politecnico di Milano, p.za Leonardo da Vinci 32, 20133 Milano, Italy. E-mail: gianfranco.becciu@polimi.it

INTRODUCTION

Water distribution networks, especially in urban areas, are usually designed as looped systems, where water can flow in different directions and demand nodes can be served through different paths. In such systems multiple flow paths prevent stagnation from occurring, allow water pressure to be uniform, and provide a high level of reliability to the supply service, because the demand nodes can be served even in the case of pipe failures. For these reasons, looped distribution networks are preferred for urban water supply.

However, sometimes it may be convenient to divide the network into district metered areas (*DMA*s), that are areas in the water network, independent from each other, created by the closure of valves or disconnection of pipes and where the inlet and outlet flow is metered (Figure 1). As a matter of fact, experience has shown that the implementation of district metered areas provides a series of benefits for water distribution systems: for example it allows for the reduction of leakages, due to easier and faster identification and location of leaks, and for the creation of a permanent pressure control system, which enables a low level of leakage to be maintained (AWWA 2003). It also improves the water distribution network management thanks to the simplified evaluation of the water balance and reduces water security risks, since the potential movement of contaminants throughout the system is minimized.

Since *DMA*s have been introduced in water distribution systems, considerable work has been done to improve the design, planning and management of *DMA*s: indeed, a number of studies and technical reports containing guidelines, design criteria and recommendations are available in literature (see e.g. Farley 1985, Morrison et al. 2007, Baker 2009). For example, indications are given about the minimum and maximum number of customer connections that a district should contain (*DMA*'s size); the main transmission system should be kept separated from the *DMA*s in order to ensure a flexible and reliable water supply; each

DMA should be connected directly with the transmission main and be independent, i.e. without any connection with other *DMA*s; other factors to take into account are pressure constraints at demand nodes, final leakage level target, implementation and maintenance costs.

The re-design of a water distribution network into *DMA*s is not a trivial issue and if not undertaken with care, can lead to supply problems, reduction of reliability and worsening water quality (Grayman et al. 2009). Due to its complexity, the design of *DMA*s in practice has always been made empirically: an initial division into districts, each having appropriate size, is considered; then hydraulic simulations are performed under different demand scenarios (average demand, maximum demand, fire flow, pipe burst) and modifications are performed manually in the water network if pressure constraints are not verified. In other words, an iterative trial and error approach is adopted until an acceptable solution is found.

Murray et al. (2010) compared the performances of looped water distribution networks with the corresponding districted version. They showed how, if the criteria mentioned above are followed, the creation of *DMA*s brings advantages in terms of water security and leakage reduction, without compromising either the reliability or the quality of the water supplied. Performance diminishes when connections between *DMA*s cannot be avoided and water from some districts flows into downstream *DMA*s, emphasizing the importance of the independence between them.

Recently, some new methods for designing *DMA*s in existing water distribution networks have been developed. Award et al. (2009) used genetic algorithms for individuating the optimal setting of Pressure Reducing Valves (*PRV*s) in water distribution networks. The methodology also involved finding the optimal *DMA* boundaries such that the excessive outlet hydraulic pressure at nodes can be minimised throughout the day. They adopted a fitness function that assesses the annual benefits deriving from the implemented pressure management scheme, and therefore the cost saving. The water network is considered as a graph and the creation of individual solutions in initial population is made by applying the Depth First Search algorithm (Cormen et al. 2001). A spanning tree is grown from the water source, then a certain number of PRVs are placed on tree branches at random and the remaining links are reinstated. This procedure ensures the connectivity and reachability of each demand node.

Di Nardo and Di Natale (2009) developed a decision support system for *DMA* design using a graph theory approach based on a shortest path algorithm (Dijkstra 1959). Their method is based on the determination of a set of candidate pipes to be closed in order to form the boundaries of the districts. The minimum head loss paths from the water source to each demand node in the network are found and each pipe is associated with a frequency value, which is proportional to the number of times that the pipe is found in a path. Candidate pipes are those having a frequency lower than a certain threshold fixed by the decision maker.

The same authors (Di Nardo and Di Natale 2011) applied the graph partitioning methodology by Karapys and Kumar (1995) to the *DMA*s design problem. They achieved the division of network nodes into a certain number of districts (partition), approximately equally sized, so that the number of links to be closed, also called cutting edges, was minimized. The methodology required that weights related to hydraulic properties were assigned to links and vertices of the graph. Pipe flow, dissipated power or diameters were chosen as hydraulic properties assigned to the links while water demand was associated to the vertices. The choice was made by analyzing all the possible combinations among the boundary pipes found with the graph partitioning algorithm.

Alvisi and Franchini (2012) proposed an automatic procedure based on the generation of different *DMA*s network layouts that are further compared in order to find the most resilient one. The creation of a solution, i.e. a possible *DMA* layout, is made with the aid of the Breadth First Search algorithm (Pohl 1989): from each node of the network a tree is made to grow and all nodes are grouped in sets, according to their distance from the source node, and the cumulative water demand is evaluated. Districts are defined as groups of sets having total water demand within the recommended limits.

Di Nardo et al. (2013) recently developed a methodology for water network sectorization that allows to divide the water network into isolated areas (sectors) that are completely separated from one another and fed by their own water source. The authors employed graph theory to model the water system and applied a heuristic optimisation technique aimed to minimise the dissipated power in order to determine the sectors' boundaries.

Lastly, Diao et al. (2013) determined DMAs boundaries through an approach based on the assumption that water distribution systems can be seen as 'community structures'. A 'community' is defined as a group of nodes characterised by a high density of edges between its vertices, whereas the number of edges connecting a community with another is significantly lower. The authors identified communities in a water distribution network starting from a condition in which each single vertex is a community and adopting a greedy strategy, joining communities until an increase in the modularity, an indicator of how well a graph is divided into communities, is observed. The size of the communities was then calculated and those exceeding the upper limit for DMA size were further decomposed, in order to ensure that each community had a suitable number of customer connections. Finally, feed lines, i.e. pipes connecting communities, and pipes to be closed were identified such that the minimum pressure requirements at each node are verified.

The aforementioned studies are based on graph theory principles and techniques. This means that the water distribution network is handled as a graph, whose vertices are represented by demand nodes, reservoirs and tanks, while edges are represented by pipes, pumps, and any other link between two vertices in the water system (Deuerlein 2008, Kesavan 1972). As a matter of fact, graphs have been proved to represent a useful and powerful tool for modelling a water distribution network, especially when dealing with the *DMA*'s design problem (Ostfeld and Shamir 1996, Ostfeld 2005, Tzatchkov 2006). Furthermore, graphs are quite simple structures, so the complexity of the problems can be reduced significantly. Of particular interest, regarding the design of districts in a water distribution network, are the properties of connectivity among nodes within a DMA and reachability of each node from the water supply source. The former indicates that all the nodes belonging to the same district are connected, while the latter that all the nodes in the network have a direct path either to the water source, or to the main transmission system.

However, water distribution networks are not just graphs composed by arcs and vertices. There are functional requirements that need to be satisfied, such as the delivery of a certain amount of water to each node and the maintenance of adequate pressures. Moreover, flow directions depend on hydraulic laws, in particular on the continuity equation at nodes and the energy conservation around each elementary loop. Therefore, when using graph theory for water distribution network analysis, the hydraulic and connectivity constraints have to be taken into account. The former ensures the physics of the water flow is taken into account, and the latter makes sure that no nodes are isolated, and thus all the customers in the water supply distribution networks are properly served.

As mentioned above, when designing *DMA*s in an existing water distribution network some criteria have to be followed, in order to not lower the performance of the water supply system. The essential factors to take into account are the size limits of the *DMA*s, the connectivity properties and respecting the minimum pressure constrains. Nevertheless, the approaches developed in literature, although able to determine *DMA*'s boundaries, do not take into account all the aforementioned factors in the design process. Either the methodology by Award et al. (2009) or the one by Di Nardo and Di Natale (2009) consider the size of the resulting *DMA*s as a design criterion: the first aims only to minimise the excess of pressure at nodes, and the second chooses the pipes to be closed only on the basis of minimum head loss paths from the water source to each demand node. Moreover, the majority of the approaches illustrated cannot ensure the resulting *DMA*s be supplied independently. Non-independent *DMA*s and *DMA*s which lack a direct connection to the transmission mains are likely to occur, despite the shortcomings in terms of reliability and water quality this would produce. Only the methodology developed by Di Nardo et al. (2013) allows to identify DMAs that are independent, but the authors did not consider the recommendations about the number of customer connections per district as a design criterion.

This paper presents a methodology based on graph theory that can partition an existing water distribution network in *DMA*s. Its purpose is to determine the boundaries of *DMA*s ensuring that each *DMA* be between the recommended size limits, and be directly connected to the main transmission system. The methodology will also ensure that no links exist between different *DMA*s, thus there is no flow exchange between adjacent *DMA*s, and all the demand nodes are adequately served during the simulation period. This paper will also show how it is able to divide a water network into a certain number of *DMA*s which are spaced at specified intervals and with the important characteristic of being independent of one another.

THE METHODOLOGY

The methodology developed, represented in Figure 2, consists of three main steps (corresponding to the square boxes in the flow diagram):

- A preliminary analysis of the water distribution network, to identify the transmission mains, the "independent" districts (see Section 1 below) and the number of *DMA*s to create, k.
- A recursive bisection algorithm, which has been tailored to take into account the design criteria described in the previous section. It determines the *DMA*'s boundaries, where valves are assumed to be available, and therefore the pipes that have to be closed. If, in practice, valves were not available at the boundaries, the expenditure for their installation would need to be taken into account in the total cost of *DMA*s implementation.
- A hydraulic simulation to verify that minimum pressure requirements are satisfied at each node.

This section aims to provide a detailed analysis of the steps listed above, along with a description of the output.

1. Input data

The proposed methodology requires the following input data:

- the model of the water distribution network, in particular the topology, i.e. the spatial coordinates of nodes and links between them, the hydraulic characteristics of the network, among which nodal base demands, demand patterns, head loss formula, pipes roughness, etc., and the hydraulic analysis method, for instance the gradient method (Todini and Pilati 1987)
- minimum required pressure to ensure the delivery of water to the customers, to verify the hydraulic feasibility of the resulting network
- the minimum and maximum allowable size for a single *DMA*, C_{min} and C_{MAX}, usually expressed in terms of the number of customer connections per district and typically equal to 500 and 5000 respectively (Farley 1985, Morrison et al. 2007). In those cases

when the information about the number of connections at each node is not available, the minimum and maximum size of a *DMA* can be expressed in terms of total water demand within the *DMA*.

An average relationship between the number of connections and water demand is adopted: let C_{tot} be the total number of customer connections in the network and W_{tot} the total average water demand (l/s); $W_{d,min}$ and $W_{d,MAX}$, respectively the minimum and maximum average water demand of a generic independent district, represent the minimum and maximum size per district and are defined as follows:

$$W_{d,min} = \frac{W_{tot}}{C_{tot}} C_{MAX} \qquad \text{Eq. 1a}$$

$$W_{d,MAX} = \frac{W_{tot}}{C_{tot}} C_{min} \qquad \text{Eq. 1b}$$

The proposed methodology allows only for the creation of *DMA*s whose sum of nodal water demands is in between the boundary values given in Equations 1a and 1b.

2. Preliminary analysis

A preliminary analysis of the water distribution network to be divided into *DMA*s is necessary, in order to acquire some basic information, fundamental for performing the further steps of the methodology. At first, water transmission mains have to be identified: they are large water pipes used to extend and convey water between sources, such as storage facilities, facilities, external water supply networks, wells, springs, etc., and the distribution mains, and are not intended to serve as a distribution system. The distinction made in the definition of transmission mains and distribution mains is a function of the size of the water system itself. Thus, this distinction is not universal and varies from system to system, and the actual function of the mains in the system is more important than the size in determining what category a main falls under. According to usual practice, transmission mains can be considered as the series of connected pipes having a diameter equal to or greater than a threshold of approximately 300 - 350 mm (12 - 14 inches).

When the transmission mains have been identified, the network is analysed in order to find the "independent districts", i.e., groups of nodes linked to the transmission main and having no connections with any other group (an example of independent districts is represented in Figure 3). This analysis is performed by exploring the water distribution network with the Breadth First Search (BFS) algorithm. All the independent districts are automatically detected by considering every node that belongs to the transmission mains as sources from which to start the exploration of adjacent nodes. From each of them, the algorithm, in its first stages, detects all the nodes that are at the distance of one edge from the source and store them in a list. In the second stage, all the nodes placed at a distance of two edge from the source are found and added to the list. The BFS algorithm terminates when there are no more nodes reachable from the source. These steps are performed for each node of the transmission mains system. The order in which nodes are explored does not influence the result of the algorithm, i.e. the list of nodes reachable form each source. Some of these lists might contain the same nodes, because it can happen that a group of nodes can be reached from more than one source. This group, having a number of links with the transmission mains equal to the number of sources from which its nodes can be explored, represents an independent district.

With the aid of BFS algorithm, all the independent groups of nodes are easily and quickly identified, along with their connections with the main transmission system; their size in terms of total nodal water demand is also calculated.

The dimensions and the characteristics of these groups of nodes can vary considerably and depend on the water distribution systems being analysed.

Specifically, the size, i.e. the number of customer connections C_d of each independent district, can be:

- lower than the minimum acceptable for a single *DMA* ($C_d < C_{min}$),
- between the minimum and the maximum ($C_{min} < C_d < C_{MAX}$),
- greater than the maximum ($C_d > C_{MAX}$).

A similar distinction can be made with respect to the water usage, as seen in equation 1.

Such a characterization allows for the location of *DMA*s that might already exist in the water distribution network, that are groups of connected nodes having size between the limits, that are 500 and 5000 customer connections in this study.

In the first case ($C_d < C_{min}$), the independent district is too small to be considered a *DMA*. Therefore, these nodes keep being supplied directly from the transmission mains, no meters are installed and no modifications are made to the network. Conversely, if the size is between the boundary values C_{MAX} and C_{min} given as input, the independent district is considered as a *DMA*. Therefore during the preliminary analysis some *DMA*s might be identified. Generally, it is not necessary to insert isolating valves, because there are no connections with any other district (i.e., the *DMA* is already isolated from the rest of the network), and a meter is installed on the feeding pipe that link the district with the transmission main. However, if the district is fed from multiple sources, further analyses need to be carried out to determine whether it is convenient to set valves on one or more feeding pipes. Lastly, if the size of the independent district is greater than the maximum allowable for a single *DMA,* it needs to be divided into smaller districts.

The preliminary analysis ends with determining the number *k* of *DMA*s to create in each independent district belonging to the third group ($C_d > C_{MAX}$). This number is chosen between a minimum and maximum value, depending on the minimum and maximum size allowable. Nevertheless, the number of *DMA*s that can be created within an independent district, cannot exceed the number of pipes linking the district itself with the transmission main, because the methodology aims to create *DMA*s all having a direct path to the water source, as recommended by the design criteria. In the case there were only few connections between the independent district and the transmission mains, an appropriate number of linking pipes should be added to the existing network, in order to allow the creation of a certain number of *DMA*s. However, this situation is very unlikely to occur, because the supply of an independent district having such a high number of customer connections could hardly come from a limited number of feeding pipes from the transmission main system.

Therefore, if n_c is the number of connections with the transmission main, the minimum and maximum number of district are given by equations 2 and 3:

$$k_{min} = \frac{w_d}{w_{d,MAX}} \qquad \text{Eq. 2}$$

$$k_{MAX} = \min\left(\frac{w_d}{w_{d,MAX}}, n_c\right) \qquad \text{Eq. 3}$$

Since small *DMA*s have high implementation and maintenance costs whilst large *DMA*s do not provide as much benefit in terms of water network control and management, *k* should be chosen close to the mean value.

3. The recursive procedure

Given a certain value of *k*, the *DMA*'s boundaries are defined through the application of a recursive bisection algorithm. It is based on the principle of recursively bisecting the set of demand nodes until the required number of subsets is obtained. The procedure for performing a single bisection, illustrated in the following section, is crucial to achieve the desired result, i.e., a feasible water distribution network that is divided into a certain number of subsets (*DMA*s) having characteristics that reflect the design criteria. The diagram in Figure 4 illustrates the recursive procedure.

Let SET={n_1, n_2, n_3, ..., n_n} be the set of nodes in a graph that are required to be divided into a certain number of subsets n_{SS}, equal to or greater than 2. When the graph is a water distribution network, SET corresponds to the set of demand nodes, the subsets are the *DMA*s and n_{SS} is equal to *k*, set by the decision maker between the minimum and the maximum value previously evaluated. The first bisection produces two subsets A and B. The number of subsets to be further created from A and B is given by rounding down half of n_{SS} and rounding up half of n_{SS} respectively (equations 4 below).

$$n_{SS,A} = floor\left(\frac{n_{SS}}{2}\right)$$
$$n_{SS,B} = ceil\left(\frac{n_{SS}}{2}\right) \qquad \text{Eq. 4}$$

Since $n_{ss,A}$ and $n_{ss,B}$ can have differing values, the size of A and B can be different as well: for instance, if the number of subsets to be created from A is higher, the size of A has to be greater than the size of B. The size of a subset, for water distribution networks, is represented either by the total number of customer connections within the subset or by the water use within the subset. Since the number of customer connections per district is one of the factors to be taken into account in the *DMA*'s design, the size of a subset has to be defined appropriately. A correct definition of the subset's size leads to *DMA*s whose number of customer connections satisfies the recommended design criteria.

Let S_A and S_B be the sizes of the first two subsets, respectively A and B; their values have to be between the boundaries given in equations 5.

$$S_{A/B,min} = n_{ss,A/B}\, W_{d,min}$$
$$S_{A/B,MAX} = n_{ss,A/B}\, W_{d,MAX} \qquad \text{Eq. 5}$$

The definition of the subset's size as in these equations has two main purposes: dividing the network into a number of districts, *k*, having roughly the same size and following the guidelines given in literature. However, it is not recommended that the subset size is too close to the boundary values, especially if $n_{ss,A}$ and $n_{ss,B}$ are greater than two. In fact, a size value very close to the boundary could lead to final *DMA*s having vastly different dimensions, implying a higher number of valves to be closed and thus a higher cost. An acceptability range for the size of subsets A and B is then defined (Figure 5) and its boundary values are given by equations 6 – 8:

$$S_{A/B,low} = al\ n_{ss,A/B}\left(W_{A/B,med} - W_{d,min}\right) \qquad \text{Eq. 6}$$

$$S_{A/B,up} = al\ n_{ss,A/B}\left(W_{d,MAX} - W_{A/B,med}\right) \qquad \text{Eq. 7}$$

$$W_{A/B,med} = n_{ss,A/B}\frac{\sum_{i=1}^{n} w_i}{k} \qquad \text{Eq. 8}$$

where *al* is a parameter between 0 and 1 that indicates the width of the interval around the medium water usage value and allows for flexibility in the *DMA*'s size. The choice of *al* is arbitrary and different values can be set for obtaining different *DMA* layouts and to check the sensitivity of the result with respect to the parameter. If *al* is set equal to 1, the upper and lower bounds for the subset size are respectively the maximum and minimum allowable. As previously mentioned, the adoption of a restricted admissible interval for subset size proves to be particularly useful when *k* is high (greater than 5), since this provides for approximately equally sized districts.

After the first bisection has been performed and subsets A and B have been determined, if $n_{ss,A}$ is equal to or greater than 2, A becomes the new set to be bisected and the previous steps are repeated until $n_{ss,A}$ is lower than 2. The same is done with the subset B.

The steps of the recursive bisection, illustrated in Figure 4, are summarized below:

1. Given SET={$n_1, n_2, n_3, ..., n_n$} and $n_{ss} = k$;
2. Evaluate $n_{ss,A}$ and $n_{ss,B}$ with equations 4;
3. Bisect SET into two subsets A and B such that their size S_A and S_B are between $S_{A,low}$ and $S_{A,up}$ and $S_{B,low}$ and $S_{B,up}$ respectively
4. a) If $n_{ss,A} \geq 2$ update SET = A , $n_{ss} = n_{ss,A}$ and go back to step 2)

b) If $n_{ss,B} \geq 2$, update SET =B, $n_{ss}= n_{ss,B}$ and go back to step 2)

5. Stop when both $n_{ss,A}$ and $n_{ss,B}$ are lower than 2.

4. The bisection algorithm

The bisection algorithm presented here is the key part of the whole methodology. By its recursive application, it allows for the creation of districts having appropriate size, i.e. the total water demand is between the limits given in equation 1, and that are independent of each other. This independence ensures that each *DMA* is connected to the transmission mains by at least one pipe, so that at least one direct flow path exists between the water source and each demand node, and there are no flow exchanges between contiguous *DMA*s. A random component has been included in the process defining *DMA*s, so that each single run of the algorithm produces a different layout of the *DMA* where all the districts have at least one connection with the transmission main. This capacity to provide a number of alternative solutions is particularly useful because the decision maker is provided with a number of feasible options. In fact, when dealing with large water distribution systems, there are many possible divisions of the network into districts. The decision maker can then choose the division that best suits his requirements from the various options provided.

Figure 6 shows the bisection algorithm that is applied recursively to identify the desired number *k* of districts in the water distribution network. Figure 7 shows an example of an independent district in a real large water distribution network. Nodes belonging to SET are highlighted in grey, nodes belonging to C in black, and the transmission mains are indicated with a black bold line. The adjacency matrix, which is an n-by-n matrix that embodies the information about the connections between the nodes belonging to SET, is determined.

Let SET= $\{id_1, id_2, id_3, ..., id_n\}$ be the set containing the n nodes belonging to an independent district having size greater than C_{MAX}, i.e. the set of nodes that need to be grouped into districts; let C= $\{c_1, c_2, c_3, ..., c_{nc}\}$ be the set of the n_c nodes of SET adjacent to the transmission main.

The procedure for performing a bisection of a certain set of nodes SET, illustrated in Figure 6 in flow chart form, consists in the following steps:

1. A node $c_i \in C$ is chosen at random
2. A spanning tree is made to grow from c_i using the BFS algorithm
3. Let L= $\{c_i , n_1, ...\}$ be the list of nodes in the order they are discovered with the BFS: since all the nodes in SET are connected, they will all be discovered sooner or later, L is thus an n-dimensional array
4. Evaluate the cumulative water usage for the set L, W=$\{W_1, W_2, ... W_i, ... , W_n\}$
5. Let I be the set of nodes n_j such that $S_{A,low} < W(n_j) < S_{A,up}$ (see Figure 8)
6. Choose a node $n^I \in$ I at random
7. Let SET1 be a subset of L formed by all the nodes in the list between the tree source c_i to n^I : SET1 = $\{c_i, n_j, ..., n^I\}$
8. Let SET2 be the complement of L with respect to SET1: SET2 = SET - SET1
9. Consider the set C'= C $\subset$ SET2, containing the nodes of SET2 adjacent to transmission main
10. Grow a tree T_k from each $c_k \in$ C' and evaluate their corresponding water usages W_k
11. a) If there is a W_k such that $S_{B,min} < W_k < S_{B,MAX}$, set \$ B =T_{k}, A = SET - T_{k}, otherwise

    b1) update set I: I=I – $\{n^I\}$; if I is not empty, go back to 6), otherwise

    b2) C = C – $\{c_i\}$ and go back to 1)
12. Evaluate the number of connections that A and B have with the transmission mains: $n_{C,A}$=C$\cap$A, $n_{C,B}$=C $\cap$B

13. If $n_{C,A} \geq n_{ss,A}$ and $n_{C,B} \geq n_{ss,B}$, A and B are the result of the bisection, otherwise go back to 11.b1).

The last step verifies that the number of pipes connecting A and B with the transmission main is equal to or greater than the number of subsets to be further made from A and B, because every *DMA* must have a direct path to the water source. Therefore each district that is to be created will have a direct connection with the transmission main and the independence of every *DMA* is ensured.

5. Outputs

The procedure illustrated above allows for the division of an independent district into a certain number of smaller districts that meet the size and connectedness requirements. It gives as its output the list of nodes belonging to each subset, that is the list of nodes belonging to each *DMA*.

Its application to all the independent districts provides a possible *DMA* layout for the water distribution network under examination. The integration of a random component in the bisection algorithm (steps 1 and 6) allows for each run of the procedure to generate a different solution.

The solution obtained describes where a certain number of pipes have been closed in order to create the *DMA*'s boundaries in the original/starting water distribution network. Hence, the hydraulic feasibility of the solution found needs to be checked. A hydraulic simulation is then performed using the software EPANET and the pressures at each demand node are evaluated. The last step is to verify that each demand node during the whole simulation period is associated a hydraulic head equal to or greater than the minimum required. If this constraint is not verified, then the solution is discarded as unfeasible, and a new division onto the *DMA*s is sought until one that respects the minimum pressure requirements is determined. Here, the ability of the resulting system to meet fire flow conditions has not been checked. However, the creation of DMAs has been proven to affect the capability of meeting fire flow conditions, as well as the reliability and the quality of the water delivered. Therefore, further analysis will be required to determine the best solution among those generated in relation to these three aspects.

The value of *k* chosen by the designer can also be changed in order to find a wider variety of possible solutions. A number of solutions can be found and compared on the basis of performance indicators, in order to determine which one is the best to adopt in a specific water network.

CASE STUDY

The effectiveness of the proposed methodology was proven by applying it to a real case study. The water distribution network used is a modified real world system, frequently used as a test bed for various modelling exercises including the Battle of the Water Sensor Networks competition (Ostfeld et al. 2008). Its geographic representation and the names of the components have been distorted in order to protect the identity of the system. The adjustments made are related only to the appearance and do not have any influence on the connectivity and the hydraulic behaviour of the network.

The network serves approximately 150,000 people and includes two reservoirs, two tanks, four pumps and five valves, and it represents a typical example of a looped urban distribution system. Its topology is illustrated in Figure 9.

The hydraulic analysis was performed by using the simulation software EPANET (Rossman 2000). The head losses were evaluated with the Hazen-Williams formula, while other basic characteristics of the network that has been used in the hydraulic analysis, are reported in Table 1.

The minimum and maximum number of customer connections per district were set equal to 500 and 5000 respectively. Nevertheless, the data about the number of connections per node was unknown, and the only information available was the total number of customer connections in the system, that is equal to 77916. Therefore, an average relationship between connections and water demand was used (Grayman et al. 2009): let $\overline{W}_c$ be the average water use per connection, given by the ratio between the total average water use and the number of connections in the network. The upper and lower limits for the size of a district can be expressed in terms of water usage, $W_{d,MAX}$ and $W_{d,min}$, and can be evaluated as the product between the average water use per connection and the upper and lower limit of connections (equations 1a and 1b). The resulting values are reported in Table 2, along with the variables used for their evaluation.

The first step of the preliminary analysis of the water distribution network, whose purpose is to gain detailed information about the water system, is to identify the pipes belonging to the transmission mains. In this case study, the main transmission system is considered to be composed of all the connected pipes having a diameter greater than or equal to 350 mm (14 inches). Transmission mains are shown in Figure 9. They are composed of 870 pipes, and represent 9.2% of all the pipes in the network, totalling 162.86 km in length.

The second step is to identify the independent districts in the water distribution network, that are groups of fully connected nodes having at least one link between the district and the transmission main system (that is a direct connection with the water source). The independent districts are then classified according to their size, i.e. according to the total water use, that is the sum of the water demands of nodes within the district. This step is performed by exploring the network as a graph with the Breadth First Search (BFS) algorithm: all the nodes belonging to the transmission main are eliminated from the graph, i.e. the values of the corresponding rows and columns in the adjacency matrix are set equal to zero. Their neighbour nodes are instead considered and set as sources from which to start growing a spanning tree performing BFS (but also Depth First Search can be employed). The list of nodes explored from each "source" represents an independent district, i.e. a group of connected nodes linked by at least one pipe to the transmission main.

The total water use W_d of each independent district is then evaluated: groups of nodes characterized by $W_d < W_{d,min}$, that have water use lower than 8 l/s, are too small to be considered *DMA*s, and will be ignored in the further analysis. The total number of this type of nodes is 946, and the corresponding total water demand is 125.3 l/s; they thus represent 9.3% of the total water demand of the system and their location on the map is shown in Figure 10.

Groups of nodes characterized by W_d between $W_{d,min}$ and $W_{d,MAX}$ are considered as *DMA*s, because their size respects the requirements given in Table 2: they are districts already existing in the network, that do not require any valve to be inserted to define their boundary; only a meter on the pipe linking them with the transmission main will be installed in order to measure the inlet and outlet flows. In the water system in analysis there are 20 districts of this type, they are illustrated in Figure 11 and their characteristics are summarized in Table 3. The total water demand of the nodes belonging to the these *DMA*s is equal to 514.9 l/s, which corresponds to the 34.06% of the total water demand of the system.

Finally, groups of nodes having total $W_d > W_{d,MAX}$ are identified (Figure 12): there are three of them in the water network, and their total water use is 819.9 l/s that is the 54.23% of the water demand of the entire network. Each of these three "large districts", whose characteristics are summarized in Table 4, will be divided into smaller *DMA*s applying the developed methodology; the solutions obtained will subsequently be combined in order to define a number of alternative solutions for the whole network, for which the performances will be evaluated.

The minimum and maximum number of districts that can theoretically be created in each one of the three large independent districts, depending on their water use and on the size limits reported in Table 2 are evaluated by the equations 2 and 3 (results are in Table 5).

The analysis described so far was performed on the water system under examination using a laptop having a processor of 1.74 GHz and a RAM of 4 GB and took a *Matlab* running time of roughly 24 seconds.

The recursive bisection procedure was now applied to the three large independent districts separately. The number *k* of *DMA*s to be created was fixed equal to 9, 4 and 3 for the first, the second and the third large independent districts respectively. A minimum required hydraulic head h_{req} of 20 m (30 pound per square inch) was considered. The solution obtained, which envisages the closure of 152 pipes for the creation of the boundaries of the 16 *DMA*s, is shown in Figure 13, while the size of the *DMA*s are reported in Table 6.

CONCLUSIONS

A new methodology for designing *DMA*s, based on graph theory, was proposed: it uses graph theory principles and algorithms for automatically determining the boundaries of each district. It allows the automatic creation of a feasible division of the network into the required number of districts, from the start. At the same time, the constraints on the size limits for a single district and the connectivity properties of the districts with the water source are considered in the process of creating the *DMA*s. It ensures that the *DMA*s are roughly equally sized in terms of water use, have an appropriate number of customer connections, and are independent one from another, i.e., there are no connections between the districts, and each one of them has a direct flow path to the water source. Finally, the performing of a hydraulic simulation verifies the compliance of the minimum pressure requirements.

The methodology developed represents an improvement in respect to the ones found in literature: in fact each *DMA*, besides being characterized by an adequate number of customer connections, is also supplied directly by the transmission main and does not exchange water with adjacent *DMA*s. Therefore the distribution system is more reliable, the quality of the water delivered is higher and the possible spread of contaminants is greatly reduced.

Further improvements of the methodology could include node elevation among the decision variables, in order to ensure that the difference in altitude of nodes belonging to the same *DMA* is lower than an appropriate threshold. Secondly, it would be interesting to check the fire flows of the resulting water distribution network, as it has been shown that the division into districts could cause a reduction in fire flows. However, it should be taken into account that a higher number of decision variables would imply a higher level of complexity to deal with, which might result in an increased computational effort and in a reduction of efficiency.

The methodology developed was applied to a real case study, in order to test its applicability and effectiveness in real large urban water distribution systems. Results highlighted how the methodology successfully provides for the division of the water distribution network into a predetermined number of districts that are characterized by the desired properties: appropriate size, connection with the main transmission system, hydraulic independence from each other (i.e. no flow paths are available between two districts) and hydraulic feasibility of the resulting network. Therefore, the methodology proposed has been proved to represent a useful tool for *DMA*s design in large water distribution networks.

REFERENCES

Alvisi, S., Franchini, M. (2013). "A heuristic procedure for the automatic creation of district metered areas in water distribution systems." *Urban Water Journal*, 1-23.

AWWA, (2003). "Applying worldwide BMPs in Water Loss Control." *J. Am. Water Works Association,* 95(8), 65-71.

Baker, M. (2009). "The Baker Report: Municipal water distribution system security study-recommendations for science and technology investments." *Tech. rep.,* U.S. Department of Homeland Security.

Berge, C. (1958). "Thorie des graphes et ses applications. Collection Universitaire de Mathmatiques." *Paris: II Dunod*, English Edition, Wiley 1961.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., Stein, C. (2001). "Introduction to algorithms." *MIT Press and McGraw-Hill.*

DeNeufville, R., Schaake, J., Stafford, J. (1991). "Systems analysis of water distribution networks." *Journal of Sanitary Engineering Division,* 97(6), 825-842.

Deuerlein, J. (2008)." Decomposition model of a graph water supply network graph." *Journal of Hydraulic Engineering,* 134(6), 822-832.

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.

Dijkstra, E. (1959). "A note on two problems in connections with graphs." *Numerische Mathematik,* 1, 269-271.

Di Nardo, A., Di Natale, M. (2009). "A heuristic design support methodology based on graph theory for district metering of water supply networks." *Engineering Optimization,* 43(2), 193-211.

Di Nardo, A., Di Natale, M., Santonastaso, G., Venticinque, S. (2011). "Graph partitioning for automatic sectorization of a water distribution system." *Proc. Control and Computing for the Water Industry*, Exeter, UK.

Di Nardo, A., Di Natale, M., Santonastaso, G., Tzatchkov, V., and Alcocer-Yamanaka, V. (2013). "Water Network Sectorization Based on Graph Theory and Energy Performance Indices." *Journal of Water Resources Planning and Management,* doi:10.1061/(ASCE)WR.1943-5452.0000364.

Farley, M. R. (1985). District Metering. Part 1- System Design and Installation. WRc.

Giustolisi, O., Kapelan, Z., Savic, D. (2008). "An algorithm for automatic detection of topological changes in water distribution networks." *Journal of Hydraulic Engineering,* 134(4), 435-446.

Grayman, W. M., Murray, R. E., Savic, D. A. (2009). "Effects of redesign of water distribution system for water security and water quality factors." *Proc. EWRI-ASCE World Water & Envir. Resources Congress.*

Jacobs, P., Goulter, I. C. (1989). "Optimization of redundancy in water distribution networks using graph theoretic principles." *Engineering Optimization*, 1(15), 71-82.

Karapys, G., Kumar, V. (1995). "METIS - Unstructured Graph Partitioning and Sparse Matrix Ordering System. Version 2.0." University of Minnesota, Minnesota (USA).

Kesavan, H. K., Chandrashekar, M. (1972). "Graph theoretic models for pipe network analysis." *Journal of Hydraulics Division,* 98(2), 345-363.

MacDonald, G., Yates, C. (2005). "DMA design and implementation, an American context." *Proc. IWA Specialised Conference Leakage.* IWA, Halifax, Nova Scotia, Canada.

Morrison, J., Tooms, S., Rogers, D. (2007). *DMA Management Guidance Notes.* IWA Publication.

Murray, R.E., Grayman, W.M., Savic, D.A., and Farmani, R. (2010). “Effects of DMA Redesign on Water Distribution System Performance.” *Integrating Water Systems,* Boxall & Maksimovi´c (eds). Taylor & Francis Group, London.

Ostfeld, A., Uber, J., Salomons, E., Berry, J., Hart, W., Phillips, C., Watson, J., Dorini, G., Jonkergouw, P., Kapelan, Z., di Pierro, F.,Khu, S., Savic, D., Eliades, D., Polycarpou, M., Ghimire, S., Barkdoll, B., Gueli, R., Huang, J., McBean, E., James, W., Krause, A.,Leskovec, J., Isovitsch, S., Xu, J., Guestrin, C., VanBriesen, J., Small, M., Fischbeck, P., Preis, A., Propato, M., Piller, O., Trachtman, G., Wu, Z., and Walski, T. (2008). “The battle of the water sensor networks: a design challenge for engineers and algorithms.” *Journal of Water Resources Planning and Management*, 134(6), 556-568.

Ostfeld, A. (2005). “Water distribution systems connectivity analysis.” *Journal of Water Resources Planning and Management,* 131(1), 58-66.

Ostfeld, A., Shamir, U. (1996). “Design reliable multiquality water-supply systems.” *Journal of Water Resources Planning and Management,* 122(5), 322-333.

Perelman, L., Ostfeld, A. (2011). “Topological clustering for water distribution systems analysis.” *Environmental Modelling and Software*, 26, 969-972.

Pohl, I. S. (1989). “Bi-directional and heuristic search in path problems.” Ph.D. thesis, Stanford Linear Accelereator Centre, Stanford University, Stanford, CA (USA).

Roger, D. (2005). “Reducing leakage in Jakarta, Indonesia.” *Proc. IWA Specialised Conference Leakage,* Halifax, Nova Scotia, Canada.

Rossman, L. A. (2000). “EPANET 2 - Users Manual.” USEPA.

Todini, E., Pilati, S. (1987). “A gradient method for the analysis of pipe networks.” *Proc. International Conference on Computer Applications for Water Supply and Distribution*. Leicester Polytechnic, UK.

Tzatchkov, V., Alcocer-Yamanaka, V., Ortiz, V. (2006). "Graph theory based algorithms for water distribution network sectorization projects." *Proc. 8th Annual Water Distribution Systems Analysis Symposium.* Cincinnati (USA), 323-330.

Walski, T, M., Chase, D. V., Savic, D. A. (2001). "Water distribution modeling." Haestad Methods, Inc., Waterbury, CT.

Walski, T., Brill, E. D., Gessler, J., Goulter, I. C., Jeppson, R. M., Lansey, K., Lee, H.-L., Liebman, J. C., Mays, L., Morgan, D. R., Ormsbee, L. (1987). "Battle of the network models: epilogue." *Journal of Water Resources Planning and Management,* 113(2), 191-203.

Walski, T., Chase, D. V., Savic, D., Grayman, W. M., Beckwith, S., Koelle, E. (2003). "Advanced Water Distribution Modelling and Management." Haested Press, Waterbury, CT.

FIGURES

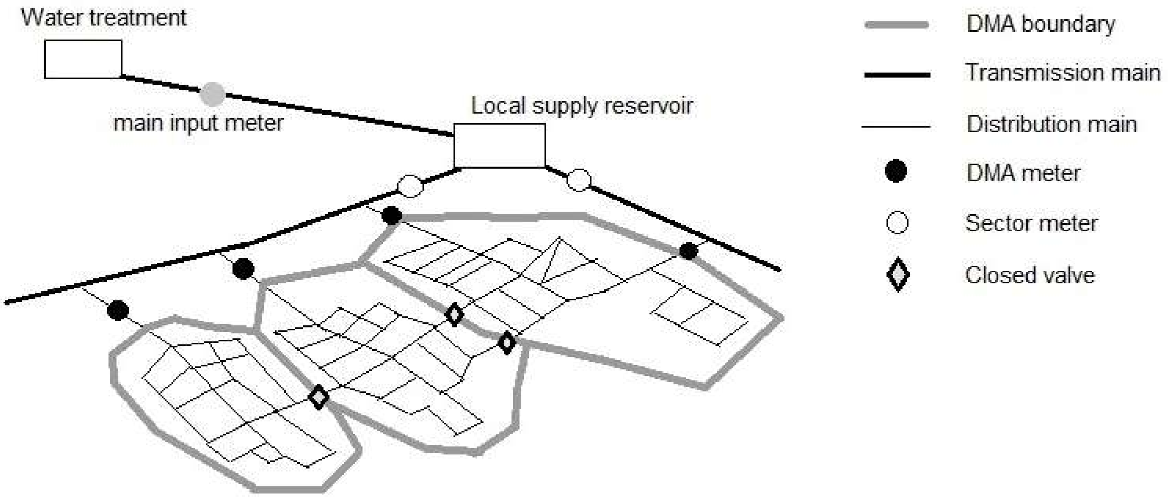


Figure 1 - Typical configuration of a water supply system with DMAs.

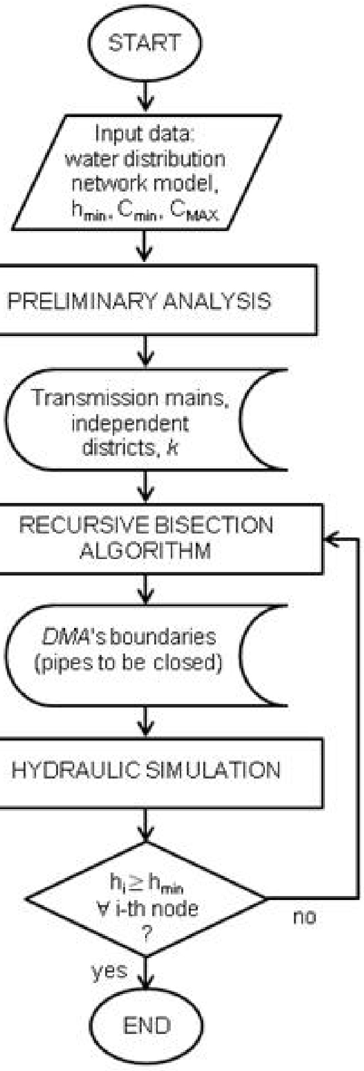


Figure 2 - The methodology proposed

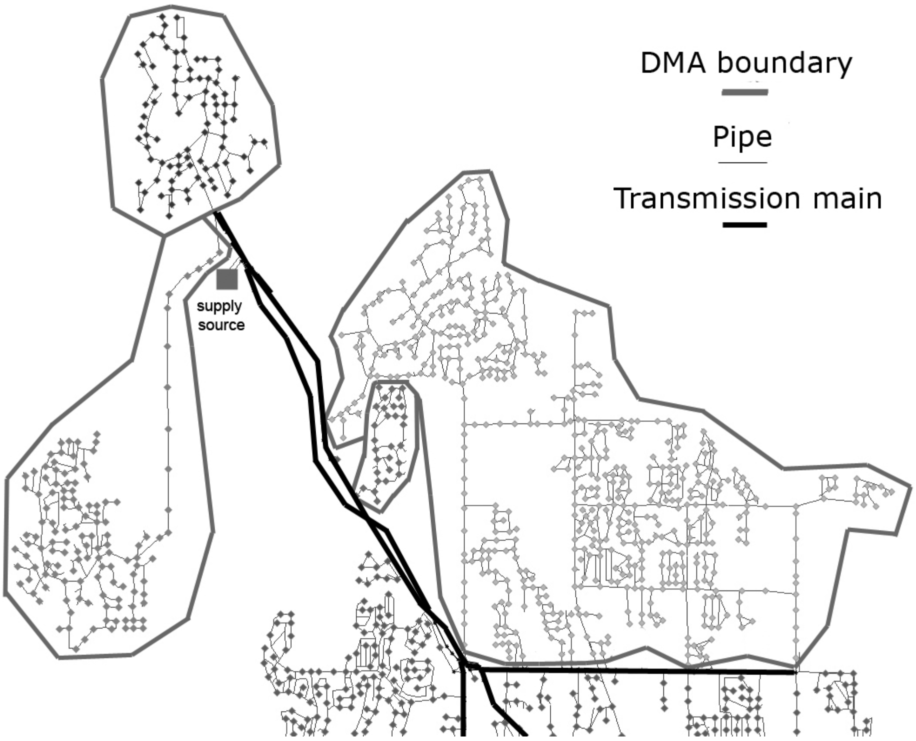


Figure 3 – Example of independent districts.

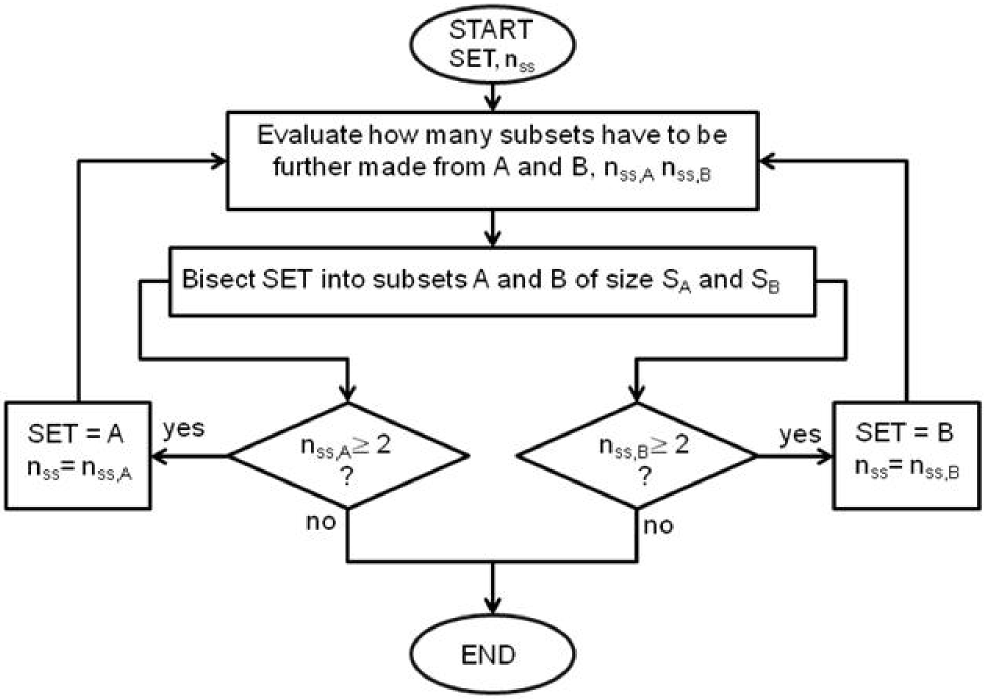


Figure 4 - The recursive bisection procedure.

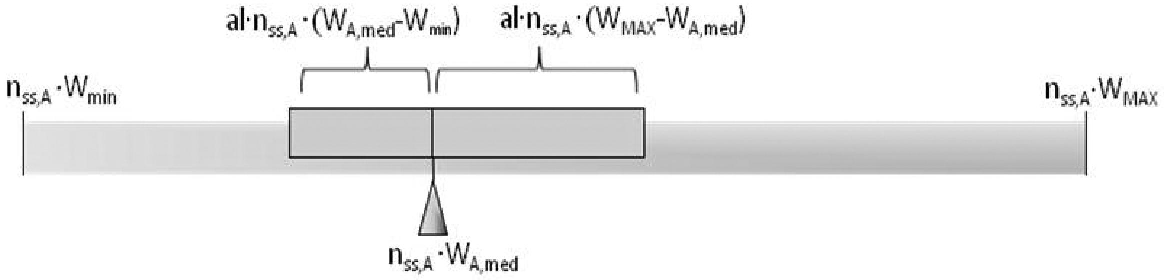


Figure 5 - Acceptability interval for subset size.

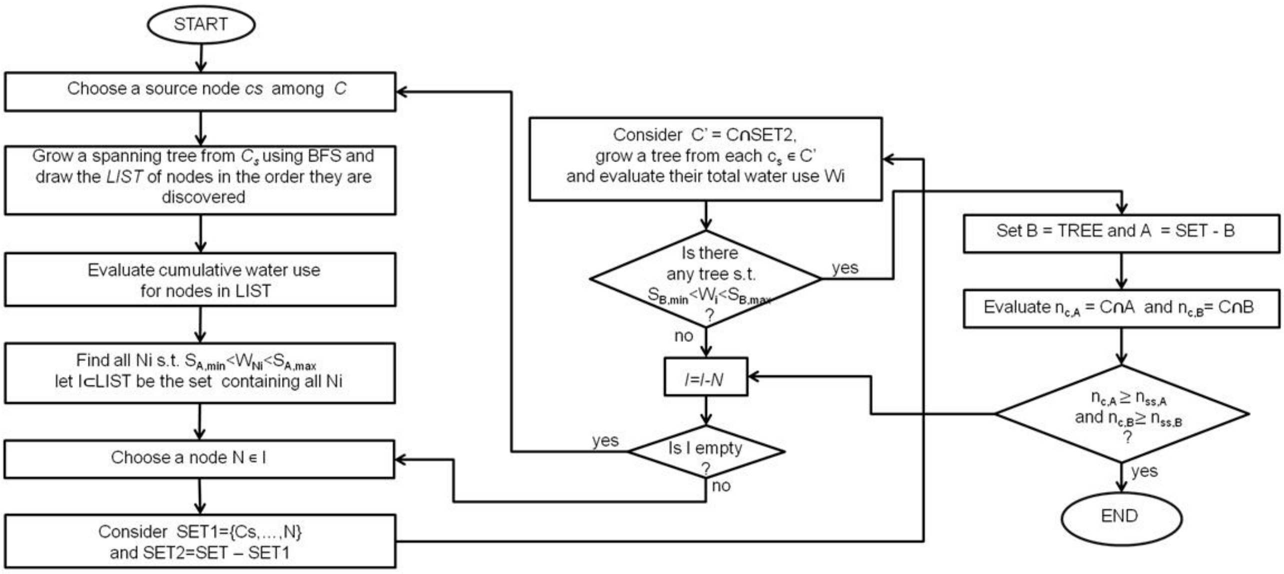
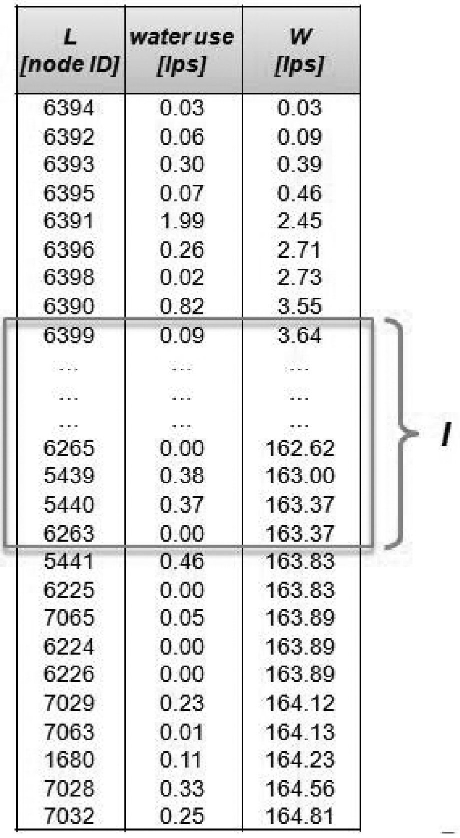


Figure 6 - Bisection algorithm flow chart

Figure 7 - Example of independent district to be divided into DMAs ($C_d > C_{MAX}$).



***L* [node ID]**	***water use* [lps]**	***W* [lps]**
6394	0.03	0.03
6392	0.06	0.09
6393	0.30	0.39
6395	0.07	0.46
6391	1.99	2.45
6396	0.26	2.71
6398	0.02	2.73
6390	0.82	3.55
6399	0.09	3.64
...	...	...
...	...	...
...	...	...
6265	0.00	162.62
5439	0.38	163.00
5440	0.37	163.37
6263	0.00	163.37
5441	0.46	163.83
6225	0.00	163.83
7065	0.05	163.89
6224	0.00	163.89
6226	0.00	163.89
7029	0.23	164.12
7063	0.01	164.13
1680	0.11	164.23
7028	0.33	164.56
7032	0.25	164.81

Figure 8 - Steps 2-5 of the bisection algorithm

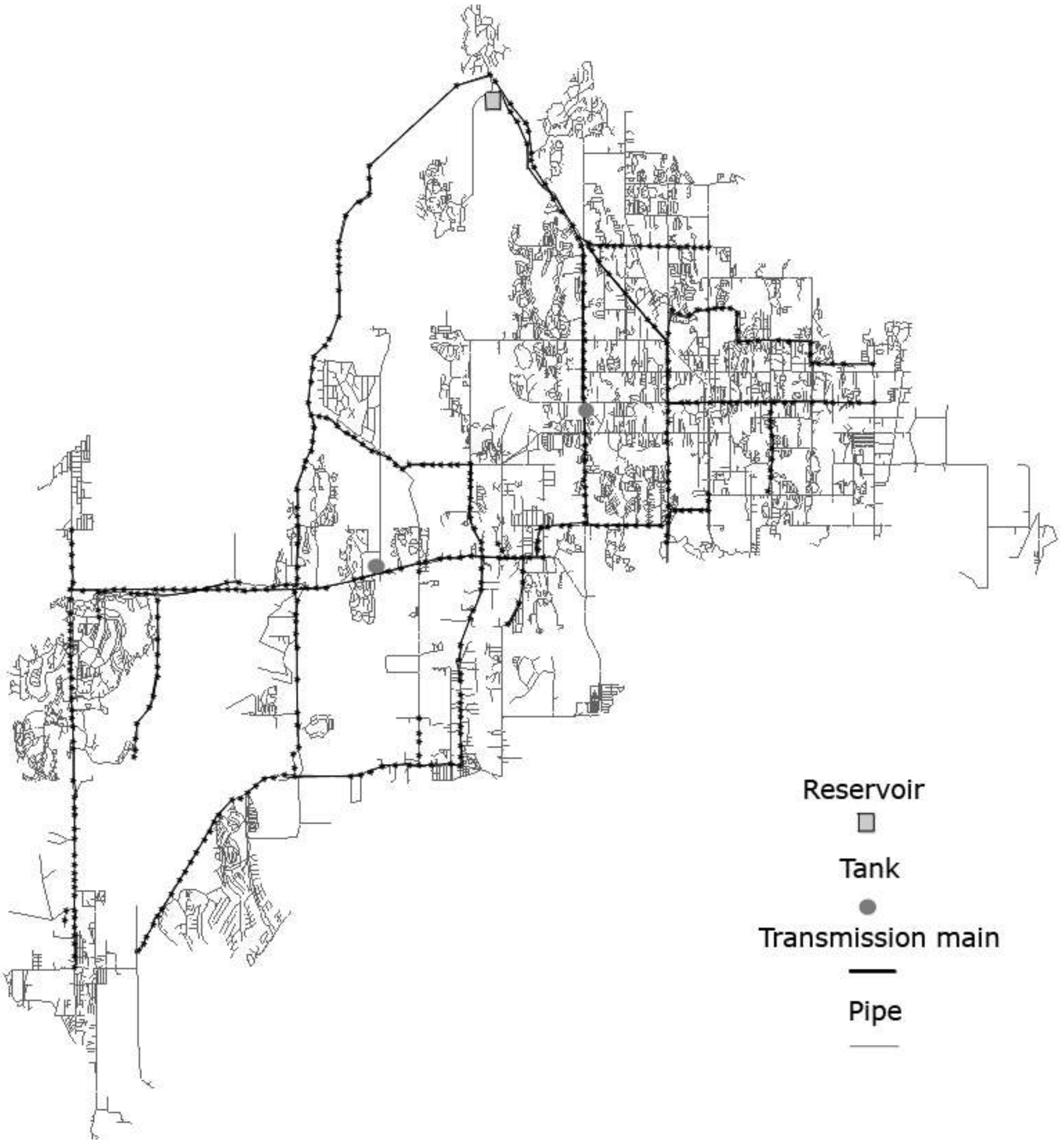


Figure 9 - The water distribution network in analysis and its transmission main system.

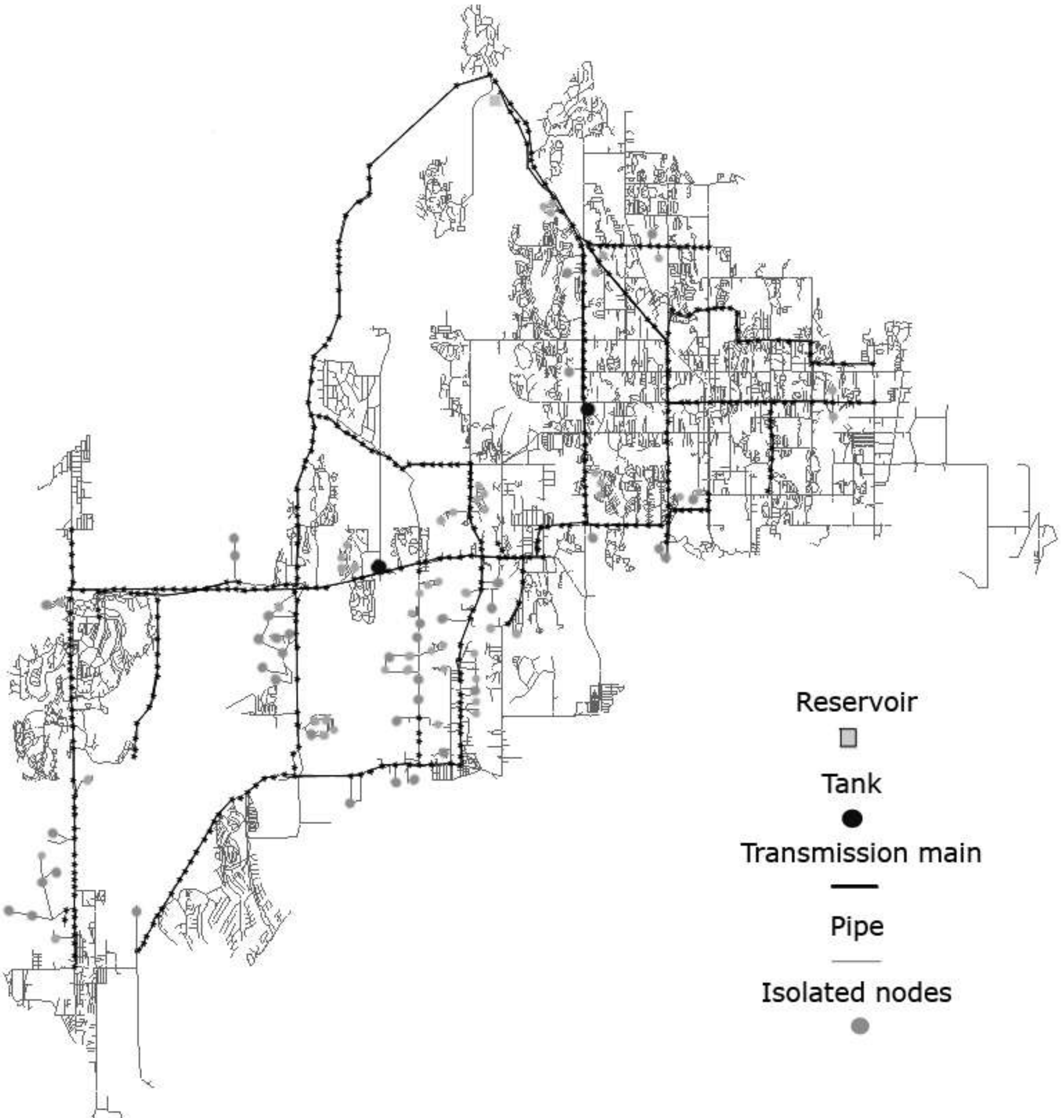


Figure 10 - Groups of nodes whose water use is lower than the minimum value for a DMA ($W_d < W_{d,min}$).

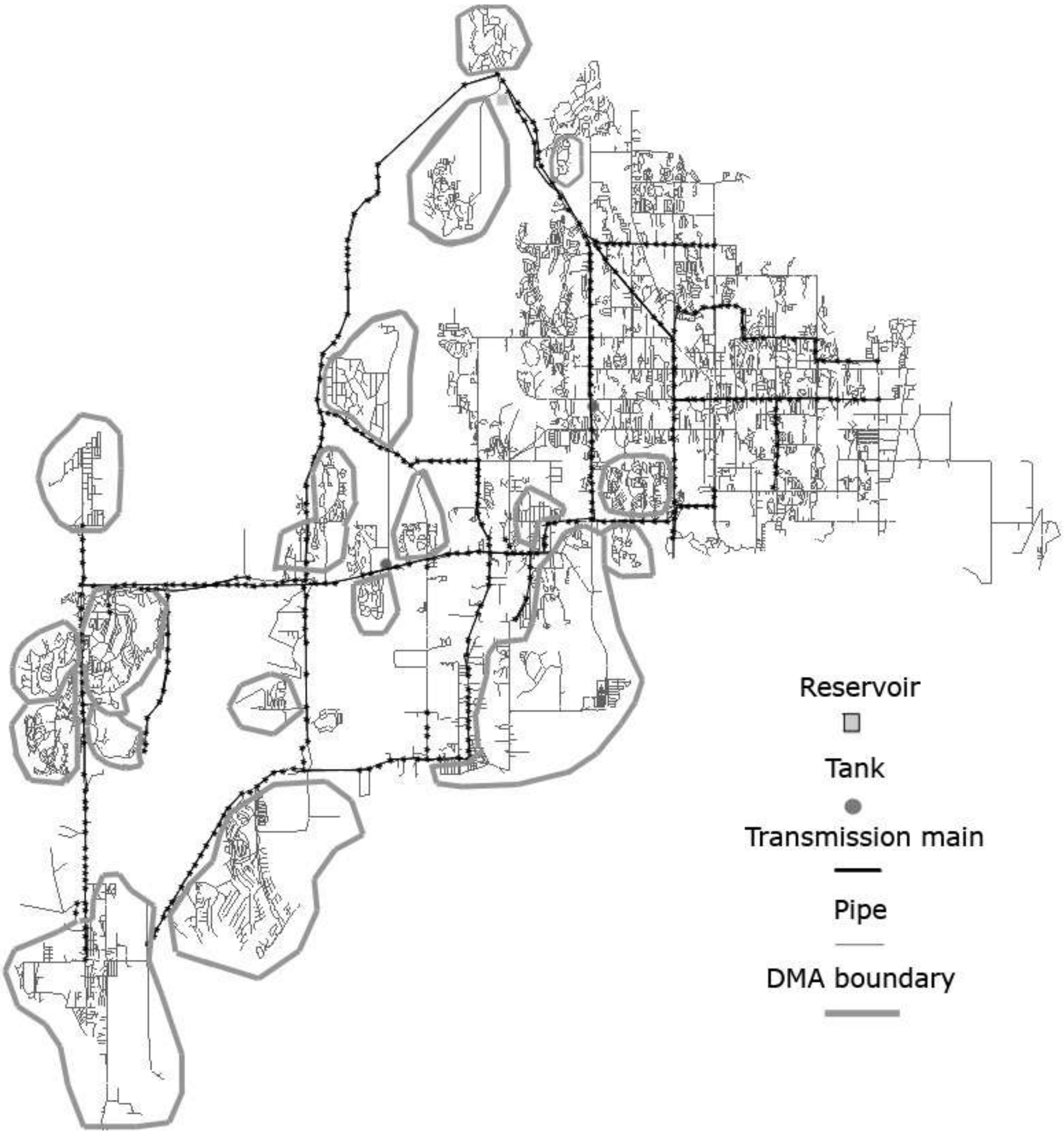


Figure 11 - Groups of nodes having total water use that is suitable for being considered a DMA ($W_{d,min} < W_d < W_{d,MAX}$)..

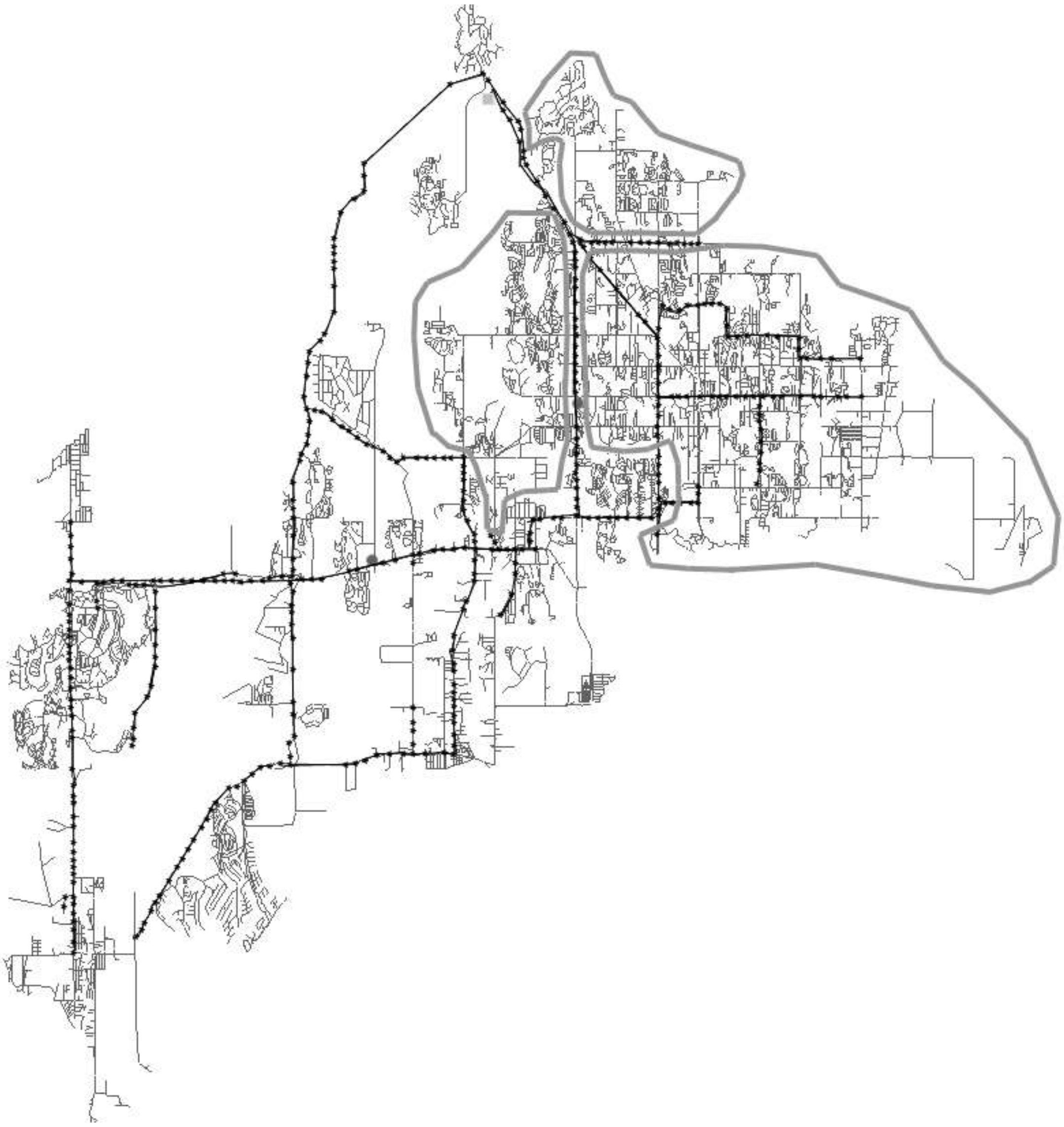


Figure 12 - Large independent districts: groups of nodes having total W_d greater than $W_{d,MAX}$

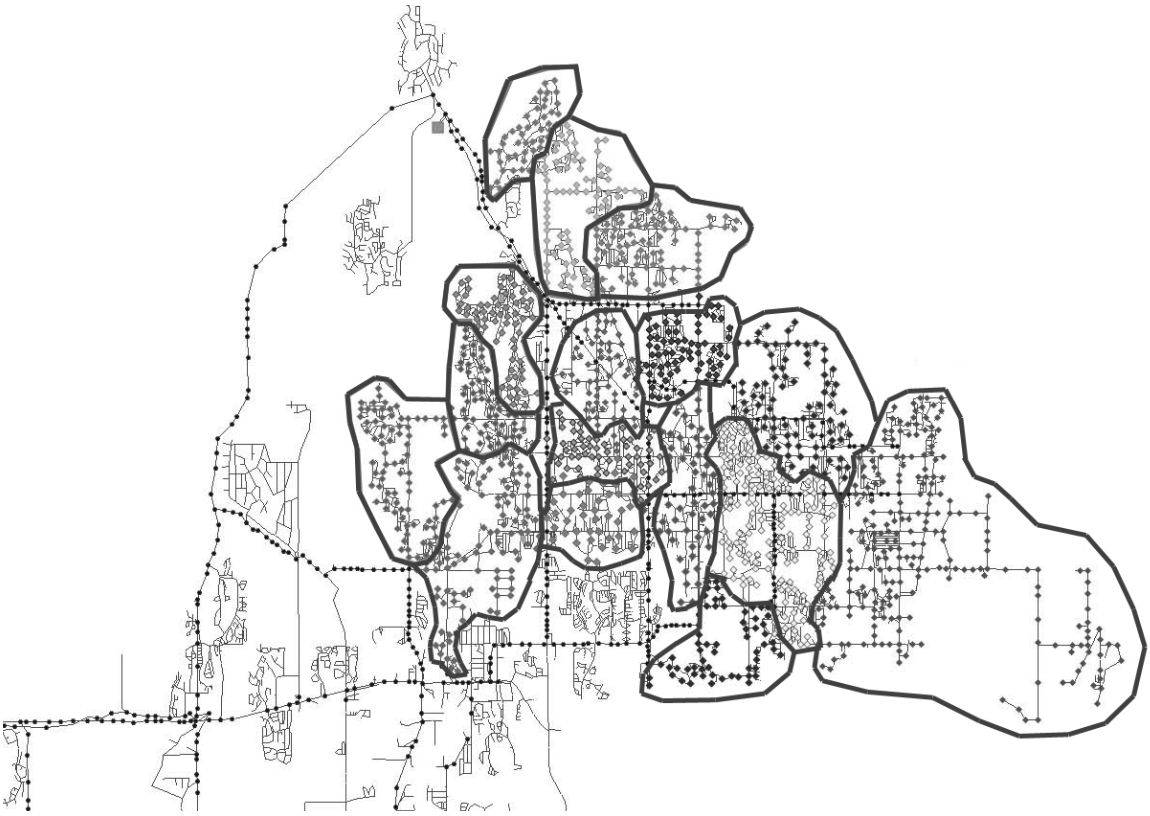


Figure 13 – Division of the independent districts into 16 *DMA*s.

TABLES

Number of Nodes	12523
Number of Links	14822
Number of Reservoirs	2
Number of Tanks	2
Number of Pumps	4
Total pipe length	1844.04 km
Total water demand	1512 l/s
Average water demand	0.121 l/s

Table 1 - Water distribution network's characteristics

Property	**Value**
C_{tot}	77916 connections
W_{tot}	1243.2 l/s
$\overline{W}_c$	0.016 l/s/connection
C_{min}	500 connections
C_{MAX}	5000 connections
$W_{d,min}$	8 l/s/DMA
$W_{d,MAX}$	80 l/s/DMA

Table 2 - Minimum and maximum size of a DMA

DMA index	**Water demand [l/s]**	**Number of nodes**	**Number of connections with the transmission main**
1	17.9	163	3
2	14.5	75	1
3	10.2	134	3
4	13.8	113	3
5	16.4	94	2
6	10.3	78	4
7	79.0	573	8
8	34.4	293	6
9	29.3	415	9
10	77.0	566	19
11	26.9	221	2
12	17.5	95	1
13	24.7	139	1
14	14.3	136	4
15	13.9	232	1
16	10.5	65	1
17	42.4	511	11
18	27.6	209	3
19	24.8	408	2
20	9.5	103	6

Table 3 - Characteristics of the existing DMAs

DMA index	**Water demand [l/s]**	**Number of nodes**	**Number of connections with the transmission main**
1	453.8	3921	73
2	201.2	1356	13
3	164.8	851	9

Table 4 - Characteristics of the large independent districts

DMA index	**k_{min}**	**k_{MAX}**
1	6	56
2	3	13
3	3	9

Table 5 - Minimum and maximum number of DMAs that can be created in the large independent district

DMA index	**Number of Nodes**	**Water use [l/s]**	**Number of connections with the transmission mains**
1	803	78.45	7
2	737	79.25	16
3	427	53.74	8
4	252	39.84	5
5	383	44.72	11
6	552	55.98	8
7	277	39.46	8
8	234	31.28	6
9	256	31.12	4
10	354	58.82	3
11	329	46.62	1
12	344	52.6	4
13	329	43.23	1
14	311	50.85	4
15	174	39.15	1
16	366	74.81	4

Table 6 - Size of the DMAs created with the methodology and number of pipes connecting each of them with the transmission mains